Proceedings of the iEMSs Fourth Biennial Meeting

Main Editors:
Miquel Sànchez-Marrè. Javier Béjar, Joaquim Comas, Andrea E. Rizzoli, Giorgio Guariso



International Congress on Environmental Modelling and Software iEMSs 2008

Integrating Sciences and Information Technology for Environmental Assessment and Decision Making

Volume 3

Universitat Politècnica de Catalunya, Barcelona, Catalonia

7-10 July 2008

Proceedings of the
iEMSs Fourth Biennial Meeting:

International Congress on Environmental Modelling and Software

iEMSs 2008

International Environmental Modelling and Software Society (iEMSs)
Universitat Politècnica de Catalunya
Barcelona, Catalonia
July 7-10 2008

SPONSORS:

ISBN: 978-84-7653-074-0



INTERNATIONAL
Environmental
Modelling
and Software
SOCIETY

International Congress on Environmental Modelling and Software
iEMSs 2008, July 7-10 2008
Universitat Politècnica de Catalunya, Barcelona, Catalonia



Impressum

Copyright: the copyright of all papers is an exclusive right of the authors. No work can be reproduced without written permission of the authors.

Each session paper in these proceedings was refereed by a member of the Editorial Board and two anonymous referees. Each workshop abstract was refereed by a member of the Editorial Board.

ISBN: 978-84-7653-074-0

Published by the International Environmental Modelling and Software Society (iEMSs)
President: Giorgio Guariso. Address: c/- IDSIA, Galleria 2, 6928 Manno, Switzerland.
E-mail: president@iemss@org. Website: http://www.iemss.org

Printed in Barcelona, Catalonia

Main Editors

Miquel Sànchez-Marrè. Javier Béjar, Joaquim Comas, Andrea E. Rizzoli, Giorgio Guariso

Editorial Board

Alexey Voinov, Andrea Emilio Rizzoli, Anthony J. Jakeman, Arno Krause, Ayalew Kassahun, Barry Croke, Bernd Page, Bert Bredeweg, Brian S. McIntosh, Carlo Giupponi, Carlo Gualtieri, Carmel A. Pollino, Chris Duffy, Claudia Turón, Dagmar Haase, Daryl Hepting, David Arctur, David Post, Dragutin T.Mihailovic, Eduardo Ayesa, Farmani Raziyeh, Geoffrey Holmes, Georgii A. Alexandrov, Gerald Schimak, Giorgio Guariso, Giulia Fiorese, Giulio Iovine, Gürkan Sin, Hirdy Miyamoto, Ian Littlewood, Ignasi Rodríguez-Roda, Ilya Zaslavsky, Ioannis Athanasiadis, Ion Irizar, Ivan Holoubek, Jaroslav Myšiak, Jean Philippe Steyer, Jennifer Koch, Jim Freer, Jiří Hřebíček, Joaquim Comas, Joaquín Izquierdo, Juan M. Lema, Karina Gibert, Keith Matthews, Kostas Karatzas, Krist V. Gernaey, Kristina Voigt, Les Oxley, Lorenz Hilty, Manel Poch, Manuel Pastor, Mark van Loodstrecht, Martin van Ittersum, Martin Volk, Michael Sheridan, Michela Robba, Michiel Blind, Miquel Sànchez-Marrè, Nigel Quinn , Onno Roosenschoon, Paulo Salles, Peter Gijsbers, Peter Struss, Piotr Magnuszewski, Raffaele Giordani, Ralf Isenmann, Ralf Seppelt, Riccardo Minciardi, Rick Sojda, Robert Argent, Roberto Sacile, Roger Moore, Rüdiger Schaldach, Sigrun Kabisch, Steven Frysinger, Thomas Usländer, Thorsten Wagener, Ulises Cortés, Ulrike Bende-Michl, Virgínia Brilhante, Werner Geiger, Werner Pillmann.

Committees

Convenor

Miquel Sànchez-Marrè FiEMSs, Knowledge Engineering & Machine Learning Group, Technical University of Catalonia, Catalonia

Co-organiser

Joaquim Comas, Laboratory of Chemical & Environmental Engineering (LEQUIA), University of Girona, Catalonia

Scientific Committee

Robert Argent, Melbourne University, Australia

James Ascough III, ARS Tucson, USA

René Bañares-Alcántara, Univ. of Oxford, UK

Bruce Beck, University of Georgia, Georgia, USA

Javier Béjar, Tech. University of Catalonia, Catalonia

Virgínia Brilhante, Federal University of Amazonas, Brazil

Luigi Ceccaroni, Tech. University of Catalonia, Catalonia

Joaquim Comas, University of Girona, Catalonia

Ulises Cortés, Tech. University of Catalonia, Catalonia

Neil Crout, University of Nottingham, United Kingdom

Brian Fath, Towson University, USA

Karina Gibert, Tech. University of Catalonia, Catalonia

Carlo Giupponi, FEEM and University of Milan, Italy

Giorgio Guariso FiEMSs, Politecnico di Milano, Italy

Holly Hartmann, CLIMAS/SAHRA, USA

Anthony J. Jakeman FiEMSs, The Australian Nat. Univ., Australia

Ari Jolma, Helsinki University of Technology, Finland

Kostas Karatzas, Aristotle University, Greece

Guy Larocque, C. de Foresteries de Laurentides, Canada

George Leavesley, USGS, Fort Collins, CO, USA

Carlo Lepori, IDSIA, Switzerland

Rebecca Letcher, The Australian Nat. Univ., Canberra, Australia

Ian Littlewood, CEH Wallingford, United Kingdom

Michael McAleer FiEMSs, Univ. of Western Australia, Perth, Australia

Kimio Morimune, Kyoto University, Japan

Jaroslav Myšiak, FEEM, Venice, Italy

Lex Oxley FiEMSs, Univ. of Canterbury, New Zealand

Claudia Pahl-Wostl, University of Osnabrück, Germany

Manel Poch, LEQUIA, University of Girona, Catalonia

Andrea Rizzoli FiEMSs, IDSIA, Lugano, Switzerland

Ignasi Rodríguez-Roda, University of

Miquel Sànchez-Marrè FiEMSs, Tech. Univ. of

Girona, Catalonia

Ralf Seppelt, UFZ Leipzig, Germany

Jean Philippe Steyer, LBE, INRA, France

David Swayne, University of Guelph, Canada

Martin van Ittersum, U. of Wageningen, The Netherlands

Javier Vázquez-Salceda, Tech. U. of Catalonia, Catalonia

Alexey Voinov FiEMSs, University of Vermont, USA

Catalonia, Catalonia

Richard S. Sojda, NRMSC-USG/Montana State University, USA

Achim Sydow FiEMSs, Fraunhofer FIRST, Berlin, Germany

Clàudia Turon, C. D. de la Conca del riu Besòs, Catalonia

Ilias Vardavas, FORTH and University of Crete, Greece

Kristina Voigt, GSF Neuherberg, Germany

Ruth Waldick, Environment Canada

Organising Committee

Georgii A. Alexandrov, CGER/NIES, Japan

David Arctur, OCGii, USA

Ioannis Athanasiadis, IDSIA, Lugano, Switzerland

Ulrike Bende-Michl, CSIRO Land and Water, Australia

Michiel Blind, RWS-RIZA, The Netherlands

Virgínia Brilhante, Federal University of Amazonas, Brazil

Joaquim Comas, University of Girona, Catalonia

Barry Croke, The Australian National University, Australia

Chris Duffy, Penstate University, USA

Krist V. Gernaey, Technical University of Denmark, Denmark

Peter Gijsbers, WL Delft Hydraulics, The Netherlands

Carlo Giupponi, FEEM and University of Milan, Italy

Giorgio Guariso FiEMSs, Politecnico di Milano, Italy

B. Harris, University of Sheffield, UK

Ivan Holoubek,Masaryk University, Czech Republic

Jiří Hřebíček, Masaryk University, Czech Republic

Ion Irizar, CEIT, San Sebastián, Spain

C. W. Anderson, University of Sheffield, UK

Robert Argent, Melbourne University, Australia

Eduardo Ayesa, CEIT, San Sebastián, Spain

Claudia Binder, University of Zürich, Switzerland

Bert Bredeweg, University of Amsterdam, The Netherlands

Brett Bryan, CSIRO Land and Water, Australia

Ulises Cortés, Tech. University of Catalonia, Catalonia

Neville Crossman, CSIRO Land and Water, Australia

Jim Freer, Lancaster University, UK

Karina Gibert, Tech. University of Catalonia, Catalonia

Raffaele Giordani, CNR-IRSA, Bari, Italy

Carlo Gualtieri, University of Napoli, Italy

Dagmar Haase, UFZ, Leipzig, Germany

Geoffrey Holmes, University of Waikato, New Zealand

A. Holzkaemper, University of Sheffield, UK

Giulio Iovine, CNR-IRPI, Italy

Joaquín Izquierdo, Technical University of Valencia, Spain

Anthony J. Jakeman FiEMSs, The Australian Nat. Univ., Australia

Sigrun Kabisch, UFZ, Leipzig, Germany

Kostas Karatzas, Aristotle University, Greece

Ayalew Kassahun, Wageningen University, The Netherlands

Christof Knöri, EMPA/University of Zürich, Switzerland

Jennifer Koch, University of Kassel, Germany

Vikas Kumar, University of Sheffield, UK

Juan M. Lema, University of Santiago de Compostela, Spain

D. N. Lerner, University of Sheffield, UK

Ian Littlewood, CEH Wallingford, United Kingdom

Piotr Magnuszewski, Wroclaw University of Technology, Poland

Keith Matthews, Macaulay Institute, UK

Brian S. McIntosh, Cranfield University, UK

Dragutin T.Mihailovic, University Novi Sad, Serbia

Riccardo Minciardi, University of Genova/CIMA, Italy

Roger Moore, CEH Wallingford, UK

Jaroslav Myšiak, FEEM, Venice, Italy

Lex Oxley FiEMSs, Univ. of Canterbury, New Zealand

Manel Poch, LEQUIA, University of Girona, Catalonia

David Post, CSIRO Land and Water, Australia

Nigel Quinn, Berkeley National Laboratory, USA

Farmani Raziyeh, IRI, University of Exeter, UK

Michela Robba, University of Genova/CIMA, Italy

Ignasi Rodríguez-Roda, University of Girona, Catalonia

Roberto Sacile, University of Genova/CIMA, Italy

Paulo Salles, University of Brasilia, Brazil

Miquel Sànchez-Marrè FiEMSs, Tech. Univ. of Catalonia, Catalonia

Gerald Schimak, ARC Seibersdorf research, Austria

Huub Scholten, Wageningen University, The Netherlands

Ralf Seppelt, UFZ Leipzig, Germany

Rüdiger Schaldach, University of Kassel, Germany

Gürkan Sin, Technical University of Denmark, Denmark

Richard S. Sojda, NRMSC-USG/Montana State University, USA

Jean Philippe Steyer, LBE, INRA, France

Peter Struss, Technical University of Munich, Germany

B. Surridge, University of Sheffield, UK

Martin van Ittersum, U. of Wageningen, The Netherlands

Mark van Loodstrecht, Tech. Univ. of Delft, The Netherlands

Alexey Voinov FiEMSs, University of Vermont, USA

Martin Volk, UFZ, Leipzig, Germany

Thorsten Wagener, Pennsylvania State University, USA

Patrick Wäger, EMPA, Switzerland

Ilya Zaslavsky, University of California, San Diego, USA

Local Organising Committee

Sergio Álvarez, KEMLG, Technical University of Catalonia, Catalonia
Montse Aulinas, LEQUIA, University of Girona, Catalonia
Cristian Barrué, KEMLG, Technical University of Catalonia, Catalonia
Javier Béjar, KEMLG, Technical University of Catalonia, Catalonia
Guiem Bosch, KEMLG, Technical University of Catalonia, Catalonia
David Cabanillas, KEMLG, Technical University of Catalonia, Catalonia
Luigi Ceccaroni, KEMLG, Technical University of Catalonia, Catalonia
Angela Chieppa, KEMLG, Technical University of Catalonia, Catalonia
Joaquim Comas, LEQUIA, University of Girona, Catalonia
Roberto Confalonieri, KEMLG, Technical University of Catalonia, Catalonia
Ulises Cortés, KEMLG, Technical University of Catalonia, Catalonia
Jordi Dalmau, LEQUIA, University of Girona, Catalonia
Patricia García, KEMLG, Technical University of Catalonia, Catalonia
Karina Gibert, KEMLG, Technical University of Catalonia, Catalonia
Juan Carlos Nieves, KEMLG, Technical University of Catalonia, Catalonia
Xavier Flores, LEQUIA, University of Girona, Catalonia
Carlos Mérida, KEMLG, Technical University of Catalonia, Catalonia
Luis Oliva, KEMLG, Technical University of Catalonia, Catalonia
Fernando Orduña, KEMLG, Technical University of Catalonia, Catalonia
Sofia Panagiotidi, KEMLG, Technical University of Catalonia, Catalonia
Manel Poch, LEQUIA, University of Girona, Catalonia
Pau Prat, LEQUIA, University of Girona, Catalonia
Thania Rendón, KEMLG, Technical University of Catalonia, Catalonia
Ferran Ripoll, LEQUIA, University of Girona, Catalonia
Ignasi Rodríguez-Roda, LEQUIA, University of Girona, Catalonia
Gustavo Rodríguez, KEMLG, Technical University of Catalonia, Catalonia
Miquel Sànchez-Marrè, KEMLG, Technical University of Catalonia, Catalonia
Beatriz Sevilla, KEMLG, Technical University of Catalonia, Catalonia
Joao Sousa, KEMLG, Technical University of Catalonia, Catalonia
Arturo Tejeda, KEMLG, Technical University of Catalonia, Catalonia
Pancho Tolchinsky, KEMLG, Technical University of Catalonia, Catalonia
Clàudia Turon, Consorci per a la Defensa de la Conca del riu Besòs, Catalonia
Javier Vázquez-Salceda, KEMLG, Technical University of Catalonia, Catalonia

Organising Groups

- iEMSs, the International Environmental Modelling and Software Society
- KEMLG, Technical University of Catalonia, Catalonia
- LEQUIA, University of Girona, Catalonia

Collaborating Groups

- ISEM, the International Society for Ecological Modeling
- BESAI, Binding Environmental Sciences and Artificial Intelligence
- MODSS, International Conference on Multi-objective Decision Support Systems for Land, Water and Environmental Management
- ISESS, International Symposium on Environmental Software Systems
- ERCIM, the European Research Consortium for Informatics and Mathematics
- TIAS, The Integrated Assessment Society
- IMACS International Association for Mathematics and Computers in Simulation

Acknowledgements

The main editors and all the members of the editorial board would like to acknowledge the work of more than one hundred of ***reviewers***, whose names cannot be listed here for lack of space. Due to their wonderful work, it has been possible to provide a fair and thorough review to all the papers and abstracts in the proceedings.

The convenors would like to thank the great efforts of all the ***Local Organising Committee members***. Without them, the iEMSs 2008 would not have been possible. Thank you all for your collaboration.

Contents

Session 17: Modelling and simulation of dangerous phenomena, and innovative techniques for hazard evaluation, mapping, mitigation **1386**
(*organised by Giulio Iovine, Michael Sheridan, Manuel Pastor and Hirdy Miyamoto*)

Land subsidence induced by groundwater extraction: the case of Bologna
G. Darini, G. Modoni, M. Saroli and P. Croce 1386

TO EARTHQUAKE PREPARATION MODELS, BASED ON HYDRO GEOCHEMICAL DATA
F.H. Karimov, N.Gh. Salomov, R.U. Vakhobova and G.B. Norova 1394

Use of Two Nonparametric Methods (CART and MARS) to Model the Potential Distribution of Gullies in Spanish Rangelands
Á. Gómez Gutiérrez, S. Schnabel and F. J. Lavado Contador 1398

On study of the geodynamic processes by geoelectrical methods
A. Kamshilin, E. Volkova , V. Leontiev, O. Kuzichkin 1406

Electrochemical stripping analysis, a powerful technique for real-time controlling of environment pollution from heavy metals
G. Aragay, A. Puig-Font, S. Laschi, M. Mascini, L. Sanfilippo, A. Merkoçi 1414

Mathematical modelling of aerosol transport and deposition: Analytic formulae for fast computation
R. McKibbin 1420

A Mathematical Model for Numerical Simulation of Shallow Water Flow: Description and Practical Application of GUAD|2D
J. Murillo, M. Rodríguez Pallarés, A. Andrés-Urrutia, P. Brufau and P. García-Navarro 1431

Effect of Mud Rheology on its Inertial Dam-Break Flow
B. Nsom, K. Bouchlaghem and N. Latrache 1439

Mud-flow and lava-flow susceptibility and hazard mapping through numerical modelling, GIS techniques, historical and geoenvironmental analyses
G. Iovine 1447

Early warning of fast landslides triggering based on instrumented slope data analysis
E. Damiano, R. Greco, A. Guida and L. Olivares 1461

Landslide Data Analysis with Gaussian Mixture Model
V. Timonin, S.B. Bai, J. Wang, M. Kanevski and A. Pozdnoukhov 1469

Modelling the Portalet landslide mobility (Formigal, Spain)
J.A.Fernández-Merodo, G.Herrera, P.Mira, J.Mulas, M.Pastor, L.Noferini, D.Mecatti and G.Luzi 1476

The role played by mountain tracks on rainfall-induced shallow landslides: a case study
L. Cascini, S. Cuomo and M. Pastor 1484

Additional torrent control strategies on debris flow alluvial fans with extremely high vulnerable settlements
B. Mazzorana, C. Scherer and N. Marangonia 1492

Modelling of the Risks of the Sidoarjo Mud Flow, Indonesia
S. Kivva and M.Zheleznyak 1500

Rheological Behaviour of Volcanic Granular Flows
K. Kelfoun 1508

Geomechanical modelling of triggering mechanisms for rainfall-induced triangular shallow landslides of the flow-type
L. Cascini, S. Cuomo, M. Pastor and J.A. Fernández-Merodo 1516

A Semantic WEB Services Platform to support Disaster and Emergency Management
M. Gloria, V. Lersi, G. Minei, D. Pasquariello, C. Monti and A. Saitto 1524

Drought Mitigation using Operative Indicators in Complex Water Systems
G.M. Sechi and A. Sulis 1532

Living With Peak Discharge Uncertainty: The Self-learning Dike
J. L. de Kok and A.Y. Hoekstra 1542

An integrated multimedia model and multicriteria analysis approach to managing sewage sludge application on agricultural soils: framework and methodology
A. C. Passuello, M. Schuhmacher, V. Kumar, M. Nadal and J. L. Domingo 1550

Assessment and Management of Environmental Risks Resulting from Operation of Large-scale Irrigation Delivery Systems: a Case Study in Southern Italy
D. Zaccaria and M. Scimone 1558

Geo-communication for risk assessment and catastrophe prevention of flood events in the coastal areas of Chennai
S. Glaser, R. Glaser, A. Drescher, C. Pfeiffer, E. Schliermann-Kraus, M. Lechner and J. Vencatesan 1569

Simulating earthquakes effects: the SHAKYGROUND simulation tool
M. Cristaldi, U. Delprato, C.Lupo and S. Gresta 1578

Thunderstorm Strike Probability Nowcasting, a New Algorithm
S. Dance, B.Ebert and D.Scurrah 1586

Session 18: Environmental Informatics: History and Current Trends **1594**
(*organised by Werner Pillmann, Bernd Page, Lorenz Hilty, Werner Geiger, Ralf Isenmann and Kristina Voigt*)

Environmental Informatics, Environmetrics, Chemoinformatics, Chemometrics: Integration or Separation
K. Voigt 1594

Informatics and the Life Cycle of Products
L. M. Hilty, R. Hischier, T. F. Ruddy, C. Som 1602

Current Trends in E-Environment and its Role in E-Democracy
J. Hřebíček, R. Legat and M. Nagy 1612

Material Flow Analysis for Eco-Efficiency with Material Flow Network Models-Concepts & Case Study
B. Page, V. Wohlgemuth and M. Raspe 1620

The Growing Importance of Open Service Platforms for the Design of Environmental Information Systems
T. Usländer 1628

Environmental Management Information Systems – Illustrations from Online Communication and Sustainability Reporting
R. Isenmann 1636

Developments in urban environmental information perception and communication
K. D. Karatzas and J. Lee 1645

Environmental Information Systems of Public Administration – Status and Trends
W. Geiger and W. Pillmann 1653

Session 19: Interaction Design for Environmental Information Systems **1660**
(*organised by Daryl Hepting and Steven Frysinger*)

Interaction design for rural agricultural sensor networks
K. Walker, A. Kabashi, J. Abdelnour-Nocerac, K. Ngugid, J. Underwood, J. Elmirghani and M. Prodanovic 1660

Graphical User Interface Design for Climate Impact Research Data Retrieval
M. Wrobel, T. Nocke, M. Flechsig and A. Glauer 1669

On the Design of an Immersive Environment for Visual Mining: Challenges and Results from an Early Study
E. Mozaffari, A. Seffah, A. Blakney and R. Naghshin 1677

Reclaiming Garbage: Interaction Design for Participation
M. Rijsberman 1686

An Analysis of Interface Quality for Decision Support
T. Maciag and D.H. Hepting 1694

On the Role of Reflection and Representation in Environmental Decision Support Systems
T. Maciag and D.H. Hepting 1701

Session 20: Data Mining as a Tool for Environmental Scientists **1708**
(*organised by Karina Gibert, Miquel Sànchez-Marrè, Joaquim Comas, Ioannis Athanasiadis, Joaquín Izquierdo and Geoffrey Holmes*)

Monolith2 - An On-line Database for Cement/Waste Products
M. O'Shea, J.A. Stegemann and M. Levene 1708

Coupling empirical and deterministic models to assess surface water contamination
M.Terrado, M.P. Lavigne, D. Barceló, S. Duchesne, J.P. Villeneuve, A.N. Rousseau, S. Tremblay and R.Tauler 1716

A data mining approach to discover weather patterns contributing to PM10 exceedances
Sfetsos A. and Vlachogiannis D. 1727

Virtual Sources for Spatio-temporal Monitoring Data Analysis
A. Nuzhny, E. Saveleva, S. Kazakov and S. Utkin 1734

Applying Data Mining Methods for Forest Planning Data Validation
A.Mäkinen, A.Kangas and T.Tokola 1742

Entropy-based fuzzy set optimisation for reducing ecological model complexity
A.M. Mouton, B. De Baets, A. Peter, G. Holzer, R. Müller and P.L.M. Goethals 1750

Identifying Biological Echoes in Radar Scans Using Machine Learning
R. Mead, J. Paxton and R.S. Sojda 1758

Statistical analysis of spatial plant patterns under the effect of forest use
I. López, T. Standovár, J. Garay , Z. Varga and M. Gámez 1766

Multivariate Spatio-Temporal Clustering (MSTC) as a Data Mining Tool for Environmental Applications
F.M. Hoffman, W.W. Hargrove, R.T. Mills, S. Mahajan, D.J. Erickson and R.J. Oglesby 1774

A Particle Swarm Optimization derivative applied to cluster analysis
J.L. Díaz, M. Herrera, J. Izquierdo, I. Montalvo, R. Pérez 1782

Automatic generation of conceptual descriptions of classifications in Environmental Domains
A. Perez-Bonilla, K. Gibert and D. Vrecko 1791

A neural network approach for selecting the most relevant variables for foaming in anaerobic digestion
J. Dalmau, J. Comas, I. Rodríguez-Roda, E. Latrille and J.-P. Steyer 1799

Trajectories' Mining Between Subprocesses in a Wastewater Treatment Plant
K. Gibert, G. Rodríguez Silva and I. Rodríguez Roda 1805

Identifying State Variables in Multivariate Hydrological Time Series Using Time Series Knowledge Mining
O. Gronz, M. Casper, P. Gemmar 1813

Prediction of Significant Wave Height Based on Regression Trees
A. Etemad-Shahidi and J.Mahjoobi 1821

Session 21: Environmental Modelling and Decision Support Systems **1827**
(*organised by Miquel Sànchez-Marrè, Joaquim Comas, Ignasi Rodríguez-Roda, Rick Sojda, Andrea Emilio Rizzoli, Clàudia Turón, Karina Gibert, Ulises Cortés and Manel Poch*)

A spatial decision support system for the optimal environmental reclamation of open-pit coal mines in Greece
F. Pavloudakis, M. Galetakis and Ch. Roumpos 1827

Design and Implementation of the Land Use Management Support System (LUMASS)
A. Herzig 1835

Sustainable Use of Pesticides: an Innovative Experience in Lombardy Region, Italy
C. Riparbelli, A. Di Guardo, A. Serrano, E. Capri and S. Brenna 1843

Influencing Factors for Scheduling Optimal Pipeline Replacement with Budgetary Constraints 1851

C.D. Alonso, R. Pérez, P. A. López and X. Delgado

The Emission Inventory Water, a Planning Support System aimed at reducing the pollution emissions in the surface waters of Flanders, Belgium
L. Van Esch, G. Vos and G. Engelen 1859

Predicting Winners and Losers: Translating Air Pollution Policy Initiatives into Indicators of Habitat Quality Change
E.C. Rowe, C.D. Evans, S. Smart and B.A. Emmett 1869

Monitoring the effects of human activities and environmental change on a population system
M. Gámez, I. López, J. Garay and Z. Varga 1877

III. Workshop Position Papers and Abstracts

Workshop 1: Creating robust sensor networks architecture and infrastructure **1885**
(*organised by Gerald Schimak, Thomas Usländer*)

Creating robust sensor networks - architecture and infrastructure
G. Schimak, Th. Usländer, D. Havlik and R. Argent 1885

A taxonomy of robustness in sensor service networks
K.S. Watson and T. Usländer 1898

Advanced Cascading Sensor Observation Service
D. Havlik, G. Schimak, R. Barta 1901

Coupled Modelling of Forest Fires and Floods in Multi-risk Scenarios under SCIER System
T. Metelka and D. Caballero 1905

Towards Process Oriented Environmental Monitoring Systems for Intelligent Risk Management
O. Faust, B.H.C. Sputh and A.R. Allena 1909

Workshop 2: Evaluating the impact of data and associated uncertainty on hydrological model predictions **1912**
(*organised by Barry Croke, Thorsten Wagener, David Post, Jim Freer and Ian Littlewood*)

Evaluating the information content of data for uncertainty reduction in hydrological modelling
B. Croke, T. Wagener, D. Post, J. Freer and I. Littlewood 1912

Effect of Rainfall Data Quality on Rainfall-Runoff Model Calibration and Regionalisation
D. A. Post, J. Vaze, F. H. S. Chiew, J.-M. Perraud, J. Teng and N. R. Viney 1927

Prediction of Streamflow and the Effect of Input Parameters in Small Urban Catchments
H. J. Kim, S. J. Noh, C. H. Jang and Y. J. Lee 1929

Ensemble modelling of runoff across southeast Australia and quantification of uncertainty
F.H.S. Chiew, N. R. Viney, J. Vaze, D.A. Post, J.-M. Perraud and J. Teng 1931

Lessons Learned from Experiments in Auto-Calibration of Large Scale Environmental and Hydrological Models
D.A. Swayne, M. Sloboda, S. Hood, V. Sharma, W. Booty and D. Lam 1932

Quantifying input and model errors in conceptual rainfall-runoff models using Bayesian total error analysis
D. Kavetski, B. Renard, G. Kuczera, M. Thyer and S. Franks 1933

New hybrid method to simulate discharges on ungaged catchments by combination of hydrological and hydraulic models
J. Lerat, V. Andréassian and C. Perrin 1934

Uncertainty and evolution of rating curves: a key issue for the reliability of rainfall-runoff models?
T. Mathevet, L. Oudin, V. Andreassian, C. Perrin 1936

Workshop 4: Data Mining as a Tool for Environmental Scientists **1937**
(*organised by Karina Gibert, Miquel Sànchez-Marrè, Joaquim Comas, Ioannis Athanasiadis, Joaquín Izquierdo and Geoffrey Holmes*)

On the role of pre and post-processing in environmental data mining
K. Gibert, J. Izquierdo, G. Holmes, I. Athanasiadis, J. Comas and M. Sànchez-Marré 1937

Data mining using k-means clustering and classification and regression trees (CART) as post-processing methods: identifying management and environmental factors for explaining sugarcane yield in Northern Argentina (1971-2005)
D.O. Ferraro 1959

Experiments in Auto-Calibration of Large Scale Environmental Models using a Data Mining Approach
D.A. Swayne, M. Sloboda, S. Hood, V. Sharma, W. Booty and D. Lam 1961

Swarm Behaviour for Visual Data Mining
I. Montalvo and J. Izquierdo 1962

Data Mining vs. Mathematical Modelling: Nonlinear Dynamics and Chaos Theory
Morales-Acoltzi T., Medel-Rojas A., Morales-Aceves F. J., Cortés-Pérez E and Morales-Calva X. 1963

Pre and Postprocessing in KLASS
K. Gibert and R. Nonell 1965

GESCONDA: from Environmental Data Mining to Environmental Decision Support
M. Sànchez-Marrè and K. Gibert 1967

Workshop 5: Qualitative models for Sustainability **1971**
(*organised by Bert Bredeweg and Paulo Salles*)

W5 Description 1971

B. Bredeweg and P. Salles

Garp3 - Workbench for Qualitative Modelling and Simulation
B. Bredweg, A. Bouwer, F. Linnebank and J. Liem 1972

A qualitative model on sustainability of the Riacho Fundo basin (Brasília, Brazil)
P. Salles, B. Bredeweg and A.L. Rios Caldas 1975

Qualitative modelling of limiting factors for a salmon life cycle in the context of sustainable river rehabilitation
R. A. A. Noble, B. Bredeweg, P. Salles & I. G. Cowx 1978

Evaluating the potential of qualitative reasoning (QR) models to contribute to a sustainable catchment management
A. Zitek, S.Muhar, S. Preis, B. Bredeweg, P. Salles and S. Schmutz 1981

Collaborative Capture and Reuse of Environmental Knowledge
J. Liem and B. Bredeweg 1984

A new approach for river exploration : River Mesta case study for qualitative modeling sustainable development
E. Nakova, B. Bredeweg and P. Salles 1987

A library of model fragments on sustainability
P. Salles, B. Bredeweg E. Cioaca, E. Nakova, A. Zitek and R. Noble 1989

A Qualitative Reasoning–based Curriculum for learning about Sustainable Development
T. Nuttle and A. Bouwer 1992

Danube Delta Biosphere Reserve Model of Environmental Sustainability based on Qualitative Reasoning Concept
E. Cioaca, B. Bredeweg and P. Salles 1995

A qualitative model about the Island Biogeography theory
Débora Leite Silvano and Paulo Salles 1998

Simulating an uncertain world: using qualitative reasoning to model a plant-resource system
Tim Nuttle, Bert Bredeweb, Paulo Salles, and Michael Neumann 2001

A Simple Tool for Preference Rating or Multi-criteria Assessment
Peter K A Barz 2003

Workshop 6: Integrated solution support systems for water management **2006**
(*organised by Ayalew Kassahun, Michiel Blind, Arno Krause and Onno Roosenschoon*)

Integrated Solution Support System for Water Management
A. Kassahun, M. Blind, A. Krause and O. Roosenschoon 2006

Supporting deliberation processes within complex decisions in natural resources management
Máñez M. and S Panebianco 2024

System Dynamics Modelling: A Tool for Participatory Simulation of Complex Water Systems within AquaStress
L.S. Vamvakeridou-Lyroudia and D.A. Savic 2025

An experimenting protocol for extracting and exploring knowledge from citizens interviews for participatory water management
G. Abrami, N. Ferrand and J.E. Rougier 2029

A Multi-Agent system for water management in river basins
T. Rendón-Sallard, M. Sànchez-Marrè, M. Aulinas and J.Comas 2031

WARGI: a Tool for Water System Drought Mitigation within EU and National Projects
G.M. Sechi and A. Sulis 2033

Workshop 7: Community modeling, data and model interoperability **2035**
(*organised by Alexey Voinov, Chris Duffy, David Arctur, Ilya Zaslavsky and Ralf Seppelt*)

Community modelling, and data-model interoperability
A. Voinov, I. Zaslavskiy, D. Arctur, C. Duffy and R. Seppelt 2035

The Great Rivers Modeling Framework and Decision Support System
B. Eckman, C. Barford and P. West 2051

A solute and water flux library for catchment models
P. Kraft, K.B. Vaché, L. Breuer, H.-G. Frede 2053

Promoting access to and reuse of modelling projects - A picture book approach
R. Seppelt 2056

Workshop 8: Actors and factors – bridging social science findings and urban land use change modelling **2059**
(*organised by Dagmar Haase and Sigrun Kabisch*)

Introduction. The challenge of bringing actors and factors together and social science issues to be addressed
D. Haase and S. Kabisch 2059

Household formation framed by the Second Demographic Transition
S. Kabisch and D. Haase 2074

Mapping out the social impacts of land use: Soft-GIS and Choice-Based Conjoint Analysis as integrative methods
A. Zuin 2076

Mapping Survey Data into Agents' Behavioural Rules for ABMs: Motivation and Challenges
T. Filatova and A. van der Veen 2077

Actors as factors for land use change: Effects of demographic change on land use across Europe
T. Tötzer 2080

Land use policy impacts in the rural-urban region: a modelling framework
S. Scatasta 2082

Quantification of land use changes, modelling of processes and impacts of shrinkage and the question of scale
S. Schetke 2083

Integrated Modelling of Smallholder Land-use Decisions on the Availability of Rattan
F. von Walter 2085

Workshop 10: Participatory and Conceptual Modelling – Making Models Amenable to Stakeholders' Concerns **2092**
(*organised by Jaroslav Myšiak, Raffaele Giordani, Carlo Giupponi, Farmani Raziyeh, Nigel Quinn, Carmel A. Pollino, Anthony J. Jakeman and Piotr Magnuszewski*)

Using a Bayesian network to assess the population-level risks of endocrine disruptors to brown trout
M. Borsuk, S. Schubert, A. Peter and P. Burkhardt-Holm 2092

Assessment of agri-enviromental measures on water quality using a Bayesian Network
M. Carpani and C. Giupponi 2093

Integrated methodology for the management of uncertainty and complexity in climate change adaptation policies: a Bayesian Decision Network approach
M. Catenacci and C. Giupponi 2096

Actor Modeling and Actor-based Modeling: New Methods to Analyze the Human Component of Human-Environment Systems
P. Döll, C. Döll and A. Titz 2098

Knowledge elicitation games on stakeholder decision-making: do we get closer to agents' behaviour?
D. Haase 2101

Waterlogging in Upper Tisza and the driving forces behind
J. Mysiak and Z. Nagy 2102

Sustainability Evaluation for Shellfish Production in Gamak Bay based on the Systems Ecology
H.T. Oh, W.C. Lee, S.M. Lee , R.H. Jung and S.J. Hong 2103

Modelers and stakeholders: who is the boss?
A. Voinov and F. Bousquet 2105

Workshop 11: Bridging the gaps between design and use: evaluating decision and information support tools **2106**
(*organised by Brian S. McIntosh, Carlo Giupponi, Keith Matthews and Martin van Ittersum*)

Some Thoughts on Success in Applying Models to Water Resource Systems
G.M. Sechi and A. Sulis 2106

From INGIS to NAMOS – Operationalisation and contextualisation of sustainability at the local level
D. Rink, A. Haase, S. Kabisch, D. Haase and G. Hartmuth 2107

Workshop 12: ICT Tools For Environmental And Human Risk Assessment **2109**
(*organised by Jiří Hřebíček and Ivan Holoubek*)

ICT Tools For Environmental And Human Risk Assessment
I. Holoubek, J. Hřebíček and L. Dušek 2109

An Intelligent Decision Support System for Environmental Impact Assessment of Industrial/Agricultural Projects
Shady Aly and Ivan Vrana 2118

Workshop 13: Taking the OpenMI forward **2120**
(*organised by Roger Moore and Peter Gijsbers*)

Taking the OpenMI forward
R.V. Moore and P.J.A. Gijsbers 2120

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Land subsidence induced by groundwater extraction: the case of Bologna

G. Darini, G. Modoni, M. Saroli and P. Croce

Di.M.S.A.T. – Dipartimento di Meccanica, Strutture, Ambiente e Territorio, Università di Cassino, Via di Biasio, 42 03043, Cassino, Italia (g.darini@unicas.it)

Abstract: A comprehensive investigation of the subsidence induced in the area of Bologna (Italy) by groundwater extraction is presented. Initially, the geological pattern of the region is reviewed in order to identify the large scale subsoil features governing the water flow regime and the ground displacements. Several types of data, covering the different aspects of the problem are then combined into a Geographical Information System and interpolated data by means of geostatistical analyses. A detailed reconstruction of subsidence over the considered area is accomplished by processing available topographical measurements. The subsoil stratigraphy, the physical, hydraulic and the mechanical characteristics of the different geotechnical units are estimated by integrating the results of available geotechnical site and laboratory investigations. The groundwater regime is investigated by correlating the water head measures recorded on several pumping wells distributed over the area of Bologna. A groundwater flow simulation model has been finally formulated in order to derive the undisturbed pre-existing water level, to identify the effects of the actual groundwater withdrawal and to predict the modifications induced by possible future exploitations.

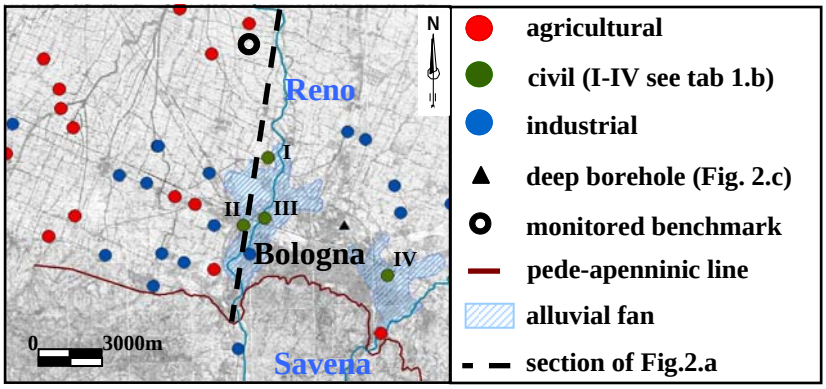
Keywords: Settlements; Seepage; Model; Spatial analysis.

1. INTRODUCTION

In the second half of the past century intensive groundwater extraction was made in the Po Valley by well pumping from the subsoil bearing strata, in order to satisfy the increasing demand of water due to the rapid growth and concentration of inhabitants and economical activities. Such a strong groundwater exploitation induced significant modifications on the groundwater flow regime and different undesired effects on the urban and natural environments. One main consequence, also recorded in other lowland areas of the world [USGS, 1999], was a marked increase of ground subsidence observed around the urban areas. In fact, while the whole Po valley is subjected to a long term natural subsidence of few millimetres per year, produced by the self weight consolidation of the high thickness alluvial deposits, in the period after 1950 some cities of northern Italy (Milano, Venezia, Ravenna and Bologna are the most enlightening cases), suffered very high settlements, reaching in some cases rates of dozens centimetres per year [Carminati and Martinelli, 2002]. This situation induced several local institutions to undertake investigations aimed to discover the causes of such unexpected movements and to quantify their extent. However, due to the large extension of the subsiding areas, to the natural delay of the phenomenon and to the lack of a comprehensive investigation method, covering all the different aspects of the problem, only qualitative conclusions could be retrieved for long time. It is nowadays possible to process the large amount of data collected over the years on large geographical

areas by means of dedicated software and to perform more complete analysis of the phenomena.
This is the situation for the city of Bologna, which is the case presented in this paper. After the second world war, the city experienced a significant demographic growth and industrial development. In particular, from 1950 to 1960, the population increased of about 33% (from 330.00 to 440.000 inhabitants) (Fig. 1.c) [Osservatorio Astronomico di Bologna, 2007]. To meet the subsequent water requirements, a large number of deep pumping wells was excavated all around the city (Fig.1). In the same period, intensive subsidence was observed, distributed with different settlement rates, all over the urban area [Pieri and Russo, 1985]. The most clear sign of correlation between well pumping and subsidence can be inferred by comparing the similar time progressions of the cumulated number of wells and of recorded settlements at a typical monitored benchmark (Fig.1.b).

(a)



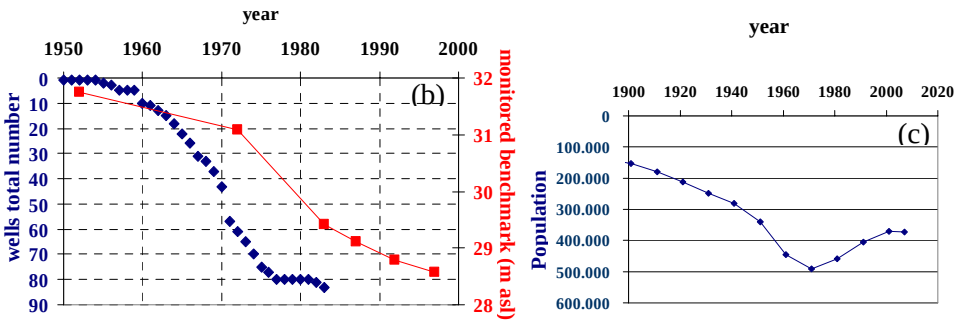


Figure 1. Map of Bologna area (a); progression with time of well's number and ground level at a monitored benchmark (b); population growth [Regione Emilia Romagna – UnionCamere, 2007] (c).

The plot shows that the most sharp increase of well number is concentrated in the period between 1960 and 1975, while the maximum settlement rate occurred after 1970. Explanation for such delay requires a more detailed analysis on the type of wells and the relative amount of water withdrawn from each of them. To this aim, a distinction of extracted water depending on its different uses (namely civil, industrial and agricultural), provided for two years by Vassena [2003], shows that the larger amount of water is withdrawn for civil activities (tab.1.a). Furthermore, it must be considered that while the agricultural and industrial activities are more numerous and widely distributed all over the country surrounding Bologna, the most active aqueduct wells (tab.1.b), excavated in a relatively short period around 1970, are concentrated only in four zones within the two alluvial fans of Reno and Savena Rivers (Fig.1.a).
Information available for the present analysis concerns ground movements, subsoil characteristics, groundwater flow and piezometric heads. These different data, collected in the past decades, have been included into a Geographical Information System (GIS), covering an area of about 200 square kilometres, to seek a mechanical interpretation of the observed phenomena. In order to correlate the different variables, each of them characterised by an independent spatial distribution of measures, interpolation over the considered area has been achieved by means of geo-statistical methods which consider each

measure as a sample of a statistical population. A mathematical model (theoretical variogram), defined to simulate the statistically observed spatial variability of measures (experimental variogram), forms the basis for interpolation of measures over the study area by means of a Kriging method [Matheron, 1971]. By this method all the different considered variables have been evaluated on the same grid of points uniformly distributed over the Bologna area with a spacing of 500 m.

Table 1. Groundwater volumes extracted in the years 1985 and 2002
(a. water extraction distinguished by use; b. water extracted from each single civil well)
[Vassena, 2003].

Use	Civil	Industrial	Agricultural	Zootechnical	Total	Year
Groundwater withdrawal (10^6 m^3/year)	43.9	18.4	5.1	1.1	68.5	1985
	38.5	22.7	9.5	0.6	71.3	2002

(a)

Well location name	S. Vitale.	Borgo Panigale	Tiro a segno	Fossolo	Year
Position in Fig.1.a	I	II	III	IV	
Groundwater withdrawal (10^6 m^3/year)	12.1	12.1	13.2	6.5	1985
	12.7	12.5	7.9	5.4	2002

(b)

2. SUBSOIL CHARACTERISTICS

The city of Bologna is located in the southern part of the Po Valley, at the foothills of the Apennines. From a geological viewpoint, the Po Valley can be considered as the result of a foreland basin evolution process [Carminati and Martinelli, 2002] with the Apennines and the plain presently interested by different evolution processes. The Po Valley links the padanian-adriatic sector with the external portion of the thrust belt composed by the tectonic units of the Upper Miocene – Quaternary, currently subjected to tectonic up-lift movements and consequent erosive processes [Bartolini *et al.*, 1996]. Starting from the middle Plio-Quaternary, when the whole Po valley was an extension of the present Adriatic sea, the basin was gradually filled by alternations of marine and continental sediments, the latter coming from the erosion of the Apennines [Amorosi *et al.*,1996]. This dynamic evolution caused the transition line between the continental slope and the marine basin to progressively advance eastward [RER and ENI-AGIP, 1998].
A comprehensive subsoil reconstruction of the southern part of the Po Valley is available thanks to a study carried out for oil and water exploitation [RER and ENI-AGIP, 1998]. Combining boreholes and deep wells stratigraphies together with seismic investigation results, three thick aquifer groups, named respectively A (shallow), B (intermediate) and C (deep), were identified. Each group consists of alternation of coarse and fine grained soil strata and is separated from the others by thick impervious barriers of regional extension [Farina et al., 2000]. The group A in the Bologna area is mainly formed by the two big alluvial fans of Reno and Savena rivers (Fig.2.a), which are the most important groundwater supplies of the city [Elmi, 1984]. It is thus reasonable to consider that the water originally permeated from the alluvial fans spread into the pervious layers of the plan area, forming a relatively shallow water table which encouraged people to undertake groundwater extraction. From a more detailed investigation of the subsoil composition, obtained by the logging of a deep borehole retrieved in the valley (Fig.2.b), a recurrent stratification of relatively thin layers of coarse and fine grained soils is seen, meaning that these two soil types are mutually penetrating and cannot be clearly distinguished on the map (the border of the shaded area traced in Fig.1.a, represents only the limit of the alluvial fan portion outcropping at the ground level).
Starting from the RER and ENI-AGIP [1998] study, the analysis over the Bologna area has been further developed [Darini, 2007]. In fact, for each aquifer group, the study reports contour maps respectively of the bottom depth and of the cumulated thickness of coarse grained deposits (considering these latter as the more pervious and thus useful for water exploitation).

These contour maps have been digitalised and interpolated by means of a geostatistical analysis, over an area of 15x15 km^2 around Bologna (Figs.3.a and b), with the aim of calculating the cumulated thicknesses of fine grained soil deposits as differences between total and coarse grained soil thickness (Fig.3.c). The results of this analysis for the topmost aquifer group (A), which has been considered, due to its upper position, as the main responsible for the observed subsidence, show that the maximum cumulated thickness of fine grained soils occurs at the boundary of the two alluvial fans of the Reno and Savena rivers.

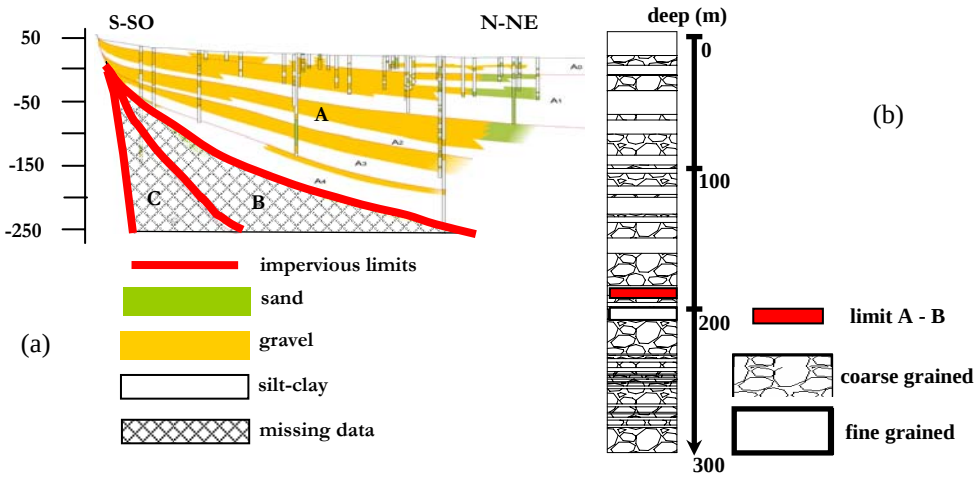


Figure 2. Litho-stratigraphical representation of the Bologna subsoil obtained from a cross section (a) traced along the Reno river [modified from Regione Emilia-Romagna , 2007] and from a deep borehole (b) whose plan position is shown in fig.1 [modified from IDROSER, 1989].

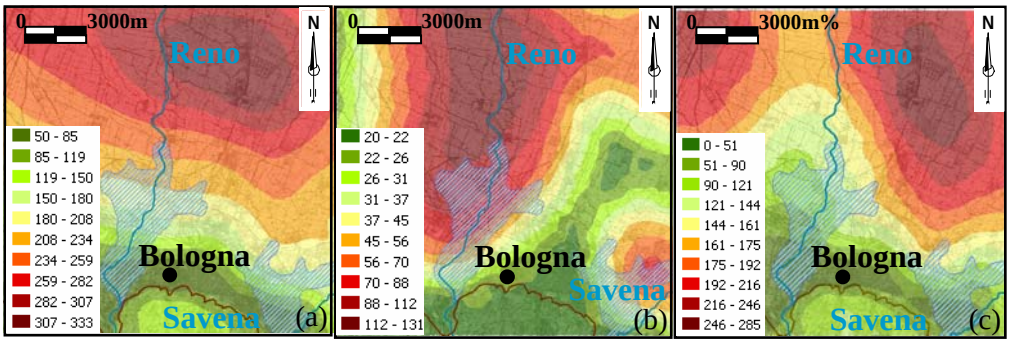


Figure 3. Aquifer Group A: total thickness (a), cumulated coarse grained soil thickness (b), cumulated fine grained soil thickness (c).

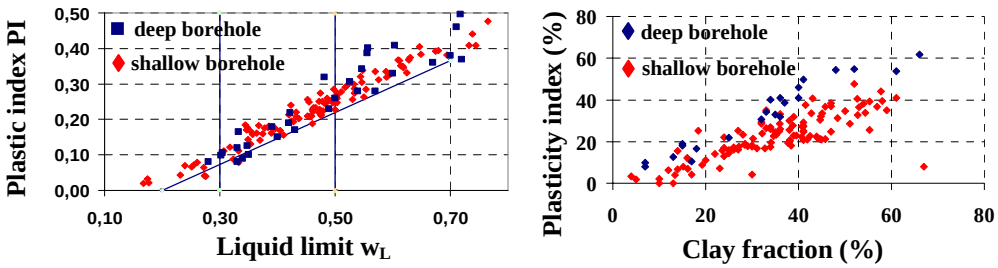


Figure 4. Casagrande plasticity chart for fine grained soil.

The physical and mechanical characteristics of the subsoil in the whole area have been investigated by collecting the available results of fields and laboratory tests performed at

different locations and depths. From this analysis a randomly distributed grain size composition of fine grained soils is seen over the considered area with a predominant fraction of silt (ranging between 40 and 60%), a slightly lower content of clays (from 10 to 50 %) and a small percentage of sand (from 0 to 20 %). The overall homogeneity of this soil can be also seen from the Casagrande plasticity chart, and the activity reported in (Fig.4) by distinguishing with different dots the values obtained on samples taken at a depth from the ground level between 50 and 300 m from those taken at more shallow depths. The plot shows a similar scattering for both sets of data, with the random position of the points on the diagram dictated by the different amount of clay.
Similar conclusions can also be retrieved for the mechanical characteristics of these deposits. In fact, even though a scattering is generally observed on the different parameters, the average values and coefficient of variation do not show significant spatial differences. The average values and the standard deviations for the compressibility, swelling and consolidation indexes, obtained from all the performed tests, are summarised in table 2. One last comment for these deposits concerns their stress history, which has been evaluated by means of the over consolidation ratios (OCR), estimated by oedometer tests. These values are equal to about 6 near the ground level and gradually decrease to 1 at about 20 m depth, showing that the fine grained deposit can be thus assumed as globally normal consolidated, apart from their shallow portion affected by water table fluctuations.

Table 2. Compressibility, swelling and consolidation coefficients of fine grained soils.

	λ	κ	c_v (cm^2/s)
Average	0,120	0,030	6,3E-03
St.Dev.	0,036	0,015	3.69E-02

Concerning the coarse grained soils, several pumping tests [RER and ENI-AGIP, 1998] gave values of the Darcy's coefficient ranging between 10^{-3} and 10^{-5} m/s without showing any clear spatial distribution. Simplifying assumptions will be introduced concerning this aspect for the analysis developed in the next paragraph.

3. GROUNDWATER SEEPAGE

The distribution of water heads in the subsoil of Bologna could be estimated only for the period after 1976, i.e. when measurement of water level on about 150 different types of wells was started [Regione Emilia Romagna, 2007]. A sample of piezometric heads contour map obtained by interpolation of the average water level measured in 1985 for the aquifer group A shows a strong water table depression within the alluvial fan of the Reno river. If flow lines are traced orthogonally to the constant head curves, a clear convergence of water is observed toward the zone where the most important pumping activity is concentrated.
Unfortunately, the available water head measurements do not allow to calculate the water table drop induced by water pumping. Based on hystorical information from the XIX century reporting groundwater outcrops near the ancient road named "via Emilia", previous analysis by Vassena [2003] estimate this head lowering by considering the water table originally located at the ground surface. In the present study, this aspect has been tackled by modelling the water flow regime with a numerical two dimensional finite-difference model [Modflow + Groundwater Vistas, version 3, 2001]. This analysis has been conducted to the upper aquifer group (A), from which most of the water is extracted, by assuming a perfectly impermeable layer at its bottom. To this aim the area reported in figure 3, extending over 15x15 squared kilometres with side boundaries falling out of the influence region of the most active wells [Kezdi e Marko, 1969], has been subdivided into a regularly distributed square grid of 500 meters spacing. According with the previously recalled RER and ENI-AGIP [1998] study the thickness on each vertical has been calculated as equal to the difference between ground and bottom levels. Two numerical simulations have been performed respectively for the two years 1985 and 2002, for which estimates of the extracted water volumes are available (see tab.1). In particular, the water flow withdrawn from each industrial well has been assigned by equally subdividing the total amount of water extracted for industrial purpose, while direct estimations are available for the four

civil wells located in the study area. The boundary conditions at the contour of the considered area have been provided by assigning nil water flow on the Apennines and the water head values interpolated from measurements on the valley and the alluvial fan. The subsoil permeability for any plan position (x,y) has been computed by the following relation:

$$K(x,y) = K^* \cdot \frac{S_{af}(x,y)}{S_t(x,y)} \quad (1)$$

where S_t and S_{af} represent respectively the total thickness of the aquifer group and the cumulated thickness of the coarse grained soils evaluated in each position (x,y) (see Fig. 3.a and b). K* is a reference permeability value, equal to $3.5*10^{-4}$ m/s (i.e. within the estimated limits of coarse grained soils), found by a trial and error procedure in order to match the measured water head. Considering the large simplification introduced in the model, the comparison between water heads obtained from the calculation and from the interpolation of measurements is generally satisfactory for both years 1985 and 2002 (Fig.5.b). This seepage model has been then applied to observe the groundwater regime under different conditions. In particular, if the industrial wells are virtually turned off in the analysis it is readily observed that their effect on the position of the groundwater table is negligible. On the contrary, the extraction for civil use is the most important responsible for the water head drop observed in the central part of the alluvial fan of the Reno river. Finally nil flow has been assigned to all wells to recover the original undisturbed position of the water table (i.e. free of well pumping) which is unknown.

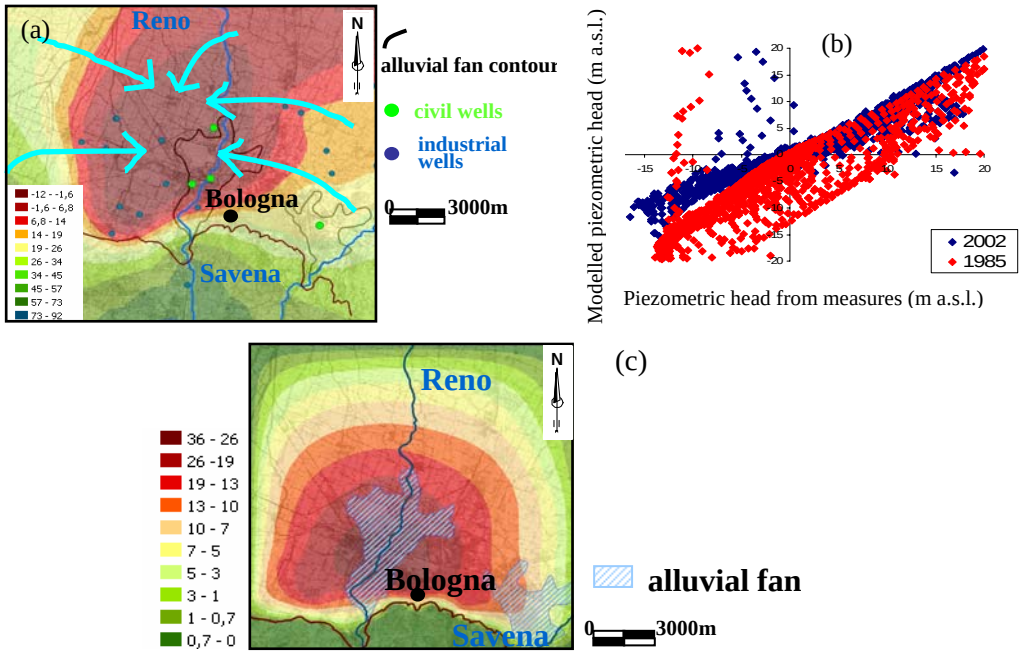


Figure 5. Water head in the Bologna area (a. contour map of water head from measurements (1985); b. modelled vs measured water heads for 1985 and 2002; c. water head reduction Δh (m) due to well pumping).

Once this position is established, the patterns of water head reduction due to any assigned extraction activity can be obtained. A representative situation of the largest observed water head drop in the area can be obtained by considering the extraction activity of 1985 and by calculating the corresponding water head drop (Fig. 5.c).

4. GROUND SURFACE SETTLEMENTS

The ground surface settlements have been monitored by different geometrical levelling campaigns carried out by local authorities. A first levelling network consisting of 200

benchmarks positioned primarily along the "via Emilia" was installed between 1950 and 1980 from the Management District of the Reno River Basin. These points were partly incorporated in a more extended and dense net of 500 benchmarks by the Municipality of Bologna which promoted four extensive measurement campaigns respectively, in 1983, 1987, 1992 and 1999. The settlements recorded on one particular benchmark common to the two networks are reported in fig.1.b. In order to calculate the settlements distribution all over the study area in a period covering the subsidence phenomenon, interpolation of levelling measures has been performed by means of geostatistical analysis. By this operation different contour maps could be drawn for different selected periods. Two examples of settlement contour maps respectively for the two periods 1972-1997 and 1983-1999 are reported in Fig.6.a and b. These maps show a typical characteristic of geostatistical interpolation. In fact, compared to the 1983-1999 map, which is based on a larger number of homogeneously distributes measures, the 1972-1997 map is less accurate, as also confirmed by the numerical tests performed on such interpolation. However, independently of the observation period which obviously influences the amount of observed settlement, both maps show the typical subsidence pattern of the Bologna's area, with larger settlements occurring around the two alluvial fans.

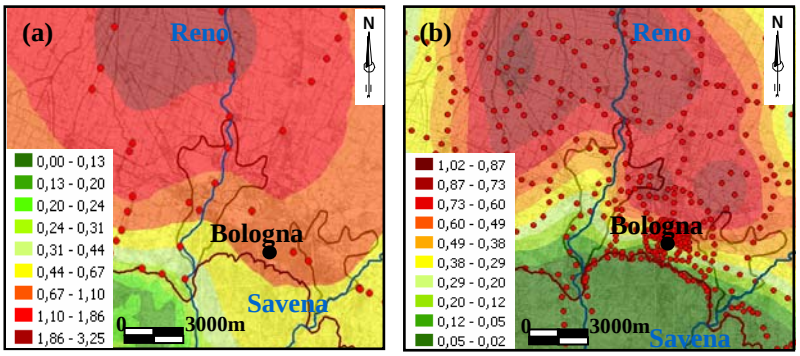


Figure 6. Geometrical levelling over the area of Bologna (a. contour map of settlements (m) in the period 1972-1997; b. same for 1983-1999).

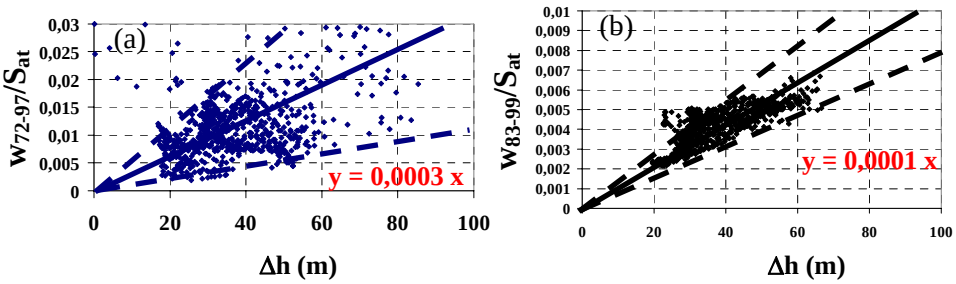


Figure 7. Settlement per unit fine grained soil thickness versus water head reduction (a. period 1972-1997; b. period 1983-1999).

Such observation suggests to relate the ground surface settlements (w) (Fig. 6.a and b) at the different positions to the corresponding cumulated thickness of fine grained soils (S_{at}) (Fig. 3.c). The ratio between these two quantities, calculated on the same grid points previously adopted for the calculation of water head reduction, is expressed in Figs.7 as a function of the water head drop Δh (Fig. 5.c). From both plots, which refer respectively to two different periods of observation (Fig.7.a to 1972-1997; Fig.7.b to 1983-1999), a clear increasing trend is observed between these two quantities. The larger scattering of points on the former plot (±50% around the average line) compared to the second (±20% around the average line) can be partly considered as an effect of the different accuracy in the interpolation of settlements for the two periods. However, by comparing the average trend on these two plots, it is seen that the rate of settlement occurred in the period 1972-1983 has been almost four times larger than after 1983.

5. CONCLUSIONS

A quantitative back analysis has been conducted in order to study the ground subsidence induced in the Bologna area by intensive water exploitation. Multi-temporal large scale analyses have been performed combining into a Geographical Information System different types of information and large amount of data regarding subsoil conditions, water levels and measured settlements. Interpretation of the complex observed pattern of subsidence has been provided by finding a close relation between settlements, cumulated thickness of fine grained soils deposits and water head reductions. In particular, the analysis on these latter performed by means of numerical modelling have revealed the predominant role of water withdrawal from the two alluvial fans of Reno and Savena rivers, carried out for civil purposes. As a more general comment, the large amount of data collected in the past decades and the availability of information tools capable of quickly processing them on large geographical scales, allowed to make profit of the past experiences on such coupled hydro-mechanical phenomenon and to derive a lesson for future more balanced managements of the groundwater source.

REFERENCES

Alessi, R., La subsidenza nel centro storico della città di Bologna. Il grado di dissesto dei fabbricati nella zona di via Zamboni. Inarcos, n. 456, 1985.

Amorosi, A., Farina, M., Severi, P., Preti, D., Caporale, L. & Di Dio, G., Genetically related alluvial deposits across active fault zones: an example of alluvial fan-terrace correlation from the upper quaternari of the southern Po Basin, Italy, *Sedimentary Geology, 102, 275-295*, 1996.

Bartolini, C., Caputo, R., & Pieri, M., Pliocene–Quaternary sedimentation in the Northern Apennine Foredeep and related denudation, *Geological Magazine, 133(3), 255−273*, 1996.

Carminati, E., Martinelli, G., Subsidence rates in the Po Plain, northern Italy: the relative impact of natural and anthropogenic causation, *Engineering Geology, 66*, 241-255, 2002.

Darini, G., Land subsidence due to groundwater withdrawal: the case of Bologna, Ph.D. thesis, Università degli Studi di Cassino (Italy), 2007.

Elmi, C., Bergonzoni, A., Massa, T., Montaletti, V., Barattella, P.L., Ronchi, A., Il territorio di pianura del Comune di Bologna: aspetti geologici e geotecnica, *Giornale di Geologia, 46(2)*, 127-152, 1984.

Farina, M., Simoni, M., Passuti, I., Il complesso idrogeologico superficiale nel contesto della città di Bologna, *Il geologo, 11*, 2000.

Groundwater Vistas-Modflow, version 3, *Environmental Simulations Inc.*, 2001.

IDROSER, Studio sul fenomeno della subsidenza nel territorio bolognese. Comune di Bologna, 1989.

Matheron, G., The Theory of Regionalized Variables and Its Application – Spatial Variabilities of Soil and Landforms. *Fontainebleau, Les Cahiers du Centre de Morphologie Mathematique 8*, 1971.

Osservatorio Astronomico di Bologna, 2007, from http://www.bo.astro.it website.

Pieri, L., Russo, P., Situazione attuale delle ricerche sull'abbassamento del suolo nel territorio bolognese. *Inarcos, 456*, 1985.

Regione Emilia Romagna and ENI-AGIP, Riserve idriche sotterranee della Regione Emilia-Romagna. S.EL.CA. (Firenze), 1998.

Regione Emilia Romagna, 2007, from http://www.regione.emilia-romagna.it website.

Regione Emilia Romagna – UnionCamere, 2007, from http://www.rer.camcom.it/datiecon website.

USGS, Land subsidence in the United States, *edited by Devin Galloway, David R. Jones, S.E. Ingebritsen*, 1999.

Vassena, C., Sviluppo e applicazione di modelli di flusso delle acque sotterranee nella conoide alluvionale del fiume Reno, Ph.D. thesis, Università degli Studi di Milano, 2003.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

TO EARTHQUAKE PREPARATION MODELS, BASED ON HYDRO-GEOCHEMICAL DATA

F.H.Karimov, N.Gh.Salomov, R.U.Vakhobova, G.B.Norova
Institute of Earthquake Engineering and Seismology
of the Academy of Sciences of the Republic of Tajikistan, Dushanbe, Tajikistan
(seismtadj@mail.tj)

***Keywords**:* Earthquake; preparation model; hydro geochemical researches

INTRODUCTION

Dynamic properties of underground waters give a basis for monitoring Earth crust geodynamics' processes. Hydrogeochemical data have been analyzed from the thermal water of Khoja-Obigarm, Yavroz, Shaambary in the Tajikistan seismic active areas on the determination of radon, pH value and electrical conductivity.

ARRANGEMENT OF THE OBSERVATIONS AND THE RESULTS

The observations have been conducted at the of thermal water deposits Khoja-Obigarm, Yavroz, Shaambary in the Tajikistan's seismic active areas. The depths of bore holes were about 50, 100 and 1300 meters, average debits were equal to 1.50, 0.35 and 0.90 liter/s with hydro carbonate-sulfate-chloride-potassium-silicon, chloride-sulfate-potassium and sulfate-chloride-potassium waters respectively. The variation of the pH in particular for Shaambary site has shown very distinct periodicity of about 14 days, caused evidently by crust deformational processes under the lunar tides, and failure of this periodicity approximately 10 days before the Dushanbe earthquake on 18 August 2006, M = 4.5, hypocenter depth 2.5-5.0 km at the epicentral distance of 20 km (Fig. 1).

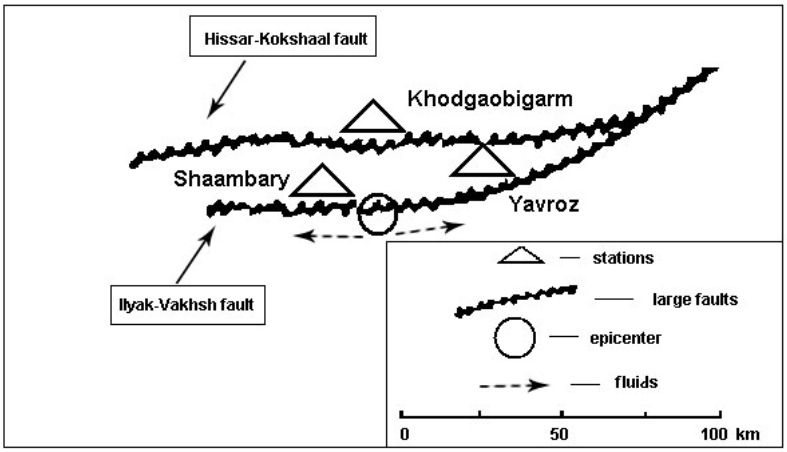


Figure1. Scheme of the observational site. The legend is at the right bottom corner.

The running average is shown by the bold line on the Fig.2. The positive long term trend stipulated seemingly by annual periodicity can be seen as well on the Fig.2.

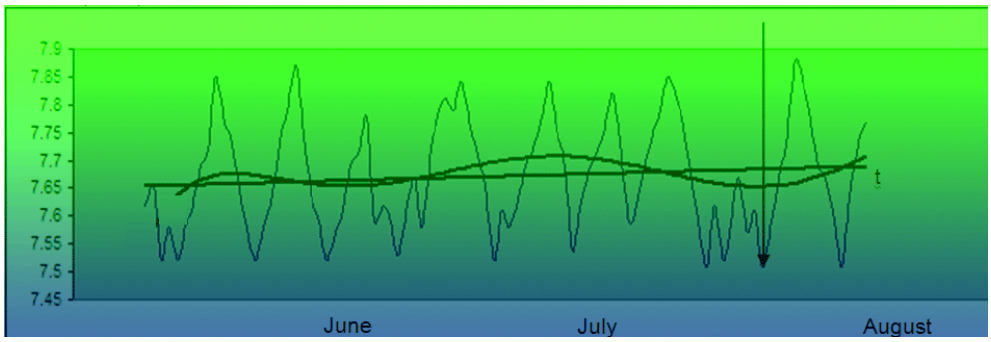


Figure 2. pH index of the Shaambary thermal water deposit in July-August 2006. The arrow shows the Dushanbe 18 August 2006 earthquake moment.

THE MODEL REPRESENTATIONS

Stresses near the focus and near to an earthquake preparation zone amount to about 10^9 Pa and they descend up to the background values within the radius r according to the formula, derived by Dobrovolsky [2000] –

$$\lg r[km] = 0.43M, \qquad (1)$$

where r is equal to about 100 km for the earthquake of magnitude 4.5.
In water samples of Khoja-Obigarm and Yavroz the pH anomalies contained some noise, probably because of the superficial chinks caused by infiltrated underground waters from atmospheric precipitation and brooks flowing nearby. Clear Shaambary site effect may confirm that elision waters play the prevailing role here. The epicenter of the earthquake is near Ilyak-Vakhsh fault, passing through the same line the Shaambary is located on (Fig.1). The fault zone represents the porous and cracked rocks and consequently the fluids can propagate through them up to long distances of hundreds to thousand kilometers. The fluid flows along the fault are shown by the dashed arrows (Fig.1).
The area of preparation of a tectonic earthquake depends on its magnitude and is described by means of the formula (1) set forth. The linear size of an earthquake source can be determined by the empirical formula, derived by Riznichenko [1976]

$$\text{Log } R_o\,[km] = 0.42M\text{-}1.37, \qquad (2)$$

where R_o is the mean radius of earthquake source.
In the case of the Dushanbe earthquake according to the formula (1) the preparation area radius is around 100 kilometer and the mean radius of the source (2) is approximately 2.5 kilometer.
Ability of fluids to extend along within a fault from area of epicenter formation on such distances has been found confirming not only by the hydro geochemical anomalies, but, as was shown by Rikitake [1976], also for the anomalies of some other geophysical fields. For example, in variations of a local geomagnetic field, obtained in particular by Skovorodkin et al. [1978], Johnston [1997] and Karimov [2007]. According to the general representations about seismic tectonic process of Dobrovolsky [2000], the integrant stress in the earthquake preparation zone displaying mean term stage has a bay-like shape. The main shock pertains to the stage of approaching of anomaly to this level. This indicates background tectonic movements of tectonic plates and geoblocks, pressing and colliding on each others.
One can observe in investigations of pH, Rn and electro conductivity parameters the time runs with periodicities about 14 days and 12 hours. 14 day period of the pH index run is shown in the Fig.2. The detailed work has been carried out to identify daily periodicity in chemical composition and conductivity of underground waters under sampling with the discreteness of 1 hour by the order of magnitude. The periodicity of pH variation for a day has been observed, but conductivity has more pronounced effect. About 12 hours periodical anomalies vary in amplitudes up to 20 mSm/m at Khoja-Obigarm, 60 mSm/m at Yavroz and 160 mSm/m at Shaambary.

The Fig.3 demonstrates 12 hour averaged periodicity data in electro conductivity of waters extracted from bore holes Khoja-Obigarm, Yavroz and Shaambary. This periodicity is superimposing on to the background trend stipulated seemingly by lunar monthly run.

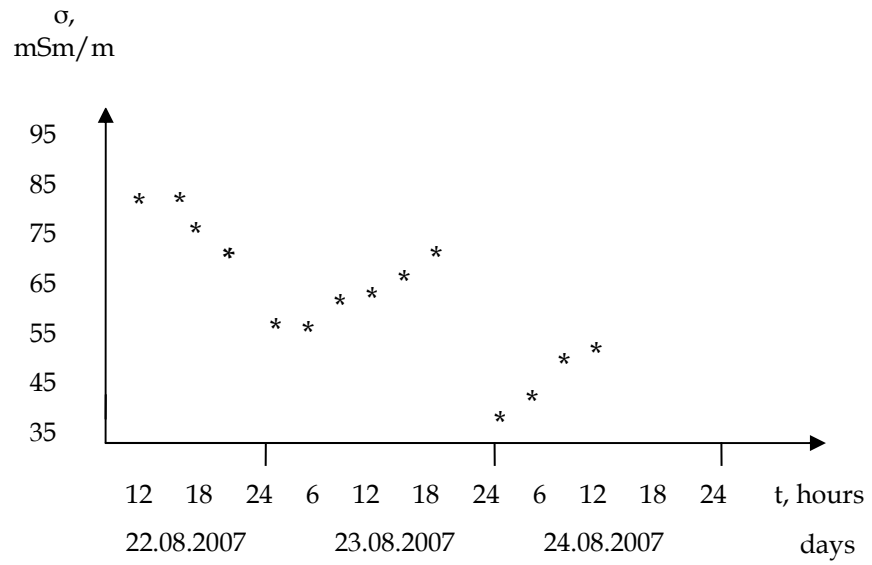


Figure 3. 12 hour periodicity in electro conductivity of water extracted from Yavroz borehole.

INTERPRETATION AND DISCUSSION

The following interpretation for such periodical variations in pH index and electro conductivity is suggested on the basis of model representations for seismotectonic processes of Dobrovolsky [2000]. The noise represents the vibrations of earth's crust The observations on lunar and solar tides by Starkov et al. [1988] conducted in the Tajikistan's seismic regions show that relative crust deformations varies between 10^{-9} to 10^{-8} in bulk parts of geoblocks and up to 10^{-6} in geoblock and fault borders, where the wings are less constrained. Relatively large crust deformations on fault vicinities promote active fluid flow dynamics along with the faults, which contain rocks with high porosity relatively to bulk geoblock rocks. So in a regular state the geoblocks are oscillating under the regular periodical tidal impact and fluids are propagating through the porous faults with the same periodicities.

Once any adjacent geoblocks have entered into cohesion there will be the inhomogeneous "inclusion" nucleated and it get started to grow in volume. The tectonic movements at the fault between geoblocks stop and therefore movement of the fluid flow fills up the local space. The energy pumping into the inclusion continues to grow and starting from critical barrier the "inclusion" begins to collapse. When hardness of the "inclusion" descends to low levels the next tidal impact will tire the geoblocks' cohesion sharply and thus the earthquake originate. Afterwards the geoblocks again start vibrating with the regular background noise level of the crust.

REFERENCES

Dobrovolsky, I.P. Theory of the tectonic earthquake preparation, UIPE RAS, 133 pp., Moscow, 2000.

Karimov, F.H. Tectonomagnetic effects in the Tajikistan's seismic regions. *Vestnik Rossiiskogo Universiteta Drugby Narodov /Bulletin of the Russian University of Friendship of Peoples, Physics*, 3-4, 162-169, 2007.

Johnston, M.J.S. Review of electric and magnetic fields accompanying seismic and volcanic activity. *Surveys in Geophysics*, 18, 441-475, 1997.

Rikitake, T. Earthquake prediction, Elsever, 357 pp., New York, 1976.

Riznichenko, Yu.V. A crust earthquake source sizes and seismic moment. In Researches on earthquakes physics, Nauka, 9-27, Moscow, 1976 (in Russian).

Skovorodkin, Yu.P., Bezuglaya, L.S., Guseva, T.V. Tectonomagnetic studies in the Tajikistan, *J.Geomagnetism and Geolectricity*, *30*(5), 481-486, 1978.

Starkov, V.I., Latynina, L.A., Karmaleeva, R.M. et al. The parameters of tidal deformations in Djerino site from the results of 19 years observations. *In Prognoz zemletryasenii*, 9, Donish, 65-77, Dushanbe-Moscow, 1988 (in Russian).

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Use of Two Nonparametric Methods (CART and MARS) to Model the Potential Distribution of Gullies in Spanish Rangelands

Á. Gómez Gutiérrez, S. Schnabel & F. J. Lavado Contador.

GeoEnvironmental Research Group (GIGA), University of Extremadura, Avda. de la Universidad, s/n, 10071, Cáceres, Spain (alvgo@unex.es)

Abstract: Gully erosion represents an important soil degradation process in rangelands. In order to take preventive or control measures and to reduce its environmental damages and economical costs it is useful to localize the points in the landscape where gullying takes place and to determine the importance of the different factors involved. The study is carried out in Extremadura, southwest Spain. The main objectives of this work are: 1) to model the spatial distribution of gullies, 2) comparison of two nonparametric schemes to construct predictive models 3) to analyze the role of prevalence and of scale for susceptibility models in geomorphology, 4) evaluating the importance of the different factors involved and 5) implementing and mapping the results with the help of a Geographical Information System (GIS). Two methods were used to model the response of a dependent variable (gullying) from a set of independent variables: Classification And Regression Trees (CART) and Multivariate Adaptive Regression Splines (MARS). Three different datasets were used; the first one for constructing the model (training dataset) and the others for validating the model (external datasets). These datasets are formed by a target variable (presence or absence of gullies) and a set of independent variables. The dependent variable was obtained mapping the locations of gullies with the help of a GPS and high resolution aerial ortophotographs. We used 32 independent variables reflecting topography, lithology, soil type, climate and land use and vegetation cover of each area. To evaluate the performance of the models we used a non-dependent threshold method: the Receiver Operating Characteristic (ROC) curve. The results show a better performance of MARS for predicting gullying with areas under the ROC curve of 0.98 and 0.97 for the validation datasets, while CART presents values of 0.96 and 0.66.

Keywords: Gully erosion; nonparametric modeling; CART; MARS; ROC Curve.

1. INTRODUCTION

A gully is a channel originated by concentrated and intermittent water flow. Gully erosion is one of the most important soil degradation phenomena in several environments around the world and has long been neglected because it is difficult to study and to predict [Valentin *et al.* 2005]. Nevertheless, the negative on-site and off-site effects of gullying are well known: soil loss, reduction of soil water retention, landscape dissection (hampering the movement of vehicles, farm machinery and animals), water contamination and, sedimentation of channels and reservoirs. Several studies have tried to model the location of gullies in the landscape based on the concept of topographical thresholds [Patton & Schumm 1975]. However, these models are difficult to apply at the regional scale with success. In this line, Gómez Gutiérrez *et al*. [2007] developed a statistical multivariable model using MARS to predict the potential distribution of gullies in rangelands of southwest Spain. In the present paper, we compare the results of this previous work with another statistical model,

Classification And Regression Trees (CART). A further objective is to analyze the role of prevalence (different proportions of presence/absence data in the dataset) and of scale for susceptibility models in geomorphology. The importance of the different factors involved is evaluated and the results are implemented and mapped with the help of a Geographical Information System (GIS).

2. STUDY AREA

The study was carried out in 54 farms, representative of *dehesas* and pasturelands in the Iberian Peninsula (Figure 1). Dehesas commonly have a savannah-like vegetation and agrosylvopastoral land use, with farm sizes in excess of 100 ha. The relief is diverse with an average elevation of 413 m (ranging from 115 m to 902 m) and a mean slope angle of 5º. Lithology is quite diverse, though acid rocks dominate, being schists and granites the most frequent. Soils are mainly Cambisols (80%), Luvisols (10%) and Acrisols (5%). Climate is Mediterranean with continental and Atlantic influences. Mean annual temperature is approximately 16º. Precipitation presents a high spatial and temporal variability. Median annual rainfall of the study farms is 607 mm, ranging from 465 to 926 mm. More information on physical conditions of the 54 farms can be found in Tables 2 and 3.

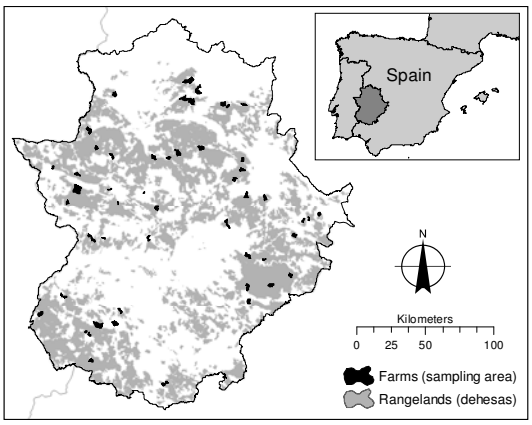


Figure 1. Location of the farms and the extension of dehesas in Extremadura region.

3. MATERIAL AND METHODS

3.1 CART and MARS

Classification And Regression Trees [CART; Breiman *et al.* 1984] is a popular data mining technique based on recursive binary partitioning. The result of CART (1) is a hierarchical binary tree which subdivides the prediction space into regions (R_m) where the values of the response variable are similar ($\cong a_m$):

$$f(x) = a_m ; \forall x \in R_m \quad (1)$$

Basically, CART predicts the pertinence of cases to a category of the dependent variable from a set of values of the independent variables. CART presents a root node with all the cases and analyzes if these belong to a single class. If the answer is yes, the algorithm finishes, but if it is no, CART divides the original dataset based on a split criterion (the Gini index of impurity) into two classes (the one which fulfills the split criterion and the rest of the dataset). The performance of the algorithm continues until a stopping rule is completed. The objective of the split criterion is to minimize the variability within each of the resulting subsets. The principal inconveniences of CART are the possible complexity of the resulting model, the hierarchical dependence between nodes at different levels and the difficulties that CART presents to reproduce smooth variations on the response. The model based on CART theory was computed using the Classification and Regression Tree tool implemented in *Statistica 7*© [Stasoft-Inc. 2004]. In this software, the variable importance is computed through the sum of the reductions in node impurity over all nodes in all computed

trees when a variable is eliminated. Then, these values are referenced as percentages of the most important variable.

On the other hand, MARS [Friedman 1991] is a relatively new technique which combines the classical linear regression, the mathematical construction of splines, the binary recursive partitioning and brute search intelligent algorithms to produce a model capable to predict the value of a target variable (categorical or continuous) from a set of independent variables. To do this, MARS approaches the underlying function through a set of piecewise functions called basic function (BF). The BF represent the information included in one or more independent variables and are selected through a step by step process. MARS general expression can be written as follows:

$$y = f(x) = \beta_0 + \sum_{m=1}^{M} \beta_m h_m(x) \qquad (2)$$

where y is the value predicted by the model by means of a function $f(x)$, which can be decomposed into an initial constant β_0 and a sum of M terms, each of them are formed by a coefficient β_m and a BF $h_m(x)$. The first step of MARS is to approximate its response with a constant ($y=a$). Then, MARS starts the search of the couple node-variable that produces the best fit of the model. This search continues until all possible BF have been added to the model (or a stopping rule is completed), obtaining as a result a very complex and overfitted model. Based on the overfitted model, MARS identifies which is the less important BF which is eliminated. This process is repeated until all BF are suppressed. The outcome is a sequence of candidate models with different BF numbers, from which MARS selects the optimal model via Generalized Cross Validation [GCV; Craven & Wahba 1979]:

$$GCV(M) = \frac{1}{N} \frac{\sum_{i=1}^{N} (y_i - f_M(x_i))^2}{\left(1 - \frac{C(M)}{N}\right)^2} \qquad (3)$$

where N is the number of cases, y is the dependent variable, $f_M(x_i)$ represents the model and $C(M)$ is a measure of cost-complexity for a model with M terms. The software programme MARS 2.0© [Salford-Systems 2000] was used for modeling. It enables the estimation of the importance of the independent variables. In order to determine the importance of an independent variable, MARS drops all terms including this variable and calculates the reduction of the goodness of fit. Finally, MARS references these values to the most important variables as a percentage.

3.2 Evaluation and validation of the models

In this paper, gullying was treated as a binary variable (1: presence and 0: absence). Therefore, the results of CART are easily comparable with reality. However the output of MARS is a continuous value, normally between -0.5 and 1.5 for a target variable encoded as 0 or 1. To evaluate the results of predictive models for categorical or binary target variables traditionally a confusion matrix was used. This matrix consists in crossing the predicted and the observed values. However, the confusion matrix and its derived statistics (misclassification rate, model efficiency, odd ratio, etc.) depend on a predefined threshold which divides the output into categories (1,0). However, these are influenced by prevalence of the dataset [Beguería 2006], typical in geomorphology studies. Therefore the use of a method is recommended which is not based on a predefined threshold and with an external dataset different from the training dataset [Muñoz & Felicísimo 2004; Beguería 2006].

The results of the models were analysed using the Receiver Operating Characteristic curve [ROC; Deleo 1993] and the Area Under the Curve (AUC). The ROC curve shows the skill of a model to discriminate presences from absences (representing the complementary of specificity versus sensitivity). A model with a good performance should present a ROC curve fitted to the upper-left side of the graph and the AUC should be a value close to one.

Two different external datasets were used to validate the model: *Guadalperalón* and *Monroy* (Table 1). These datasets were obtained in areas having similar environmental characteristics than the areas of the training dataset.

Table 1. Characteristics of the validation areas.

Dataset	Area	Elevation	Slope	Lithology	Soil types	Tree Cover	Annual Rainfall
Guadalperalón	131 ha	353 m	5.47 º	Schists	Distric Cambisol	30.2 %	595 mm
Monroy	114 ha	397 m	2.01 º	Schists	Distri-endolep. Cambisol	30.0 %	601 mm

3.3 Elaboration of the dataset

The general scheme of this part of the work was to generate a map showing the presence or absence of gullies in the study area and to obtain maps of different factors related with the gully formation. The presences of the target variable were obtained mapping existing gullies in the field with a GPS and the absences were obtained with a random point generator tool implemented using the Hawth's Analysis Tools [Beyer 2002] of ArcGis 9.0 [ESRI 1999]. A total of 662 Presences were mapped, as compared to 7001 Absences.

Different types of independent variables were used (Tables 1 and 2). To represent the topography a Digital Elevation Model (DEM) was used, with a pixel size of 5 m, extracted from contour lines and elevation points of the topographical map of Spain (scale 1:10,000) using the *Topo to Raster tool* [Hutchinson 1993] implemented in ArcGis 9.0 [ESRI 1999]. With this DEM we generated derived Digital Terrain Models to represent slope, curvature, ruggedness and some hydrological features of the surface (Table 1). Soil type was extracted from maps of 1:300,000. This small scale represents a handicap, but it is the only soil map available for the whole study area, and we think that it can represent regional trends in some soil types for the development of gullies. Lithology is based on the digital version of the geological cartography of Spain (scale 1:50,000). Maps of mean annual, seasonal and monthly rainfall were elaborated using a dataset (1960-1990) from 222 meteorological stations in the study area, [using as interpolation method thin plate splines; Hutchinson 1991]. Finally we used five variables from the forest map of Extremadura [scale 1:50,000, EGMASA 2002]: land use and management of the farms, dominant tree species, tree cover, total vegetation cover and structure of the vegetation cover.

Table 2. Quantitative independent variables used and their values for the study area (training dataset).

Variable	Average	Deviation	Units
Elevation	413	125	Meters
Slope	5.06	5.60	Degrees
Curvature	0.00	1.15	10^{-2} meters
Vertical curvature	0.01	0.56	10^{-2} meters
Horizontal curvature	0.01	0.81	10^{-2} meters
Total upslope length	2677	23,046	Pixels
Longest upslope length	80.0	218.7	Pixels
Drainage area	1,057.9	9,093.0	Pixels
Total vegetation cover	76.3	25.4	%
Tree cover	27.1	17.2	%
Monthly rainfall	(12 variables)	-	Millimeters
Seasonal rainfall	(4 variables)	-	Millimeters
Annual rainfall	695.6	145.3	Millimeters

The effect of prevalence was analysed varying the number of absences on the training dataset and maintaining the same number of presences and calculating the success of the resulting models using the AUC.

Table 3. Qualitative independent variables used in the model.

Variable	Number of categories	Dominant categories
Lithology	139	Schists and granites
Soil type	47	Distri-endoleptic Cambisols and distric Cambisols
Soil family	8	Cambisols and Luvisols
Vegetation structure	63	Mixed woodland and low shrubs
Land use	8	Pasturelands
Dominant tree specie	15	*Quercus rotundifolia* Lam.

4. RESULTS AND DISCUSSION

4.1 CART

The result of CART is a tree with 34 non-terminal nodes and 35 terminal nodes. Table 3 shows part of the model classification rules. This model presents an AUC of 0.97 for the training dataset. The values of the AUC for the validation datasets are quite different, while *Guadalperalón* presents a high value (0.96), the AUC of the *Monroy* dataset is only 0.66.

Table 4. CART model classification rules. These expressions must be compiled and implemented on a GIS to obtain susceptibility maps for gullying.

Terminal node	Classification rule
4	If [(lithology=9 or 11 or 40 or 80 or 99 or 109) and (Soil type=0 or 11 or 16) then the pixel is classified as ungullied area (0).
…	…
54	If [(lithology ≠ 9 or 11 or 40 or 80 or 99 or 109) and (vegetation structure = 22) and (lithology = 84 or 90 or 127) and (planimetric curvature ≤ 0) and (longest upslope length grid ≤ 225.13) and (elevation >251) and (rainfall in 4th trimester ≤ 343) and (december rainfall >84.5) and (elevation ≤ 474.5) and (elevation ≤ 380.5)] then the pixel is classified as gullied area (1).
…	…
123	If [(lithology ≠ 9 or 11 or 40 or 80 or 99 or 109) and (vegetation structure ≠ 22) and (rainfall in 3rd trimester >35.5) and (soil type ≠ 18) and (lithology ≠ 20 or 85) then the pixel is classified as ungullied area (0).

Regarding the influence of independent variables, elevation (DEM) and annual rainfall seem to be the most important variables. In addition, lithology and soil type are observed as root node or nodes in the upper part of the tree structure. Furthermore, variables related with topography are situated in the lower part of the tree. This trend seems to be related with the scale of the independent variables (Figure 2). Since the variables were obtained from maps with only three different scales (1:10,000, 1:50,000 and 1:300,000), we can deduct from the tree structure that the quality and resolution of the model is supported by large scale variables (elevation, slope, curvature, drainage area, etc.) due to their position in the lower part of the tree. This argument could be an important key in the future, because the tree structure shows us which variables must be elaborated carefully (with high resolution and precision), and this will probably improve the success of the model for predicting the location of gullies in similar environments. In the case of our study the variables in the upper part of the tree indicate areas with a risk of gully formation and the variables in the lower part of the tree which have higher spatial resolution enable the location of the gully.

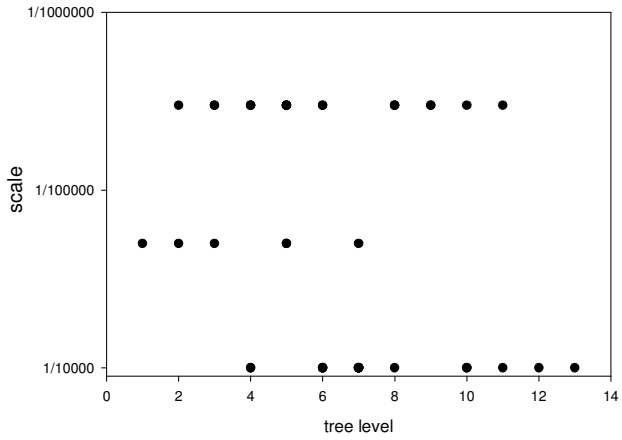


Figure 2. Relation between the scale of each cartographic source and the position where its derived variables are located on the tree.

4.2 MARS

The outcome of MARS is a model with 78 BF, 57 terms and a ROC curve that is fitted to the upper left side of the plot. The AUC for the training dataset is 0.98. Using a preliminary threshold of 0.5 we get a sensitivity of 0.953 and a specificity of 0.86. For the *Guadalperalón* validation dataset the AUC reaches a value of 0.98. For the *Monroy* validation dataset the obtained AUC is similar (0.97), showing a good performance of the model. A portion of the MARS model for predicting gullying in the study area is as follows:

$$Y = 0.065 + 11.973 * BF1 + ... + 0.317 * BF108 + \ldots + 0.000104 * BF120 \quad (4)$$

where *y* is the predicted value and BF1= (lithology = 9,11,40,60,80,99 or 109), BF108 = [(areas with rainfall in march below 73 mm – 73) * BF1] and BF120 = [(lithology = 40, 90 or 109) * BF4], and finally, BF4= [areas with a value of total upslope length bellow 27186 pixels - 27186].

MARS used 27 variables to construct the model. Individually, the most important variable is lithology followed by vegetation structure, rainfall in the 3rd trimester of the year and elevation. By groups, we can observe that climatic variables are the most important in absolute terms, principally due to their large number (17). However, if the variables are grouped (topography, soil type-lithology, climate, and land use-vegetation cover) the highest variable importance corresponds to lithology and soils. These results should be interpreted carefully because MARS and CART models can explain a portion of the spatial distribution of gullies, but cannot establish causal relationships between independent variables and gully occurrence [Donati & Turrini 2002].

4.3 The role of prevalence on the training dataset

No clear tendency was observed. The AUC values for training datasets with different proportion of presences/absences in the training dataset are unstable (Figure 3). Although some authors have attributed an important role of prevalence on the training datasets for modelling geomorphological phenomena, we have not found an influence on the success of the models. Possibly the independence of the AUC validation statistic from prevalence [Beguería 2006] and the quality of the presences and absences data can minimize this effect.

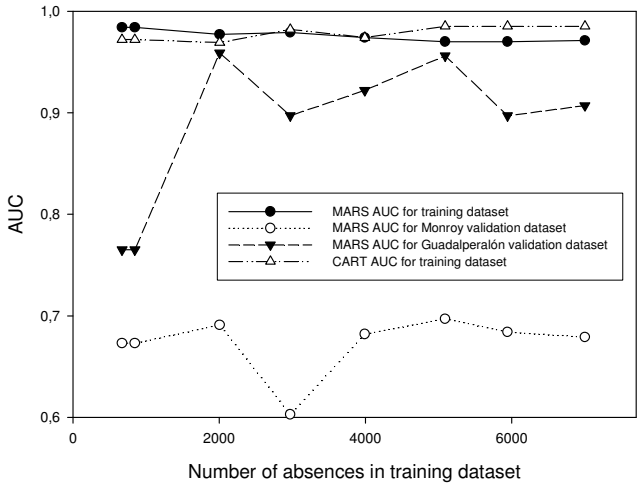


Figure 3. AUC values using different number of absences in the training dataset and with a fixed number of presences.

4.4 Implementing the models and mapping the results

After obtaining the models, they were implemented into a GIS to obtain maps of gully susceptibility. Finally, a composition map was elaborated combining the results of MARS and CART. For this the strategy consisted in using the AUC of each model as a weight for transforming the values and afterwards summing the results and generating the map.

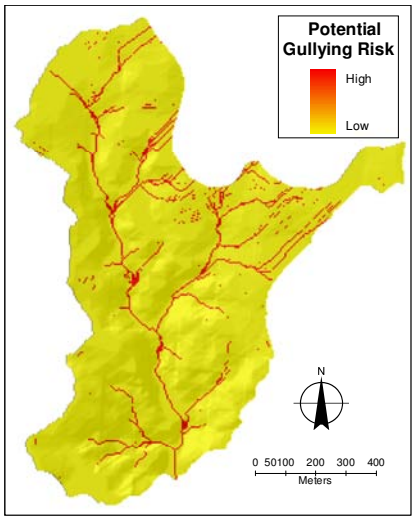


Figure 4. Composition map for the *Guadalperalón* validation area.

5. CONCLUSIONS

The two models present a good performance, with AUC from 0.66 to 0.98. MARS produced better results for the two validation datasets. Some differences between the two models regarding the importance of variables were found. Although variable importance cannot be interpreted as a causal relation, in CART the position of each variable in the tree can help us to understand the relation between variable scale, model structure and processes. Regarding the influence of prevalence on the success of the models, no clear

tendency was found. Finally we have used a simple method to create a map combining the results of both models and using the AUC values as weights. Susceptibility maps of gully erosion can be a useful tool for regional planners.

ACKNOWLEDGEMENTS

The work has been supported by the Junta de Extremadura and European Union funded project Montado/Dehesa (INTERREG IIIA, SP4.R13) and by the Spanish Ministry of Education and Science (CGL2004-04919-C02-02).

REFERENCES

Beguería, S. Validation and Evaluation of Predictive Models in Hazard Assessment and Risk Management. *Natural Hazards* **37**: 315-329, 2006.

Beyer, H. L. Hawth's Analysis Tools For ArcGis. In, 2002.

Breiman, L., Friedman, J. H., Olsen, R. A. & Stone, C. J. *Classification and regression trees*. Wadsworth, Belmont: CA. 1984.

Craven, P. & Wahba, G. Smoothing noisy data with spline functions. *Numerische Mathematik* **31**: 377-403, 1979.

Deleo, J. M. Receiver operating characteristic laboratory (ROCLAB): software for developing decision strategies that account for uncertainly. In: *Proceedings of the Second International Symposium on Uncertainlty Modelling and Analysis*, pp. 318-325. College Park: Computer Society Press, 1993.

Donati, L. & Turrini, M. C. An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: application to an area of the Apennines (Valnerina; Perugia, Italy). *Engineering Geology* **63**: 277-289, 2002.

EGMASA Mapa de vegetación y recursos forestales de Extremadura a partir del Mapa Forestal de España. In: *Mapa Forestal de Extremadura*, ed. C. d. A. y. M. Ambiente, 2002.

ESRI ArcGis 9.0. In, 1999.

Friedman, J. H. Multivariate adaptive regression splines. *Ann. Statist.* **19**: 1-141, 1991.

Gómez Gutiérrez, Á., Schnabel, S. & Felicísimo, A. Modelling the ocurrence of gullies in semiarid areas of southwest Spain. In: *Progress in gully erosion research*, eds. J. Casalí & R. Giménez, Pamplona: Public University of Navarra, 2007.

Hutchinson, M. F. Climatic analyses in data sparse regions. In: *Climatic risk in crop production. Models and Management for the semiarid tropics and subtropics*, eds. M. C. Chow & J. Bellamy, pp. 55-71. CAB International, 1991.

Hutchinson, M. F. Development of a continent-wide DEM with applications to terrain and climate analysis. In: *Environmental Modeling with GIS*, ed. G. M. F. e. al., pp. 392-399. New York: Oxford University, 1993.

Muñoz, J. & Felicísimo, Á. M. Comparison of statistical methods commonly used in predictive modelling. *Journal of Vegetation Science* **15**: 285-292, 2004.

Patton, P. C. & Schumm, S. A. Gully erosion: a threshold phenomenon. *Geology* **3**: 88-90, 1975.

Salford-Systems MARS. In, 2000.

Stasoft-Inc. Statistica: Data analysis software system. In, 2004.

Valentin, C., Poesen, J. & Yong, L. Gully erosion: Impacts, factors and control. *Catena* **63**: 132-153, 2005.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

On study of the geodynamic processes by geoelectrical methods.

A. Kamshilin [a] , E. Volkova [a] , V. Leontiev [a], O. Kuzichkin[b]
[a] *Institute of the Physics of the Earth RAS*
10, Bolshaya Gruzinskaya, Moscow, 123995, Russia (kamshilin@ifz.ru)
[b] *Muromsky Institute, 23 Orlovskaj, Murom, Vladimir Region, Russia (electron@mivlgu.ru)*

Abstract: Electrical methods with artificial generators are widely used in geophysical prospecting (electrical prospecting). Last decades electrical methods were applied for investigation of geodynamic phenomena for the purpose of prediction of different catastrophes: earthquakes, landslides, karsts etc. (geodynamic electric method - GDEM).
There exists the principal difference between electrical prospecting and GDEM. In the first case absolute values of geoelectrical parameters are measured, in the second case small time-space variations of geoelectrical parameters instead of the absolute values are measured.
Change in rock mass state results in distortions of electric field of generator. These distortions can be scalar or vector. Scalar distortions, as a rule, are connected with temperature effect, precipitations. Vector changes in geoelectrical fields reflect the processes of origin and development of inhomogeneity. These vector changes of an electric field are the subject of our investigations. A special measurement procedure and the special equipment are required for the solution of this problem.
The results gained by authors in the course different field and laboratory investigations are introduced in the present paper. The special method of the active geoelectric monitoring was used in all experiments.

Keywords: Geoelectical Monitoring; Dynamic components; Equipotential lines.

1. INTRODUCTION

The aim of our investigations is to study changes of stress-deformed state of medium by the method of active geoelectric monitoring. Using the system of emitting electrodes and electrical generator we can create initial electric field $\overline{E}_1$ in medium. Due to changes of stress conditions in medium time-space character of its inhomogeneity medium will be changed. In this case initial field $\overline{E}_1$ will be changed by the value $\Delta\overline{E}$

$$\overline{E}_2 = \overline{E}_1 + \Delta\overline{E} \qquad (1)$$

where $\overline{E}_2$ – field in the changed medium, $\Delta\overline{E}$ – difference field.
It is possible to consider $\Delta\overline{E}$, as the field of some dummy source, which corresponds to multiplicative function of inhomogeneity and real energy source and is connected with them spatially and quantitatively. Difference field $\Delta\overline{E}$ includes component, orthogonal to initial field E_1. These ideas became the basis for development of measurement procedure.

2. MEASUREMENT METHODS

The following requirements are necessary for measurement of dynamic components of geophysical parameters by a geoelectric method.

- Method shall provide maximal sensitivity to difference field $\Delta\overline{E}$ and minimal sensitivity to field $\overline{E}_1$.
- It is necessary to ensure low sensitivity to temperature, precipitations etc. Otherwise it is necessary to install electrodes on considerable depth.
- Correct choice of the relative position in the system source - inhomogeneity- sensor shall ensure the given directional diagram. It will allow to control the chosen object (the active fracture, for example) or to detect position of the new inhomogeneities.

The first and the second requirements are satisfied, if measuring electrodes are located on equipotential lines of initial field $\overline{E}_1$. In this case the initial potential difference between measuring electrodes MN is equal to zero. Noise factors, uniformly distributed throughout the area (temperature, precipitates, air pressure etc.), cause scalar variations of the field $\overline{E}_1$. In this case the shape of equipotential lines does not vary and $\Delta\overline{E}$ is equal to zero.
The third requirement is satisfied by installation of measuring electrodes along the beam coming from a dummy source.
Electrical installations with location of measuring electrodes on equipotential lines of field $\overline{E}_1$ are termed as the equipotential.installations. They can have different configuration. We often used three-electrode installation with two mutually orthogonal pairs of meaeuring electrodes («three-electrode equipotential two-dimentional installation»). Measuring pairs M_1 N_1 and M_2N_2 were located on equipotential lines of emitting electrode A. The second emitting electrode B was on the considerable distance.

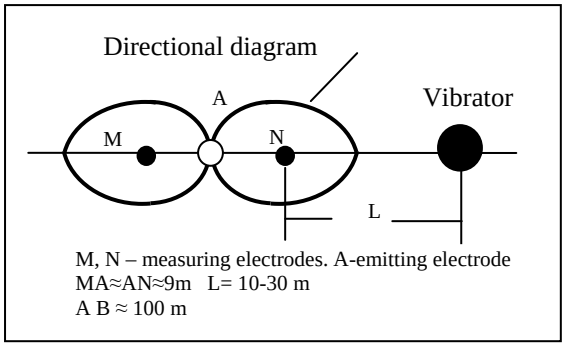


Figure 1. The experiment plan

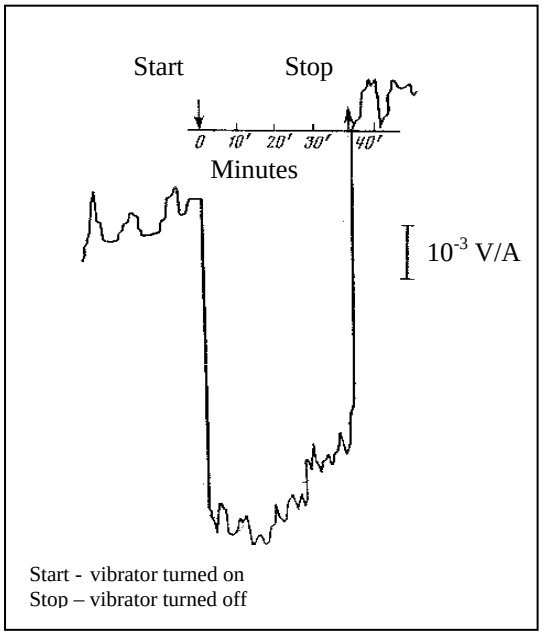


Figure 2. Result of experiment with vibrator

3. RESULTS OF EXPERIMENTS

3.1 Field Experiments with the Seismic Vibrator

Scheme of experiment is shown in Figure 1.
The used measurement instrument was the electrovariometer (EV). This device was designed for long-term continuous automated measurements of variations electrical parameters of rocks [A.N. Bogolubov et al., 2002].

EV consists of the generator, measuring unit, emitting electrodes AB and one or several pairs of measuring electrodes MN. The generator connected with the Earth through AB electrodes emits a low-frequency signal (73 Hz), the measuring unit receives active and reactive components of voltage between measuring electrodes (Re U_{MN} and Im U_{MN}). One

pair of measuring electrodes MN was oriented in direction of the vibrator. The second pair of measuring electrodes was installed orthogonally the first one (it is not shown in Figure 1).

The result of one experiment (modification Im U_{MN}) is shown in Figure 2. Frequency of the vibrator was 4 Hz, a time of work was 40 minutes. Im U_{MN} value was normalized to generator current I_G. Im U_{MN}/I_G value was changed by $5.9*10^{-3}$ V/A after vibrator was turned on (Start) and was recovered after vibrator was turned off (Stop). Re U_{MN}/I_G value was changed in the same way. Polarity of the effect in all experiments remained invariable; the dependence on vibrator frequency was not detected. The signal on orthogonal pair MN was not detected.
These results were the consequence of inhomogeneous, reversible medium change under the influence of the vibrator. We believe that it is the complex of changes of stress condition and electrical conductivity of medium under the influence of elastic vibrations This problem needs the further investigation.

3.2 Natural experiments

During several years we conducted field observations in seismically and landslide active regions using the three-electrode equipotential installation. As field experiments have shown, this method allows to decrease co-phased electrical interference by three orders of magnitude in comparison with Venner installation. The influence of the near-surface layer conductivity caused by temperature change was decreased by two orders and more. A direct effect of precipitation was not detected.

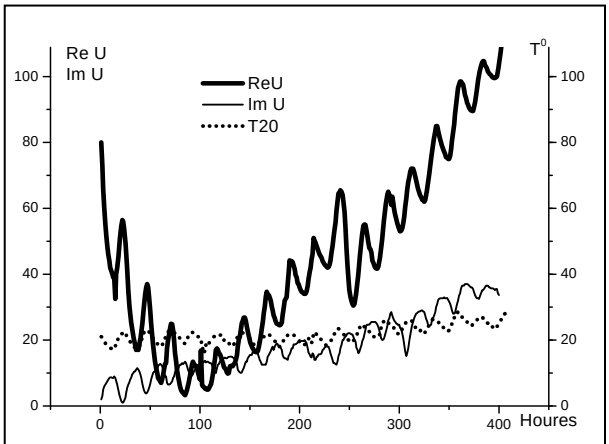


Figure 3. Synchronous measurments of Re U_{MN}, Im U_{MN} and soil temperatures

During the observations, quasiperiodic oscillations ("equipotential line breathing") with the main periods of approximately 24, 12 hours were registered. Amplitudes of the oscillations with different

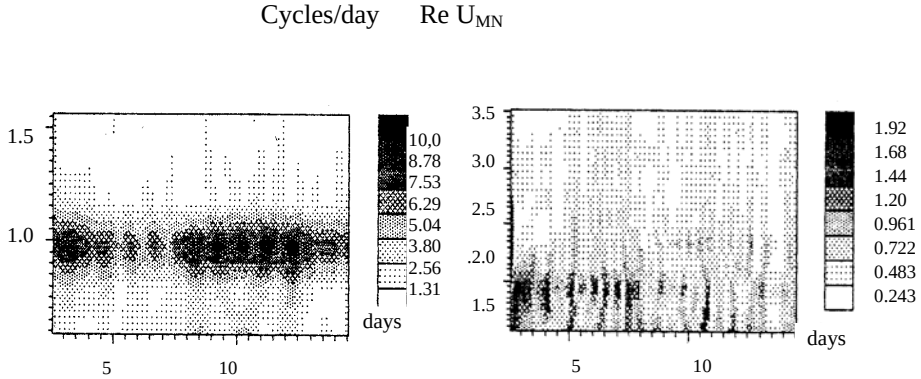


Figure 4. ST diagram of Re U_{MN}

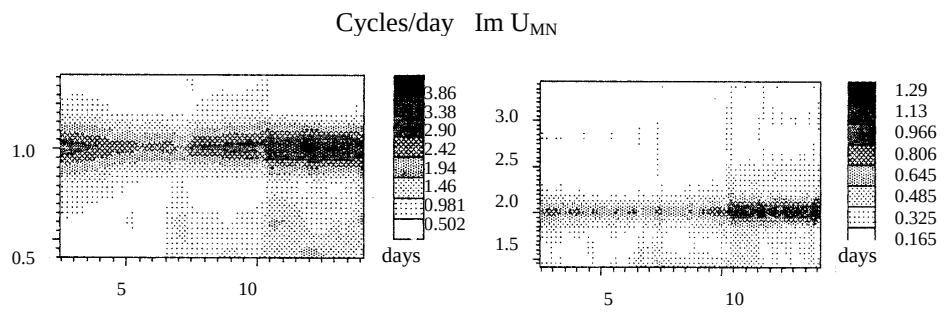


Figure 5. ST diagram of Im U_{MN}

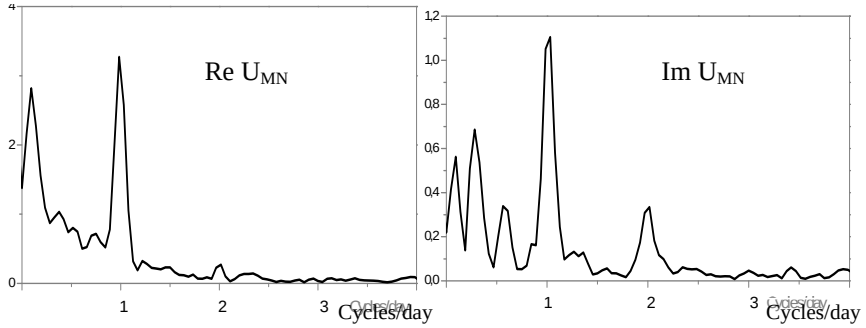


Figure 6. Spectra of Re U_{MN} and Im U_{MN}

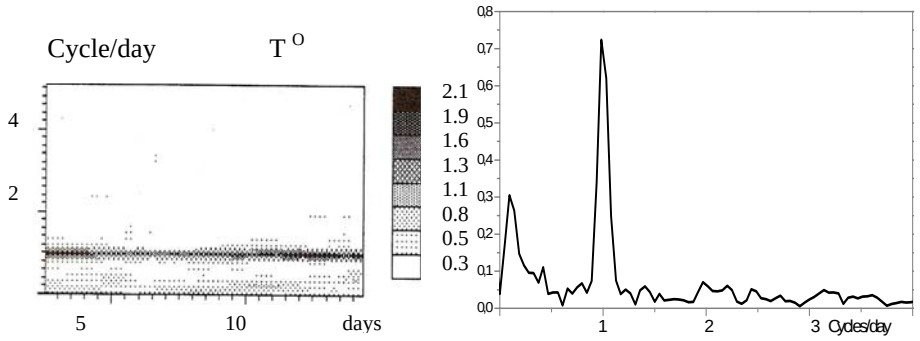


Figure 7. ST diagram and spectrum of T^O

periods varied in the time. This behavior (regime) corresponded to the “quiet” condition of the medium (the uppermost layer of the Earth’s crust). The disastrous geodynamic events (earthquakes and landslides) were accompanied by the destruction of a "normal" trend several hours before these events. Oscillations were restored several hours after these events. . These results are presented in A. N. Bogolubov et al., [2002] and Volkova E.N. et al. [1998].

The results of synchronous measurements Re U_{MN}, Im U_{MN} and soil temperatures at depth of 20 cm are shown in Figure 3. Spectral-time diagrams and Fourier spectra of these curves are shown in Figures 4, 5, 6 and 7. Harmonic components of intraday periodicity are

detected in all series. Amplitudes of diurnal and semidiurnal oscillations for Re U_{MN} are differed several times. The amplitude of the semidiurnal oscillations Im U_{MN} is comparable with amplitude of diurnal oscillations, and it was noted early [A. N Bogolubov et all., 2002]. A temperature series contains only one dominating diurnal period.

It is most likely that variations Re U and Im U were caused by geodeformation processes.. Diurnal oscillations have both "temperature" and "tidal" origin, the semidiurnal harmonics are the result of lunisolar influence. This assumption was partially confirmed by model experiments in the lab.

3.3 Modeling experiments

The experimental equipment was designed by Dr. Victor P. Khomenko [Khomenko, 2003]. The working camera of installation (95 x 74 x 30 sm) was filled with sand. At a definite moment, after the bottom of the chamber started gradually to open, an enclosed cavity was formed in the chamber due to the sand spilling out through the partially opened bottom (the cavity did not appear on the surface of the sand).

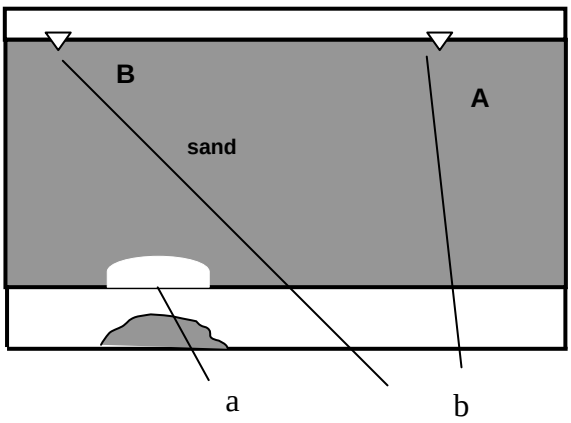


Figure 8. Location of the cavity and measuring equipment. a – cavity, b - emitting electroedes A and B

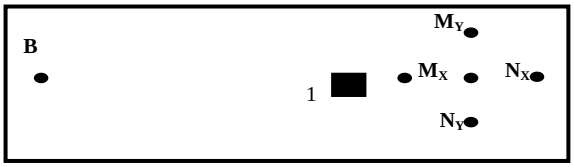


Figure 9. Location of the tilt-meter (1) and two-dimentional eqiupotential installation on the surface of the model

The possibility of registration of process of formation of the enclosed vacuity was studied by the method of geoelectric monitoring in the course of these experiments. Attempts of our colleagues to make it by means of traditional geoelectric and seismic methods have failed.

We used different versions of the equipotential installations, including two-dimensional eqiupotential geoelectrical device. In some experiments the tilt-meter (НИ-3, designed at IFE RAS) was installed on the surface of the sand (Figure 8, 9).

The formation of cavity was fixed in all experiments (more than ten experience were conducted). The typical result of experiment with 600 mm sand layer is shown in Figure 10. The background noise was measured from 0 till the moment A. A is the beginning of opening of metal bottom of model. Noise has appeared after opening of a metal bottom at this moment.
The opening in bottom of model was increased from the moment A till the moment C. Cavity (height 15 sm., width 25 sm.) appeared at the moment B. Cavity was increased at the moment C (height 18 sm., width 35 sm.). The moments of formation and increasing of the cavity were clearly registered by the measuring device.

The use of a two-dimensional eqiupotential installation allowed to define direction to cavity. The positive slope of U_{MxNx} diagram till the moment of cavity formations (moment B) was observed during experiment. The same slope was observed after a moment B.
It means, that distortion of equipotential lines of an electric field took place. We have assumed that there took place deformation of surface of model caused by continuous increase of opening in bottom of model.

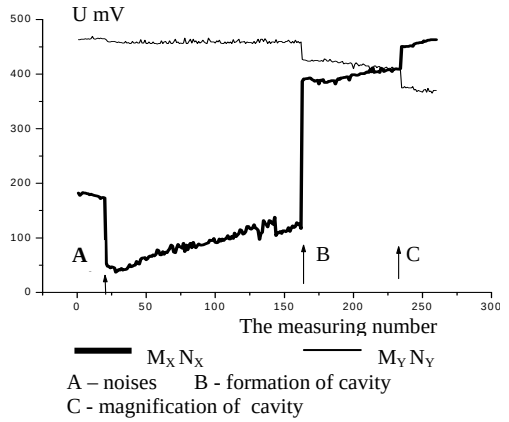


Figure 10. Detection of enclosed cavity by the two-dimensional eqiupotential installation

This assumption was checked by synchronous electrical and tilt measurements (Figure 11).
The background noise was recorded during 10 minutes (till the moment A). After that the bottom of installation started to move apart. The positive trend was observed on both diagrams. Response to formation and development of the cavity is also observed on both diagrams. Tilt-meter data show formation of the cavity was preceded by deformation of the surface of the sand in the working camera.

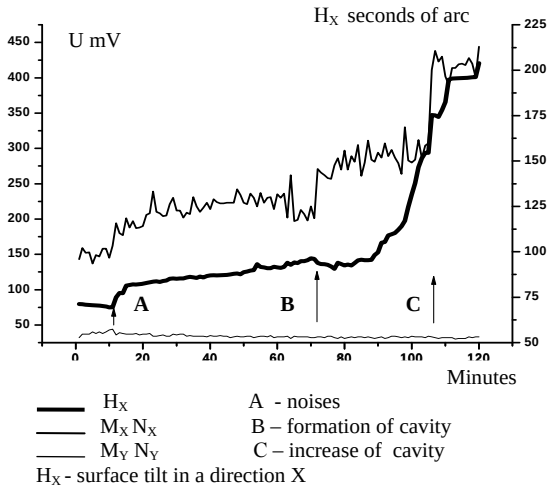


Figure 11. Synchronous measurement of an electric field and declinations of a surface of model.

4. RESUME

Investigations in the seismically and landslide active regions and model experiments show that the equipotential installations have high resolution ability and a high noise immunity.

We could compare behavior of U_{MN} during formation of cavity (Figures 10 and 11) and during work of vibrator (Figure 2). The nonreversible changes caused by formation of inhomogeneity (cavity) took place in the first case. In the second case changes of U_{MN} are similar, but reversible. So, it is possible to consider that vibrator action is equivalent to appearance of inhomogeneity, which disappears after vibrator stops working. So, experiments with vibrator and model experiments confirm that two-dimensional eqiupotential geo-electrical installation is able to register space-time development of deformational process.

Analysis of a spectral content of electrical and temperature series (Figure 3-7) confirms conclusions made in A. N Bogolubov et all. [2002] that quasiperiodic oscillations of U_{MN} are caused by temperature and tidal strains. Final inferences on nature of these oscillations could be made after further natural observations. It is necessary to carry out a cycle of continuous geoelectrical, deformational, hydrogeological, temperature and other measurements to detect qualitative and quantitative relationships between variations of physical properties.

It is necessary to pay additional attention to analysis of a low-frequency trend of temperature and electrical series (Figure 3).

Experiments with ground model shows that the presented method of the active geoelectric monitoring gives possibility of sure registration of formation of underground cavities.

The simplified measuring device was used in model experiments. Its resolution is approximately equal to 5 mV/second of arc according to data on the interval A-V, Figure 11. Resolution ability of an electrovariometer EV, which was used in natural experiments, is approximately two orders higher. It is known that amplitudes of surface tilts caused by lunisolar tides amount to 0.04 second of arc. So, surface tilts caused by lunisolar tides could be registered by the equipotential installation.

The correct choice of the relative location of the generator, inhomogeneity, and sensors (GNS) is very important for the investigation of inhomogeneity dynamics. The GNS configuration, selected with regard to features of investigated process provides the best noise cancellation and the greatest sensitivity for a useful signal, i.e., it provides a maximal signal/noise ratio.

4. CONCLUSION

In this paper we have tried to introduce maximal volume of the experimental results. The main objective of experiments was to checkout of possibilities of instrumentation and a measurement procedure.
The main requirement is to provide maximal sensitivity to dynamic components of geophysical parameters of medium. The results of field observations and model experiments demonstrate the expediency of introduced geoelectrical monitoring methods of investigation of geodynamic processes. It is necessary to investigate possibilities of different modifications of the equipotential installations for control of concrete objects.
Unfortunately volume of the paper does not allow to discuss theoretical and instrumental problems.
Evaluation of our conclusions and hypotheses will be possible after conducting of the long-term complex geophysical observations in quiet and active regions.
In our opinion, integrating of methods of active geoelectric and seismic monitoring with the methods based on mechano-elektrical transformings in rocks is very perspective.
Space-time variations of mechano-elektrical transfer function of rocks could be studies in this case. It will produce additional information.
The methods of monitoring based on analysis self-oscillation in rocks are also important for explanation. The first results have been published in Kamshilin et al., 2002.

ACKNOWLEDGEMENTS

We'd particularly would like to thank Dr. Gantov B.A., Dr. Tolmachev V.V., Dr. Leonenko M.V. and other scientists from **"ANTIKARST AND COASTAL PROTECTION"** for supporting this research. We'd also would like to acknowledge helpful comments of Dr. Khomenko V.P. and Cherniy V.I.

REFERENCES

Anatoly N. Kamshilin, Elena N. Volkova, Oleg R. Kuzichkin and Mikhail A. Sokolnikov, Self-oscillations in rocks, results of laboratory experiments, *Annals of Geophysics*, V. 47, N. 1, 93-100, February 2004.

A. N Bogolubov, A.N. Kamshilin, E.N. Volkova, Possibilities of geoelectrical and seismo-electrical monitoring in investigations of the karst phenomena, *Enviromental Geology*, V. 41, # 7, 760-764, March 2002.

Victor P. Khomenko. Regularities and forecasting of suffusion, Geos, Moscow, 2003, (In Russian).

Volkova E.N, Kamshilin A.N., Alexandrov S.I., Efendiev M.I., Geoelectrical monitoring, *Atlas of variations of natural, anthropogenic and social processes, vol 2. Cyclical dynamic in nature and the society (in Russian*,. *Scientific Word*, 168-175, Moscow, 1998.

Electrochemical stripping analysis, a powerful technique for real-time controlling of environment pollution from heavy metals

Gemma Aragay[1,2], Anna Puig-Font[1], Serena Laschi[3], Marco Mascini[3], Luca Sanfilippo[4], Arben Merkoçi[1,2]*

[1]Nanobioelectronics and Biosensors Group, Institut Català de Nanotecnologia

[2]Departament de Química, Universitat Autònoma de Barcelona

Bellaterra, Catalonia, Spain

[3]Dipartimento di Chimica, Polo Scientifico di Sesto Fiorentino

Università di Firenze, Italy

[4] SYSTEA SpA, Anagni (FR), Italy

* E-mail: arben.merkoci.icn@uab.cat

Abstract: The real-time controlling of environmental pollution from heavy metals is of a great importance. Although several instruments and technologies are already available for laboratory use up to date there is not still any fast and low cost device that can solve such a problem fulfilling all the necessary criteria for the in-situ and long-term monitoring of heavy metals in waters. The application of electrochemical stripping analysis for heavy metal analysis using a flow through system will be shown. The detection system is based on screen-printed electrodes (SPE) which allow an easy electrochemical stripping ensuring a high sensitivity in detection of heavy metal traces and allowing an *in-situ* control of water pollution. The results obtained for the measurements of lead, cadmium and copper including the response stability will be shown. Concentrations up to few ppb have been possible to be detected with SPEs immersed up to a 2 week monitoring period. The mentioned electrochemical detection system is under further development for its future integration into a complex microanalyser that aims to be with interest in the future for water risk management in Europe in the framework of WARMER, a FP6 European project.

Keywords: Heavy metals; Lead; Cadmium; Copper; Electrochemical Stripping Analysis; Screen-printed electrodes; Flow through analysis; WARMER.

1. WATER POLLUTION AND MONITORING SYSTEMS

Since many years modern civilisation has become concerned about the water quality from all rivers, lakes, underground waters, as well as in waste waters. Nowadays, society brings very severe danger of degradation for our environment not only through catastrophes (hard risks), but also through such called “soft risks”, which involves the slowly degradation of the environment leading to disaster.[1, 2] For this reason on-line chemical analyses for water quality measurements, are considered as very important. Water monitoring tools must be adapted or improved to face new objectives, such as investigation, surveillance and operation. Given that the needs for monitoring are increasing within resource limits, the emerging tools for water-quality monitoring are being based on new principles and techniques, giving either quantitative or qualitative information. For instance, ion-selective electrodes (ISEs) have been employed as a tool for process control monitoring, measuring concentrations of many ions in drinking and other types of water. [3] Although ISEs are applied in a number of fields in flow through systems, such as clinical or environmental, they have some limitations concerning the detection limits and their low selectivity.[4] Sensors based on electrochemical stripping analysis are an alternative to ISEs. The

application of this emerging technology in connection to the use of screen-printing electrodes (SPE) for water-quality monitoring is the main objective of this paper.
Monitoring systems should be built by developing small and easy-to-maintain smart sensors, being able to detect a large spectrum of pollutants, typical for monitored region. The Water Risk Management in EuRope (WARMER) is an FP6 EU project that aims to create an extended system for on-line water monitoring with main purpose of risk management, integrating mixed technologies in the areas of semiconductors, analytical chemistry, micro-mechanical fluidic systems, ICT technology, remote sensing and extensive networking of environmental water monitoring data. The system will be ready to be connected in a large network with terrestrial "in situ" sensors and satellite remote sensing of data, capable of creating risk maps for the region. This will lead to have permanent control of water resources in Europe.
Several detection techniques ranging from optical to electrochemical will be integrated into the water monitoring system. This paper will show the latest advances in the design and application of electrochemical stripping based sensors for heavy metal analysis.

2. ELECTROCHEMICAL STRIPPING ANALYSIS FOR HEAVY METALS.

Heavy metal pollution is a growing environmental problem which requires immediate attention in terms of fast, low cost and effective devices for *in-situ* control. Several techniques for heavy metals detection are being used. Nowadays, the most common methods used for the analysis of heavy metals are flame atomic absorption spectrometry, graphite furnace-atomic absorption spectrometry, inductively coupled plasma-atomic emission spectroscopy and inductively coupled plasma-mass spectrometry.[4,5] These techniques have high sensitivity, accuracy and reproducibility. They are commonly used for measuring trace metals in laboratories but they are not suitable for *in-situ* measurements due to their complex set up. Because health concerns and legal restrictions of contamination by heavy metals, it is critical to have probes that can provide rapid on-site evaluation of heavy metal contents. Towards this goal, electrochemical methods offer several advantages related to cost and simplicity. Electrochemical stripping analysis (ESA) has proved to be a powerful technique for the determination of trace heavy metals in environmental, clinical and industrial origin samples.[6-9] ESA enhances selectivity and sensitivity by combining heavy metal separation, pre-concentration and determination in one step.

ESA has a first step that consists of the electrodeposition of the heavy metal ions at constant potential for the purpose of their preconcentration onto the surface of a working electrode, normally in a stirring solution. The second step is a linear potential sweep in the anodic direction to reoxidize or strip out the plated metals after a known and measured time. The concentration of the metals onto the electrode surface is higher than in the solution achieving by this way lower detection limits.

Square Wave Anodic Stripping Voltammetry (SWASV) is the most used ESA method for heavy metals. SWASV offers several advantages including a more rapid determination step and the fact that the necessity to remove dissolved oxygen from solution is avoided.

The focus to further improvements of ESA is mainly directed to the choice of the working electrode material. Traditionally, mercury has been used as working electrode material for heavy metals analysis in ESA due to its ability to form an amalgam with the metals. It is well known that mercury is highly toxic and the increased risks associated with its use and manipulation have caused a trend for more environmentally friendly materials. [10]. The use of bismuth instead of mercury is also reported due to his properties to form alloys with different metals but at the same time its lower toxicity in comparison to mercury. Gold-based SPE have also been used for trace measurement of heavy metals in connection to potentiometric stripping analysis.[11]

3. SCREEN-PRINTED SENSORS AND THE MEASURING SYSTEM

Screen-printing technology is a well known alternative for making patterns on various substrates (i.e., paper, plastic, wood, glass, ceramics, etc.). It offers the possibility of mass and inexpensive production of the screen-printed electrodes (SPE) to be used for several applications including the ESA of heavy metals.[6-8] SPEs (Ecobioservices & Researches s.r.l., Italy) have been used in this study (see Figure 1).

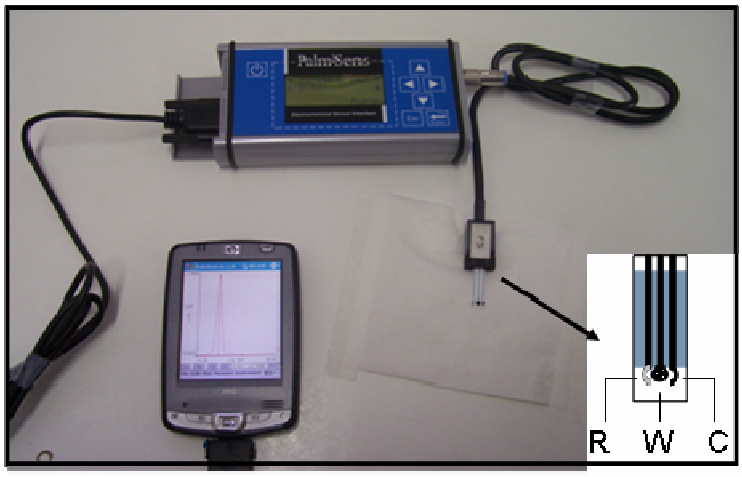


Figure 1. Image of the hand held device system used for ESA of heavy metals. The principal components are the Palmsens potentiostat; a pocket PC and the SPE the details of which are shown in the right side: a reference electrode (R), a working electrode (W) and a counter electrode (C).

They consist in three electrodes (working, W; reference, R; counter C, electrodes) in a single strip. W and C are based on graphite ink while R is a do-reference electrode based on a silver ink (see Figure 1). The graphite working electrode is modified with a cellulose-based membrane using a solution containing 100 mg of mercury acetate, 100µL of acetic acid in 10 mL of H_2O. To 1.5 mL of this solution, were added 3.5 mL of H_2O and 12.5 mg ml^{-1} of Methocel® 90HG. The 5µl of this solution were cast onto the surface of the working electrode, dried in the air at room temperature.[12] It is well known that the electrodes configuration and the graphite based inks used for printing significantly affect the electrochemical performances of the printed strip sensor being the composition of the working electrode the most important.

The measuring system used is a PalmSens potentiostat, battery powered that offers future possibilities to be integrated in the automatic analyser.
The SPEs used in metal detection have been checked in both batch and flow through measuring systems.

4. HEAVY METAL MEASUREMENTS

The measuring system (see Figure 2a) consists of a peristaltic pump used for pumping the measuring solutions to the flow cell where the SPE is integrated. The SPE is connected through the PalmSens

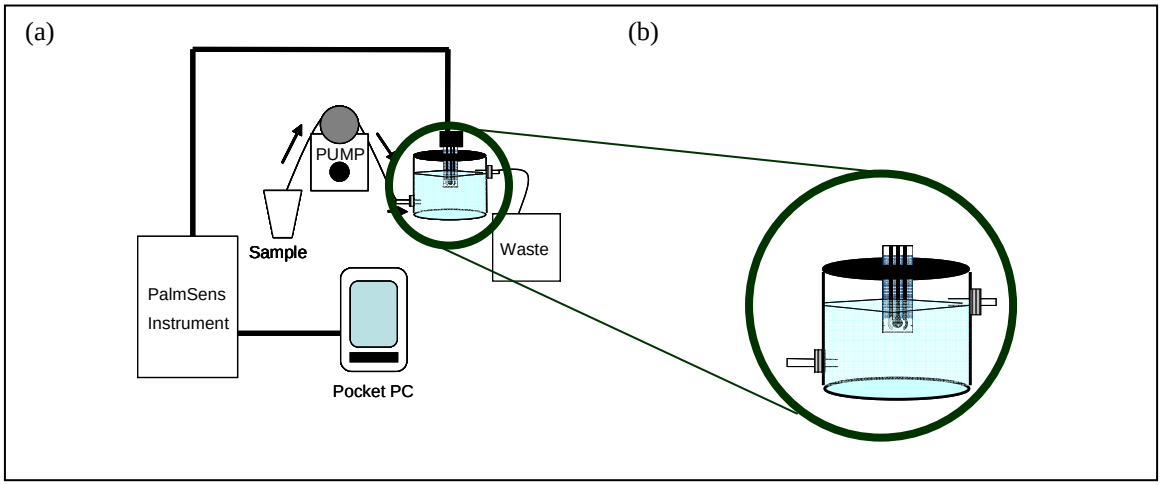


Figure 2. (a) Schematic diagram of the experimental set-up for the flow-through system. (b) Flow cell design.

potentiostat which is at the same time connected to the Pocket PC that control the applied potential and measure the currents giving the final Square Wave Voltammetry (SWV) curves used for heavy metal quantification.

The measuring SWV parameters are: potential range for stripping was between -1.1 and -0.15 V. The square wave waveform has a step potential of 3 mV, a pulse amplitude of 28 mV and a frequency of 15 Hz. A conditioning step to clean the electrode from residual heavy metals, applying a potential of -0.15 V

for 60 s, is used. The next step is the deposition step, where a potential of -1.1 V is applied to the working electrode. First two steps (conditioning and preconcentration) were performed in flow through conditions with a flow rate of 300 µL/min combined with stirring conditions. An equilibration step of 30 sec. is applied after the preconcentration step in quiescent solution (stop flow and not stirring). A pre-treatment step for the sensors is needed when the electrode (SPE) is used for the first time. It consists of applying a potential of -1.1 V during 300s and it is a necessary step to activate the mercury film of the SPE.

The flow cell used consists of a cylindrical flow cell of 10 mL of volume (Figure 2b). The advantage of this flow cell is that it avoids the risk of bubbles forming as well as the dryness of the sensor while operating during a long measuring period (up to 2 weeks). However this cell has the drawbacks of the use of stirring conditions as well as of the high amounts of sample required for measurements.

An external reference electrode (immersed in the same flow cell) has been used due to the fact that the medium did not contain chloride ions necessary for the internal reference electrode (pseudo reference electrode) to work normally.
For the batch system the measurements were performed by immersing the SPE in an acetate buffer 0.1M (pH 4.75) containing 100 ppb of lead. An external reference electrode was immersed in the same solution and was used instead of the internal one. Conditioning and deposition step have been performed under stirring conditions but equilibration step and square wave scan were performed in quiescent solution. No oxygen removing was performed. A stand-by potential of -0.15V was applied when the sensor was not used. The SPEs was immersed in the measuring solution during the whole measuring period.
For the flow through life-time measurements, the external RE was not used. Measurements were performed by immersing the SPE in a solution of 100 ppb of lead in HCl 0.1M (so the external RE was not needed). Acetate buffer 0.1M (pH 4.75) was used to perform the activation of the electrode and for stand-by conditions to avoid the damage of the electrode.

Figure 3a shows typical SWV curves obtained for standard solutions with three different heavy metals (Cd, Pb and Cu) with increasing concentrations from 0 to 100 ppb in the flow through system described previously. Three parallel measurements for each solution have been performed. The solutions were prepared using acetate buffer 0.1 M (pH 4.75). The use of the acetate buffer as measuring medium was necessary so as to avoid the damage of the SPE mercury film in the case of HCl medium.

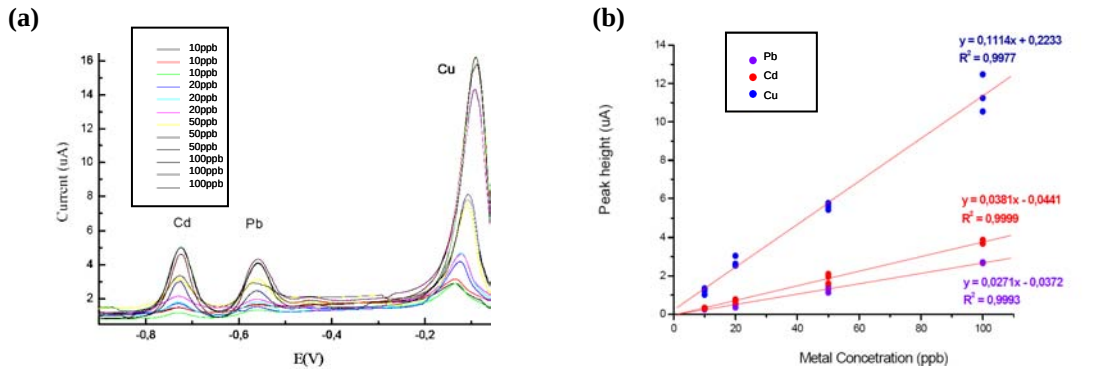


Figure 3. (a) Simultaneous detection for three heavy metals (b) Corresponding calibration curves for lead, cadmium and copper in a flow through system (10-100 ppb).

A linear trend of the SPE response (Figure 3b), with a correlation coefficient of around 0,998, toward all the heavy metals in the range from 10 to 100 ppb has been observed being the lowest concentration level measured of about 5 ppb (not shown). RSD for lead, cadmium and copper for a 50 ppb solution are 4.4 %, 6.8 % and 0.88 %, respectively.

The life-time of the SPE in batch and in flow through has been also studied. Figure 4 shows the stability of the responses during a 2 weeks measuring period under the mentioned conditions.

Figure 4a shows the trend for 100ppb concentration of lead in the batch measurements with the described parameters. There is a good stability (rsd about 10%) in the measurements performed within the same day and also during the two weeks. In the flow through system (figure 4b) the measurements are less reproducible but the peak height is higher showing that the efficiency of the stripping analysis is higher

due probably to the combined flow and stirring conditions. In both cases it can be affirmed that the SPE has a lifetime of at least two weeks being the electrode immersed in the measuring solution (acetate buffer 0.1M pH 4.75). Some parameters including the pH of the measuring solution are still being studied so as to improve the reproducibility of the measurement.

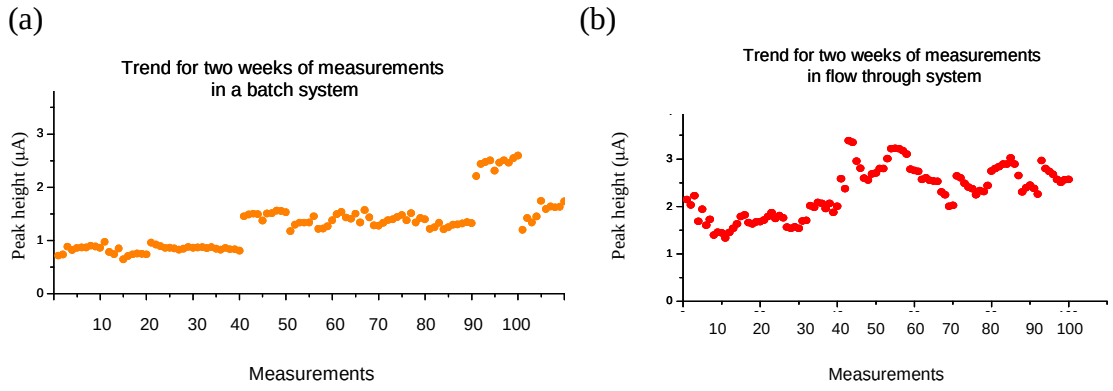


Figure 4. (a) Trend of the peak height for lead during two weeks of measurements in a batch system. b) Trend of the peak height for lead during two weeks of measurements in a flow through system.

Some tests with real sea water samples have been already performed in both systems showing good responses for different concentrations of heavy metals. Test were performed spiking real sea water samples from Mediterranian sea (Mataro area), North sea, Cantabria sea and *Mar Menor lagoon* in a heavy metal range from 5 to 100 ppb. Fresh water, contaminated with heavy metals, from different sources in Catalonia were also analysed for lead contamination. The results were compared to the standard method of heavy metals determination in sea water (Hang Drop Mercury Electrode HDME) giving rise to good results with no significant differences (results not shown).

5. CONCLUSIONS.

The designed flow through system and the flow cells presented in this paper show good response for heavy metals determination using SWASV as the selected ESA technique. A linear response to Pb, Cd and Cu at 10 to 100 ppb has been observed being the measuring system stable during a 2 weeks measuring period. The developed detection system will open the way for the fast, low cost and in-situ simultaneous analysis of trace heavy metals.
Modifications of the SPEs so as to improve the analytical performance and avoid the use of mercury are still in process at our laboratories with the aim to integrate these in an automated microanalyser for field use. Home made electrodes (mercury-free) versus mercury based electrodes are being evaluated in terms of analytical performance. The home-made SPEs have been modified with bismuth, an environmental friendly material that substitutes mercury due to its possibility to form alloys with different metals.
In addition to the SPE improvement, other flow cells designs are being considered. The use of a thin-layer wall-jet, including a lab-on-a-chip platform and new modifications of the SPEs are being considered and will be object of future publications.

AKNOWLEDGMENTS

This work has been done under the WARMER research project, FP6-034472.

REFERENCES

[1] Gagnon B., Marcoux G., Leduc R., Pouet M.-F., Thomas O.: Trends Anal. Chem. 26, 308 (2007)
[2] Kutseva N.K., Kartashova A.V., Tchamaev A.V.: J. Anal. Chem. 60, 788 (2005)

[3] Johnson R.D., Bachas L.G.: Anal. Bioanal. Chem. 376, 328 (2003)
[4] Chumbimuni-Torres K.Y., Kubota L.T.: J. Food Comp. Anal. 19, 225 (2006)
[5] Friedge H.: Umwelt 4, 291 (1983)
[6] Hansen P., "American Laboratory", Chapter 23 at Volume 6, Shelton, CT, 1991.
[7] Crivelli, F. : Prodotto Chimico & Aerosol selezione, 10, 26 (1983)
[8] Honeychurch, K.C., Hart, J.P.: Trac-Trends Anal. Chem. 22, 456 (2003)
[9] Merkoçi, A., Castañeda, M.T., Alegret, S., "Stripping analysis of heavy metals by using mercury free composite based sensors". Chapter at 'Pure & Applied Analytical Chemistry', Edited by Research SignPost, E. Palomar, p.1-22, (2005)
[10] Merkoçi, A., Alegret, S., "Graphite-epoxy electrodes for stripping analysis", Chapter 7 at Volume 49 "Electrochemical Sensor Analysis", Alegret and Merkoçi (Eds), Comprehensive Analytical Chemistry, Elsevier B.V., p.143-161, (2007)
[11] Domínguez Renedo O., Alonso-Lomillo M.A., Arcos Martínez M.J.: Talanta 73, 202 (2007)
[12] Laschi S., Palchetti I., Mascini M.: Sens. Actuators B 114, 460 (2006)
[13] Palchetti I., Laschi S., Mascini M.: Anal. Chim. Acta 530, 61 (2005)

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Mathematical modelling of aerosol transport and deposition: Analytic formulae for fast computation

Robert McKibbin

Institute of Information and Mathematical Sciences, Massey University at Albany,
Private Bag 102 904, North Shore Mail Centre, Auckland, New Zealand (r.mckibbin@massey.ac.nz)

***Keywords**:* Modelling; Advection-dispersion; Particle transport; Deposition; Hazards.

Solid and liquid particles (aerosols) ejected into the atmosphere by volcanic or hydrothermal eruptions, dust and sand swept up by storms or from other pollution sources, and droplets from crop-spraying, are subsequently dispersed by atmospheric wind currents. The particles fall under gravity while being advected by the mean wind and dispersed by turbulence. Particle sizes are generally not uniform, and may also change during flight (perhaps by particle coalescence and/or fragmentation, or, in the case of fluids, by evaporation or condensation) with consequent change to the settling speed. The wind may also change with elevation (and with time) and particles may be trapped on crop or forest foliage as they near the ground.

A quantitative model that reflects these influences on particle dispersal is outlined. It is assumed that the wind does not change over the time of particle flight, and that there is no change of particle size due to evaporation or condensation. However, the other elevation-dependent features listed above are included. Changes of conditions with elevation are treated by using a piecewise-constant wind velocity, associated dominant turbulence length scale, settling speed and trapping rates. In any case, this is the way that data are provided for most of the numerical schemes currently available. When the vertical dispersion is assumed negligible (as is commonly supposed in such modelling), analytic solutions to the advection-dispersion equations that describe the motion of the particles may be found. Results calculated directly from the analytical formulae provide examples of the method.

1. INTRODUCTION

When particles are released in the atmosphere, perhaps from volcanic eruption columns, sand or dust storms, dust pollution sources, crop-spraying, aerial top-dressing or pollen from trees, they are generally blown by the wind while falling through the air as well as being dispersed by the wind's turbulence. Particles that fall through air that has been polluted by toxic gaseous discharges from industrial processes, may adsorb chemicals which are then present in the deposits of the particles on the ground. While the paths of individual particles may be of interest, it is usually the density of the eventual distribution of the fallout on the ground that is more important. In particular, being able to predict and estimate of the thickness of volcanic ashfall deposits or chemically-polluted dust is of most value to health and safety agencies which use such information for hazard maps.

Estimation of the characteristics of the wind during any event is essential in the modelling exercise. Meteorological measurement equipment is unlikely to give more than information about average wind-speed and direction at a sequence of elevations, so any model should be able to use that information, but no more than that. The model proposed in this paper is based on dividing the atmospheric wind-flow into a sequence of layers, within each of which the speed and direction are known layer-averaged values. It

generalises previous models where the particle settling speeds remain constant, and/or the dispersion is assumed isotropic (for example, see Bonadonna *et al*. [2002]; Lim, [2006]), and includes trapping by foliage.

Early models assumed uniform wind conditions and particle sizes (for a summary, see Lim [2006]). More recently, papers by Bonadonna *et al.* [2005] and Costa *et al.* [2006], amongst others, discuss settling speed calculations that take into account air density variation with elevation as well as the effect of non-sphericity of the particles. Particle sizes from discharges or releases are generally not uniform initially, and may change during flight, either by particle coalescence and/or fragmentation, or, in the case of liquid droplets, by evaporation or condensation, with consequent changes to the settling speed. The wind (speed, direction and dominant turbulence length scales) may also change with elevation (and with time, although not considered here). The inclusion of appropriate sink terms in the conservation equations allows for modelling the trapping of some particles on forest, orchard, crop or windbreak foliage (see, for example, McKibbin [2006a]; Harper *et al*. [2007]). This aspect is important in agriculture and horticulture, where spray drift onto nearby areas may be a health hazard. Predictions of deposits of tephra ejecta from volcanic (and hydrothermal eruptions: see McKibbin *et al*. [2005], McKibbin [2006b]) are useful for scenario-planning by civil defence organizations.

Results calculated directly from the derived analytical formulae are used to provide examples of the method. The new formulae allow extremely fast direct computation of deposit distributions. Because of this, fine discretizations of release zones, non-uniform wind profiles and ejected particle size distributions can be made with little computation cost. Space precludes much detail; only a summary of the method and some aspects of its implementation can be given here.

2. MATHEMATICAL MODEL FORMULATION

It is reasonably assumed that the air is of uniform density and that its motion is not affected by the small volume fraction of particles present. A Cartesian coordinate system is arranged so that the *x*-*y* plane is the ground and the *z*-axis is vertical upwards. The mean (locally time-averaged) mass concentration of particles per unit volume of the atmosphere is denoted $C(x,y,z,t)$, and the principle of conservation of mass of the solid phase yields

$$\frac{\partial C}{\partial t} = -\nabla \cdot \mathbf{q} + m_s - m_t \tag{1}$$

where $\mathbf{q}$ is the particle mass flux per unit cross-sectional area, m_s is a mass source term (mass rate per unit volume) and m_t is a mass removal term due to trapping, such as by foliage in a forest canopy or a crop. The specific mass flux has advection, dispersion and settling components, in the form:

$$\mathbf{q} = C\mathbf{u} - \underline{\mathbf{D}} \otimes \nabla C - CS\mathbf{k} \tag{2}$$

where $\mathbf{u} = (U,V,0)$ is the mean (horizontal) wind velocity vector, with mean wind speed $W = \sqrt{U^2 + V^2}$. The particles are small, and are assumed to have quickly reached their average terminal velocity $\mathbf{u} - S\mathbf{k} = (U,V,-S)$ with respect to the mean motion of the air, where S is the particle gravitational settling speed in the downward direction ($\mathbf{k}$ is a unit vector in the positive *z*-direction). The turbulence in the atmosphere is assumed to cause mechanical dispersion of the particles; $\underline{\mathbf{D}}$ is the dispersion tensor, which may be written in the form $\underline{\mathbf{D}} = W\underline{\mathbf{L}}$ in terms of mean wind speed and a dispersion length tensor which involves characteristic dominant length scales of the atmospheric turbulence. This pragmatic model assumes a simple form where the dominant principal dispersion lengths are assumed horizontally constant, but not necessarily isotropic; this leads to Gaussian-type profiles with spreads that are directly calculable from the parameters and elapsed time (to which can be related explicitly the distance fallen and/or mean horizontal distance travelled).

For the case of a release of a mass Q of particles at point (X_0, Y_0, H) above the ground at time $t = 0$, the mass source m_s may be written using Dirac delta functions in the form

$$m_s = Q\delta(x - X_0)\delta(y - Y_0)\delta(z - H)\delta(t) \tag{3}$$

Distributed releases may be composed by superposing suitable combinations of such sources (see further below). The trapping term, considered to be proportional to the particle concentration C, is assumed to be of the form

$$m_t = k(x, y, z)C \tag{4}$$

where the trapping rate $k(x, y, z)$, which is zero outside any trapping zone, is {mass of particles trapped per unit volume of the trapping region, per unit time}/{mass of particles present per unit volume of the trapping region}; in SI units, $[k] = s^{-1}$. In general, k will depend on the foliage characteristics and the particle type and size, and possibly the wind speed (Mercer & Roberts [2005]).

It is assumed here that, in the case of fluid droplets, there is no particle mass change due to evaporation or condensation. If so, rather than include a mass loss term due to evaporation in (1), it may be better to use a particle number concentration N rather than a particle mass concentration C. Earlier work on droplet trapping by agricultural shelterbelts is reviewed in, for example, Harper *et al*. [2006, 2007], and the inclusion of evaporation is comprehensively discussed in Harper [2008]. Solid or liquid particles may agglomerate at certain elevations during their fall. For example, condensed water droplets at certain levels in the atmosphere may act as a "binder" to hold smaller solid particles together as larger agglomerates (see Rynhart [2004], for example). This may be taken into account by changing the settling speed S at levels where such agglomeration occurs.

Substitution of expressions (2) – (4) into (1) and some rearrangement gives:

$$\frac{\partial C}{\partial t} + \nabla \cdot (C\mathbf{u} - \underline{\mathbf{D}} \otimes \nabla C - CS\mathbf{k}) = Q\delta(x - X_0)\delta(y - Y_0)\delta(z - H)\delta(t) - kC\,. \tag{5}$$

For a horizontal wind, the dispersion tensor is taken to be of the form:

$$\underline{\mathbf{D}} = \begin{bmatrix} D_{xx} & D_{xy} & 0 \\ D_{xy} & D_{yy} & 0 \\ 0 & 0 & D_{zz} \end{bmatrix},$$

which is a symmetric tensor of order 2. If the wind direction is at angle θ to the x-axis (i.e $U / W = \cos\theta, V / W = \sin\theta$), and the longitudinal (downwind) and transverse (crosswind) dispersion coefficients are D_L and D_T respectively (allowed to be different here, but assumed the same in previous studies), then the dispersion tensor has the form:

$$\underline{\mathbf{D}} = \begin{bmatrix} D_L \cos^2\theta + D_T \sin^2\theta & (D_L - D_T)\sin\theta\cos\theta & 0 \\ (D_L - D_T)\sin\theta\cos\theta & D_L \sin^2\theta + D_T \cos^2\theta & 0 \\ 0 & 0 & D_{zz} \end{bmatrix}.$$

(Note that if the dispersion is isotropic, $D_L = D_T$ and $\underline{\mathbf{D}}$ is a diagonal matrix.) If the settling speed, the mean wind speed and direction, as well as the turbulent dispersion, vary with elevation, then $S = S(z)$, $U = U(z)$, $V = V(z)$ and $\underline{\mathbf{D}} = \underline{\mathbf{D}}(z)$, and (5) can be written:

$$\begin{aligned} &\frac{\partial C}{\partial t} + U(z)\frac{\partial C}{\partial x} + V(z)\frac{\partial C}{\partial y} - S(z)\frac{\partial C}{\partial z} - \frac{dS}{dz}C \\ &\quad = D_{xx}(z)\frac{\partial^2 C}{\partial x^2} + 2D_{xy}(z)\frac{\partial^2 C}{\partial x \partial y} + D_{yy}(z)\frac{\partial^2 C}{\partial y^2} + D_{zz}(z)\frac{\partial^2 C}{\partial z^2} + \frac{dD_{zz}}{dz}\frac{\partial C}{\partial z} \\ &\qquad + Q\delta(x - X_0)\delta(y - Y_0)\delta(z - H)\delta(t) - k(x, y, z)C \end{aligned} \tag{6}$$

In general, this equation has no analytical solutions, and usually a full-scale numerical approach is necessary for finding the concentration C and consequent deposits.

However, for the case where the vertical dispersion is small compared to other components, as is often supposed by other authors (for example, see Bonadonna *et al.* [2002]) for regions not very close to the ground, and the elevation-dependent profiles of wind characteristics and particle settling speeds are approximated by piece-wise (layer-wise) constant functions, then concentrations and deposits can be found as explicit analytic solutions. Fast direct computation is then possible without problems associated with accuracy and stability of spatial and temporal discretization schemes.

3 SOLUTIONS FOR THE PARTICLE CONCENTRATION

When the vertical dispersion D_{zz} is assumed negligible, and the wind and settling parameters are piecewise constant functions of elevation, it is possible to find analytic solutions for (6) with appropriate initial and boundary conditions. First, the solution for a single uniform layer is described, and then it is used to find the analytical solution for the general case.

3.1 Uniform atmosphere

A mass Q of particles with constant settling speed S is released at time $t = 0$ from a height H into a steady, uniform, horizontal wind of speed W (a constant), with constant dispersion coefficients and a uniform trapping parameter k. A set of Cartesian coordinate axes (x, y, z) is aligned so that the mean wind is in the positive x-direction, so $\mathbf{u} = (W, 0, 0)$, and the origin (take $X_0 = Y_0 = 0$ for clarity) is directly below the point of release. Then (6), with associated initial and boundary conditions becomes:

$$\frac{\partial C}{\partial t} + W\frac{\partial C}{\partial x} - S\frac{\partial C}{\partial z} = D_L\frac{\partial^2 C}{\partial x^2} + D_T\frac{\partial^2 C}{\partial y^2} + Q\delta(x)\delta(y)\delta(z-H)\delta(t) - kC \qquad (7)$$

$$C(x, y, z, 0) = 0;\ C \to 0 \text{ as } x, y \to \pm\infty,\ z \to +\infty\,.$$

The concentration at time t after release can be found using a standard solution procedure involving Laplace and Fourier transforms, and is given by:

$$C(x, y, z, t) = \frac{Q}{4\pi\sqrt{D_L D_T t^2}}\exp\left[-\frac{(x-Wt)^2}{4D_L t} - \frac{y^2}{4D_T t}\right]\exp(-kt)\,\delta[z-(H-St)]\,. \qquad (8)$$

The concentration in the atmosphere is zero except at the level $z = H - St$; the particles fall steadily with speed S, while spreading out in the x- and y-directions as time increases. The centre of mass moves at speed W in the x-direction and at speed S in the negative z-direction. The particles arrive at the ground at time $t_1 = H/S$. The mass density $f(x, y)$ of the deposit on the ground is given by integrating the downward mass flux there:

$$f(x, y) = \int_0^\infty SC(x, y, 0, t)dt = \frac{Q}{4\pi\sqrt{D_L D_T t_1^2}}\exp\left[-\frac{(x-Wt_1)^2}{4D_L t_1} - \frac{y^2}{4D_T t_1}\right]\exp(-kt_1) \qquad (9)$$

This deposition density ($[f]$ = kg m^{-2} in SI units) has elliptical level surfaces (contours), centred on the point $(Wt_1, 0, 0)$. If $k \neq 0$, the distribution of the particles trapped in the foliage (in kg m^{-3}) may be calculated explicitly (see McKibbin [2006a]). The total amount trapped is $Q_{TT} = Q(1 - e^{-kH/S})$; this leaves an amount $Q_D = Q - Q_{TT} = Qe^{-kH/S}$ distributed on the ground. Note that, interestingly, these *total* amounts do not depend on the wind speed or the dispersion parameters. Spatially-distributed trapped amounts are easily calculated – see McKibbin [2006a].

3.2 Layered atmosphere

In a system where the atmosphere is divided into n superposed layers, the solution above describes the motion of the particles in the layer in which they are released. If this region is labelled Layer 1, occupying $Z_1 < z \le Z_a$ where $Z_a \ge H$, the particles all reach the bottom of the Layer 1 (= the top of Layer 2) at time $t = t_1 = (H - Z_1)/S_1$. The particles then form a horizontally-dispersed mass source at the top of Layer 2. Because the wind in Layer 2 may be moving in a different direction to that of Layer 1, the coordinate system is rotated for Layer 2 so that the wind is in the positive "x"-direction in that layer.

The procedure is as follows. The wind direction and hence the positive "x"-coordinate axis in the system in Layer i, $1 \le i \le n$, is rotated at an angle θ_i to the global coordinate system (x, y, z), and is labelled (X_i, Y_i, z). The coordinate systems are related by:

$$\begin{aligned} X_i &= x\cos\theta_i + y\sin\theta_i \\ Y_i &= -x\sin\theta_i + y\cos\theta_i \end{aligned} \quad \text{or} \quad \begin{aligned} x &= X_i\cos\theta_i - Y_i\sin\theta_i \\ y &= X_i\sin\theta_i + Y_i\cos\theta_i \end{aligned} \tag{10}$$

The wind speed, settling speed, dispersion and trapping coefficients may vary with elevation, and so are subscripted accordingly. At time $t = t_1 = (H - Z_1)/S_1$, the particles arrive at the top of Layer 2 with an areal density distribution given by:

$$f(X_1, Y_1) = \frac{Q}{4\pi\sqrt{D_{L1}D_{T1}t_1^{\,2}}} \exp\left[-\frac{(X_1 - W_1 t_1)^2}{4D_{L1}t_1} - \frac{Y_1^2}{4D_{T1}t_1} \right] \exp(-k_1 t_1) \tag{11}$$

The (X_1, Y_1, z) coordinate system is aligned at angle $\theta_1 - \theta_2$ to the (X_2, Y_2, z) system. A small element of mass source at $(X_2, Y_2, z) = (\xi, \eta, Z_1)$ at time t_1 is given from (11) by

$$\begin{aligned} dQ = f(\xi,\eta)\,d\xi d\eta = \frac{Q}{4\pi\sqrt{D_{L1}D_{T1}t_1^{\,2}}} \exp\Bigg[&-\frac{\{\xi\cos(\theta_1 - \theta_2) + \eta\sin(\theta_1 - \theta_2) - W_1 t_1\}^2}{4D_{L1}t_1} \\ &-\frac{\{-\xi\sin(\theta_1 - \theta_2) + \eta\cos(\theta_1 - \theta_2)\}^2}{4D_{T1}t_1} \Bigg] \exp(-k_1 t_1)\; d\xi d\eta \end{aligned} \tag{12}$$

The total effect of all source elements on the particle concentration in Layer 2 is found by integration of all the small source elements $dQ(\xi, \eta)$ at (ξ, η, Z_1) given by (12).

The particle concentration in Layer 2 is then given by:

$$\begin{aligned} &C(X_2, Y_2, z, t) \\ &= \int_{\xi=-\infty}^{\infty} \int_{\eta=-\infty}^{\infty} \frac{dQ}{4\pi\sqrt{D_{L2}D_{T2}(t - t_1)^2}} \exp\left[-\frac{\{X_2 - \xi - W_2(t - t_1)\}^2}{4D_{L2}(t - t_1)} - \frac{(Y_2 - \eta)^2}{4D_{T2}(t - t_1)} \right] \exp(-k_2 t) \\ &\qquad \times \delta[z - \{Z_1 - S_2(t - t_1)\}] \end{aligned}$$

Some algebraic manipulation and use of standard integration formulae gives a compact and explicit expression for C at a given position and time, in terms of the various parameters. Proceeding via the same process from layer to layer downwards allows calculation of the concentration within all layers, and thereby the distribution of the deposit on the ground (the bottom of Layer n, say, which is at $z = 0$).

The general results for the areal mass distribution at the bottom of any Layer $i \ge 2$ are presented now. Taking $t_0 = 0$, the cohort of particles falls through Layer i between the times t_{i-1} and t_i which, respectively, correspond to their arrival at the top and bottom of that layer. The times may be calculated from:

$$t_1 = \frac{H - Z_1}{S_1}, \quad t_i = t_{i-1} + \frac{Z_{i-1} - Z_i}{S_i} = t_1 + \sum_{j=2}^{i} \frac{Z_{j-1} - Z_j}{S_j} \quad \textit{for } i \geq 2 .$$

Note that $Z_{i-1} > Z_i$; the layers are numbered from the top downwards. The areal mass density at the bottom of Layer i (at the level $z = Z_i$) is given by:

$$f(X_i, Y_i, Z_i) = \frac{Q}{4\pi\sqrt{\Delta_i(t_i)}} \exp\left[-\frac{\sigma_i(t_i)}{4\Delta_i(t_i)} - K_i(t_i) \right], \tag{13}$$

where

$$\sigma_i(t_i) = \sum_{j=1}^{i} (t_j - t_{j-1}) \left\{ D_{Tj} \left[X_j - \sum_{k=1}^{i} W_k (t_k - t_{k-1}) c_{kj} \right]^2 + D_{Lj} \left[Y_j - \sum_{k=1}^{i} W_k (t_k - t_{k-1}) s_{kj} \right]^2 \right\} \tag{14}$$

with $$c_{ij} = \cos(\theta_i - \theta_j), \; s_{ij} = \sin(\theta_i - \theta_j) ,$$

and the $\Delta_i(t_i)$ are given recursively by the formulae:

$$\Delta_1(t_1) = D_{L1} D_{T1} t_1^2, \quad \Delta_i(t_i) = \Delta_{i-1}(t_{i-1}) + (t_i - t_{i-1}) \sum_{j=1}^{i-1} (t_j - t_{j-1}) P_{ij} + (t_i - t_{i-1})^2 D_{Li} D_{Ti} \quad \textit{for } i \geq 2 ,$$

where $$P_{ij} = s_{ij}^2 (D_{Li} D_{Lj} + D_{Ti} D_{Tj}) + c_{ij}^2 (D_{Li} D_{Tj} + D_{Lj} D_{Ti}) .$$

Also, $$K_i(t_i) = \sum_{j=1}^{i} k_j (t_j - t_{j-1}) .$$

The above results are for mass distributions at the bottom of the various layers; formulae for mass fluxes at elevations between layer boundaries are similar in form. If the horizontal dispersion is isotropic ($D_{Li} = D_{Ti}$) and there is no trapping ($k_i = 0$), the resulting formulae agree with those given by Lim [2006]. If the dispersion is isotropic and uniform over all elevations ($D_{Li} = D_{Ti} = D$) and $k_i = 0$, then the results agree with those given by Bonadonna *et al*. [2002]. The deposit on the ground (at the bottom of Layer n) is given by (13) for $i = n$ at $Z_n = 0$. Note: the (X_i, Y_i) in (14) are given in terms of ground coordinates (x, y) by (10).

4. ESTIMATION OF PARAMETERS

The various parameters describing the wind speed and direction, the dominant turbulence length scales and particle settling speeds, all need to be reliably estimated. Wind mean speeds may be found from meteorological simulations or measured data. Particle settling speeds are available from correlations constructed from laboratory experiments.

Turbulence length scales – The model above assumes that the particles are spread by mechanical dispersion, caused by the turbulence in the air motion. The turbulence, caused by the wind, varies with height and in part this can be attributed to gradients in wind speed. At a given height, turbulence within the airflow is modelled as having a certain characteristic length; since turbulence has a variety of scales, the length is a typical mean value for the flow. The effect of air turbulence is incorporated using a dispersion tensor **D** whose components are constant within each layer.

Experimental observations of turbulence have suggested that the effective dispersion tensor changes with the scale of the dispersing plume. This is not allowed for by Gaussian dispersion in its simplest form. In particular, in an early paper, Sutton [1932] attempts to model dispersion using an empirical formula wherein the effective dispersion rate is given by a fractional power of the distance travelled. Essentially, by implication, this will mean having a dominant length scale that evolves with time. Pasquill [1961] also comments that it is not unreasonable to, locally with respect to height and time, assume a near steady homogeneous structure. His model also allows for a change in length scale with time due

to the spread of the release. Such effects are not excluded from the model of this paper. The particle release falls downwards and so, as the dispersion length scale and velocity constants are varied with height, this effectively allows for changes with time.

Release/emplacement heights – If a cohort of particles is released into the air by controlled means (such as from chimney stacks, aerial top-dressing, spraying machines, etc.) then the physical source position is known. In geophysical events such as dust- or sand-storms, or volcanic eruptions, release positions have to be estimated. Estimates of volcanic plume heights, distribution of ash releases from plumes and similar considerations may be found in the volcanological literature – see, for example, Sparks *et al*. [1997], McKibbin & Smith [2006], Lim *et al*. [2008a], Bonadonna *et al.* [2005], Costa *et al.* [2006]. A brief discussion about plume-type volcanic eruptions is given here.

The collective term for all particles ejected from volcanoes is *tephra*. There are many eruption types, depending on magma composition (rock type), topography, vent history, etc. One event may include several different eruption types (multiple plumes, explosions, dome collapse, pyroclastic flows, etc.); see, for example, Sparks *et al*. [1997]. The model presented above for particle deposition may be applied to thermally-bouyant plumes of hot gases which lift rock particles of various composition from the eruption site high into the atmosphere, while entraining and heating air. In this process, the plume slows and cools, eventually reaching a maximum height where it spreads laterally. As it slows, the vertical plume speed reduces from the ejection speed v_0 at the eruption vent down to zero at the top of the plume. Observed plume heights of eruptions can be used to estimate the total mass erupted through correlations; e.g., for Plinian eruptions, Carey & Sigurdsson [1989] give:

$$\log_{10}(M) = \frac{H_{\max} + 60.5}{7.18} \tag{15}$$

where M is the total mass released (kg) and $H_{\max}$ is the maximum observed column height (km).

Treating the column as a vertical line source, at least two approaches can be made to estimate the height at which tephra particles of a certain size are released into the atmosphere. One is from McKibbin & Smith [2006] who calculated the dynamics and shape of entraining plumes using a simplified model, and observed that the calculated vertical speed declined approximately linearly with height above the vent. This allows an estimate of the elevation H_S at which the plume speed is equal to the settling speed S_d of a cohort of particles with diameter d. Since the vertical plume speed v as a function of height z is given approximately by $v/v_0 = 1 - z/H_{\max}$, the release height may be estimated by:

$$H_S = H_{\max}\left(1 - \frac{S_d}{v_0}\right) \tag{16}$$

An alternative is to use the so-called *Suzuki distribution* [Suzuki, 1983], which gives a formula for the release density distribution $\zeta(z)$ of a mass of particles over the total plume height $0 \le z \le H_{\max}$. In terms of the notation used here, the traditional form contains the Suzuki constant A:

$$\zeta(z) = \frac{A^2}{H_{\max}\left[1 + (A-1)\exp(-A)\right]}\left(1 - \frac{z}{H_{\max}}\right)\exp\left[-A\left(1 - \frac{z}{H_{\max}}\right)\right] \tag{17}$$

which has the property that

$$\int_0^{H_{\max}} \zeta(z)\,dz = 1 \tag{18}$$

Here the constant has some fixed value $A \ge 1$, and the maximum density of particle release is at $z = H_{\max}(1 - 1/A)$. A typical value is $A = 5$, to give the maximum release at $0.8H_{\max}$. However, having fixed the value of A, all particle sizes have the same release distribution up the column.

Perhaps a better approach is to combine the two options above. Note that comparing (16) and the maximum density height from the standard Suzuki distribution, the role of A may be taken by the ratio v_0/S_d. It is reasonable that the maximum release density of a certain size cohort of particles be approximately at the level where the plume speed is the same as the settling speed, i.e., where the particles have neutral buoyancy, and where they may fall out of the plume into the surrounding atmosphere. Substitution for A into (17) gives:

$$\zeta_d(z) = \frac{\left(\frac{v_0}{S_d}\right)^2}{H_{\max}\left[1+\left(\frac{v_0}{S_d}-1\right)\exp\left(-\frac{v_0}{S_d}\right)\right]}\left(1-\frac{z}{H_{\max}}\right)\exp\left[-\frac{v_0}{S_d}\left(1-\frac{z}{H_{\max}}\right)\right] \qquad (19)$$

Use of (19) allows each particle size cohort to have its own Suzuki distribution, based on the relevant settling speed. For computation, the line source is partitioned into N equal sub-intervals $[(p-1)\Delta z, p\Delta z]$, $p = 1, \ldots, N$, each of length $\Delta z = H_{\max}/N$, with corresponding Suzuki density values $\zeta_d[(p-\frac{1}{2})\Delta z] / \sum_{k=1}^{N}\zeta_d[(k-\frac{1}{2})\Delta z]$. Note the normalization to ensure that the discretized distribution has a sum of unity, to maintain the condition (18).

Particle sizes and settling speeds – From geological records, volcanic deposits consist of a variety of particle sizes. Their particle size compositions are commonly estimated by sieving a sample, and retrieving size data in the form of ϕ values, which are related to particle diameters by $d = 2^{-\phi}$ or $\phi = -\log_2 d$ where d is measured in mm; a diameter of $d = 1$ mm corresponds to $\phi = 0$. Particles with large diameter d have small ϕ values and *vice versa*. Reported results of such sieving are usually stated as a set of data giving either absolute masses or volumes, or sample fraction masses or volumes, of each cohort for integral ϕ values – for typical examples, see Bonadonna *et al.* [2002]. This distribution is therefore not continuous, but represents a set of "bundled" data for a cohort of particles that are smaller than a certain standard sieve diameter but larger than the next (smaller) size with half the diameter of the first. Diameters d of particles in the cohort with $\phi = i$ (i is an integer) are constrained by $2^{-i} < d \le 2^{-(i+1)}$. The data is therefore coarse, rather than continuous as the sizes in a typical sample are.

One way to deal with this is to use a set of splines, i.e. piecewise continuous smooth functions which connect smoothly (value and slope) at the "knots" ϕ_I and take zero values at the end-points of the data set Also, for the cohort $\phi = i$, where the interpolating distribution function is given by $g(\phi)$, it is required that

$$\int_i^{i+1} g_i(\phi)d\phi = m_i\,, \qquad (20)$$

where m_i is the data value [mass or volume (or fraction of either of the total sample)] for that cohort. Quadratic splines are suitable as they require exactly the number of conditions given above.

The smoothed set of data can then be divided as finely as one wishes to use as a set of narrowly-sized cohorts, each with an appropriate settling speed. In this way, a mass release can be modelled as a set of smaller releases, each cohort having similar diameters and hence settling speeds. The deposits from each of these cohorts can be calculated directly from the formulae above and superposed to give the total deposit density at any point on the ground, and, knowing the trapping propensity for differently-sized particles, to calculate the amount left behind on foliage.

The issues of estimating settling speed are discussed by, for example, Bonadonna *et al.* [2005] and Costa *et al.* [2006]. It depends on the mass, size and shape of an individual particle, as well as on air density; a suitable correlation that includes the variation of air density with elevation can be easily implemented in the current layered model.

5. EXAMPLES

Because of space restrictions, only two examples can be presented here. The first, in an aerial top-dressing setting, is to illustrate how point sources can be analytically superposed by calculus methods before calculation of deposits. The second illustrates the use of volcanic tephra fallout (ashfall) data to compute ground deposits using the layered-atmosphere model.

Horizontal line source – If mass Q is released uniformly along a straight line between two points of equal height, the total deposit is the sum of deposits of infinitesimal releases along the line; an analytic formula may be found by analytic integration.

From (9), the ground deposit from a mass release Q from (X_0, Y_0, H) at $t = 0$ into a uniform wind of speed U in the x-direction, with no trapping, is given by

$$f(x,y) = \frac{Q}{4\pi\sqrt{D_L D_T t_1^2}} \exp\left[-\frac{(x - X_0 - Ut_1)^2}{4D_L t_1} - \frac{(y - Y_0)^2}{4D_T t_1} \right] \tag{21}$$

where $t_1 = H/S$, and S is the settling speed. If Q is now released uniformly along a straight line between (X_a, Y_a, H) and (X_b, Y_b, H), the total deposit is the sum of deposits of infinitesimal releases along the line $(X_0, mX_0 + c, H)$ [where $m = (Y_b - Y_a)/(X_b - X_a)$ and $c = Y_a - mX_a$], and is given by:

$$\begin{aligned} f_{topdress}(x,y) &= \int_{X_a}^{X_b} \frac{Q}{L} \frac{\sqrt{1+m^2}}{4\pi\sqrt{D_L D_T t_1^2}} \exp\left[-\frac{(x - X_0 - Ut_1)^2}{4D_L t_1} - \frac{(y - mX_0 - c))^2}{4D_T t_1} \right] dX_0 \\ &= \frac{\sqrt{\pi} e^{-\beta}}{2\sqrt{A}} \left\{ erf\left[\sqrt{A}(X_b - \alpha) \right] - erf\left[\sqrt{A}(X_a - \alpha) \right] \right\} \end{aligned} \tag{22}$$

where

$$A = \frac{1}{4D_L t_1} + \frac{m^2}{4D_T t_1},\ B(x,y) = \frac{x - Ut_1}{2D_L t_1} + \frac{m(y-c)}{2D_T t_1},\ C(x,y) = \frac{(x - Ut_1)^2}{4D_L t_1} + \frac{(y-c)^2}{4D_T t_1}$$

$$\alpha(x,y) = \frac{B}{2A},\ \beta(x,y) = C - \frac{B^2}{4A}$$

A suitable modification can be made in the case where $X_b = X_a$.

In Figure 1, an example where four flight passes are made at an elevation of 200 m, while a 20 m s^{-1} wind blows from the West; dispersion lengths are taken to be 1 m (dispersion parameters 20 m^2 s^{-1}), with a particle settling speed of 10 m s^{-1}. The dashed lines are a plan view of the aeroplane's path while the release is occuring and the contour lines are at intervals of 10 % of the maximum deposit.

Other examples of combining point sources to form a vertical line source and a horizontal disk are provided in Lim *et al*. [2008b].

Volcanic ashfall – Hurst [1994] gave a set of wind and mass release data (titled EX401) which he used for a demonstration of ASHFALL, a program that calculates estimates of volcanic fallout (tephra deposits) from a vertical ash plume. Input data included the wind speeds and directions over various height intervals, the masses and settling speeds of particle size cohorts, horizontal diffusion (= dispersion) coefficients (with assumed horizontal isotropy), and a Suzuki constant ($A = 5$). This data set serves as a good test of the parity of predictions from Hurst's numerical simulation (which used a finite-difference, time-stepping procedure) and the method described here, where the calculation is direct.

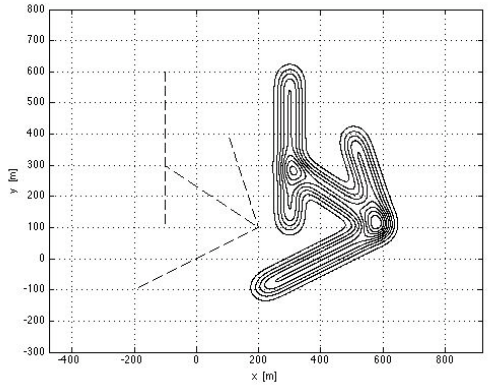


Figure 1. Contours of the deposit from a sequence of aerial top-dressing passes in a wind from the West. Dashed lines are plan views of release paths. See text for parameter values.

The method of this paper produces the contours of ashfall thickness (in mm) as shown in Figure 2. The agreement with the simulation contour plot given by Hurst [1994, Fig. 5] is very good.

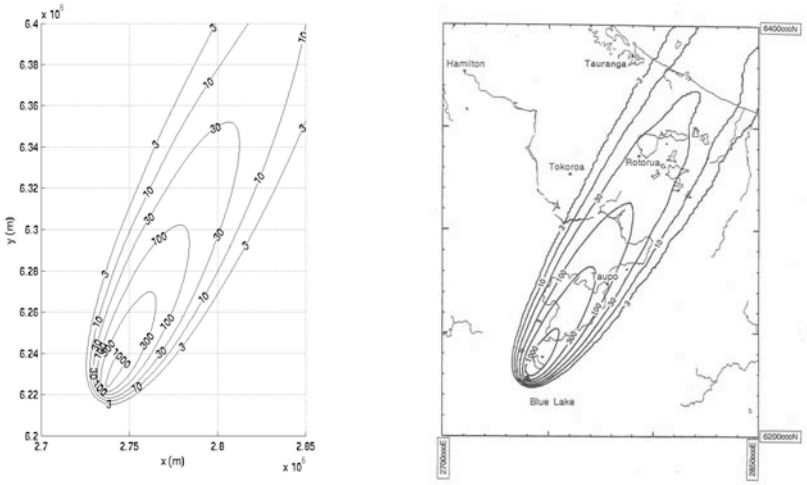


Figure 2. Left: Calculated contours of ashfall thickness using the method of this paper. Right: Hurst's [1994, Fig. 5] contour plot. Both are for Hurst's data set EX401.

6. SUMMARY

A mathematical model that allows estimation of deposit distributions from releases of particles in the atmosphere has been described. The solution for a mass release from a single point can be used as a building block for distributed sources, either analytically for some simple geometries such as uniform release along a straight line, or as non-uniform releases in any arrangement by superposition of discrete events. Calculation is by direct evaluation of analytic formulae, so no decisions have to be made about convergence or stability, such as must be considered in numerical solutions of differential equations. Good results rely on satisfactory estimation of the parameters arising in the formulae. Wind conditions, particle sizes and positions of their release, and the trapping characteristics of foliage, all need to be available. However, because the formulae are explicit, the sensitivity with respect to variation of the parameters is readily explored without the re-meshing usually required in numerical schemes. Consideration of the effects of inclusion of vertical dispersion near the ground, the evaporation or condensation of liquid droplets, and appropriate turbulence length scales in forest canopies is continuing.

REFERENCES

Bonadonna, C., Macedonio, G. and Sparks, R.S.J., Numerical modelling of tephra fallout associated with dome collapses and Vulcanian explosions: application to hazard assessment on Montserrat. *In:* Druitt, T.H. and Kokelaar, B.P. (eds), *The eruption of Soufriere Hills Volcano, Montserrat, from 1995 to 1999*, Geological Society, London, Memoirs, 21, 517-537, 2002.

Bonadonna, C., Connor, C.B., Houghton, B.F., Connor, L., Byrne, M., Laing, A. and Hincks, T.K., Probabilistic modelling of tephra dispersal: Hazard assessment of a multiphase rhyolitic eruption at Tarawera, New Zealand, *J. Geophys. Res. 110*, B03203, doi:10.1029/2003JB002896, 2005.

Carey, S.N. and Sigurdsson, H., The intensity of Plinian eruptions, *Bulletin of Volcanology*, *51*, 28-40,1989.

Costa, A., Macedonio, G. and Folch, A., A three-dimensional Eulerian model for transport and deposition of volcanic ashes, *Earth and Planetary Sci. Letters, 241,* 634-647, 2006.

Harper, S.A., Mathematical models for dispersal of aerosol droplets in an agricultural setting, PhD thesis, Massey University, New Zealand (in press) 2008.

Harper, S.A., McKibbin, R. and Wake, G.C., Modelling droplet interception by a shelterbelt: a continuum approach, *Proceedings of the Third Annual Postgraduate Conference of the Institute of Information and Mathematical Sciences, 25 October 2006*, Massey University, Auckland, 38-44, 2006.

Harper, S., McKibbin, R. and Wake, G. A simple mathematical model for spray droplet transport and interception by a shelterbelt, *Proceedings of the Fourth Annual Postgraduate Conference of the Institute of Information and Mathematical Sciences, 24 October 2007,* Massey University, Auckland, 16–23, 2007.

Hurst, A.W., ASHFALL – A computer program for estimating volcanic ash fallout, *IGNS Science Report 94/23, Institute of Geological & Nuclear Sciences*, New Zealand, 1994.

Lim, L.L., Modelling of volcanic ashfall, PhD thesis, Massey University, NZ, 2006.

Lim, L.L., Sweatman, W.S. and McKibbin, R., A simple deterministic model for volcanic ashfall deposition, *The ANZIAM Journal 49*(3), (in press) 2008a.

Lim, L.L., Sweatman, W.L., McKibbin, R. and Connor, C.B., Tephra fallout models: The effect of different source shapes on isomass maps, *Mathematical Geosciences*, *40*, 147-157, 2008b.

McKibbin, R., Modelling pollen distribution by wind through a forest canopy, *JSME International Journal, Series B, 49*(3), 583–589, 2006a.

McKibbin, R., Modelling deposition of hydrothermal eruption ejecta, *Conference Proceedings: (28th) New Zealand Geothermal Workshop & NZGA Seminar, 15–17 November 2006,* University of Auckland, New Zealand, 2006b.

McKibbin, R., Lim, L.L., Smith, T.A. and Sweatman, W.L., A model for dispersal of eruption ejecta, *Proceedings of the World Geothermal Conference 2005, Antalya, Turkey, 24-29 April 2005*, International Geothermal Association, Paper no. 0715, ISBN 975-98332-0-4 (CD), 2005.

McKibbin, R. and Smith, T.A., Hydrothermal eruption jets: Air entrainment and cooling, *Conference Proceedings: (28th) New Zealand Geothermal Workshop & NZGA Seminar, 15-17 November 2006*, (CD), University of Auckland, 2006.

Mercer, G. and Roberts, T., Predicting off-site deposition of spray drift from horticultural spraying through porous barriers an soil and plant surfaces, *Proceedings of the 2005 Mathematics-in-Industry Study Group held at Massey University, Auckland, New Zealand, 24-28 January 2005*, Massey University, 27-52, 2005.

Pasquill, F., The estimation of the dispersion of windborne material, *The Meteorological Magazine*, *90*, 33-49, 1961.

Rynhart, P.R., Mathematical modelling of granulation processes, PhD thesis, Massey University, New Zealand, 2004.

Sparks, R.S.J., Bursik, M.I., Carey, S.N., Gilbert, J.S., Glaze, L.S., Sigurdsson, H. and Woods, A.W., *Volcanic plumes*, Wiley, New York, 1997.

Sutton, O.G.A., Theory of eddy diffusion in the atmosphere, *Proceedings of the Royal Society of London, Series A, 135*, 143-165, 1932.

Suzuki, T., A theoretical model for dispersion of tephra. *In:* Shimozuru, D. and Yokoyama, I. (eds), *Arc volcanism, physics and tectonics*, Terra Scientific, Tokyo, 95-113, 1983.

A Mathematical Model for Numerical Simulation of Shallow Water Flow: Description and Practical Application of GUAD|2D

J. Murillo[a], M. Rodríguez Pallarés[b] A. Andrés-Urrutia[c] P. Brufau P[d]. and García-Navarro[e].

[a] *Fluid Mechanics, CPS, University of Zaragoza, Spain (Javier.Murillo@unizar.es)*

[b] *INCLAM, S.A.Water Engeneering, Spain (martin.rodriguez@inclam.com)*

[c] *INCLAM, S.A. Water Engeneering, Spain (a.andres@inclam.com)*

[d] *Fluid Mechanics, CPS, University of Zaragoza, Spain (Brufau@unizar.es)*

[e] *Fluid Mechanics, CPS, University of Zaragoza, Spain (Pigar@unizar.es)*

Abstract: GUAD 2D is a finite volume based two-dimensional model for the numerical simulation and analysis of flood waves caused by different factors such as extreme rainfall, gradual or sudden dam break.. The model is able to deal with critical, subcritical or mixed flow situations over irregular topography. The computation requires initial conditions, that can correspond to dry bed or any other flow state.

Keywords: Shallow Water, Hydraulic simulation model, two-dimensional, finite volume.

1. INTRODUCTION

Many engineering and environmental problems involve the study of unsteady water flows. River flows, in particular, are mostly unsteady and, as they are characterized by the presence of a vertical scale much smaller than the horizontal ones, they can be described by the shallow water model (Cunge et al. 1980) which forms a set of non linear hyperbolic equations.

A great deal of work has been devoted to develop 1D and 2D numerical models for unsteady shallow flows in the last decades and various computational techniques using finite difference, finite element and finite volume methods have been reported (Cunge et al. 1980, Bellos et al. 1991, Alcrudo and García-Navarro 1993, Sleigh et al. 1998, Bermúdez et al. 1998, García-Navarro and Vázquez-Cendón 2000). Several numerical difficulties must be adequately treated to obtain an accurate solution without numerical errors. Zhao et al. (1994) provided a good historic revision and the features required for a two dimensional river flow simulation model: it should be able to handle complex topography, dry bed advancing fronts, wet-dry moving boundaries, high roughness values, steady or unsteady flow and subcritical or supercritical conditions: natural topography is the main challenge. Dominant source terms and open boundaries are two important difficulties to face when using a conservative method since they both can damage the conservative character of the solution. Bed slope and friction source terms are of special relevance in hydraulic applications based on a shallow flow model. For that reason, a considerable effort has been recently devoted to this topic in a search for the correct source term, discretization, given a particular numerical scheme with good properties for the homogeneous case (Leveque, 1998, Roe, 1981, Glaister 1992, Vázquez-Cendón 1999, Bermúdez and Vázquez-Cendón 1994 and Bermúdez et al. 1998, Murillo et al., 2006).

Another numerical problem of relevance is the modelling of wet/dry interfaces between internal cells that have traditionally represented a difficulty for modellers wanting to solve the shallow water equations over a bed of irregular geometry. Flow over dry bed involves a complicated situation that can be analyzed as a boundary condition which is dynamically

changing in time with the moving front and continuously expanding or reducing the flow domain. The alternative is to include the wet/dry interfaces in the full domain of computation in which there may be wet cells and dry cells simultaneously. In this case the numerical scheme chosen for the discretization must be able to cope with them. In general, cells being flooded or dried during the computation tend to introduce numerical instabilities in the solution, resulting for example in negative water depths or unphysical high velocities. Different approaches have been proposed to handle them (Kramer, 2001, Beffa and Connel 2001, Heniche and Secretan 2000, Kawahara and Umetsu 1986, Khan 2000, Brufau et al. 2002 and Brufau et al. 2004, Begnudelli and Sanders 2006, Bradford and Sanders 2002, Murillo et al., 2007).

The simulation model GUAD2D is based on a finite volume method developed at the University of Zaragoza (for more details see Murillo et al. 2006, Murillo et al. 2007). The resulting software package has been adapted for standard personal computers. Furthermore it has been provided with a user-friendly interface that simplifies the definition and set-up of two-dimensional simulations (GUADCreator).

The package also offers post-processing module called GUADView that is fed from the GUAD2D results and is very helpful to plot, visualize and analyze them. It includes the most necessary GIS capabilities to analyze this kind of simulations making possible the visualization of raster images and vector layers together in the same frame with the results of the simulation generated by the model, simplifying the right identification of flood zones. Also the visualization package can show time histories of depth and velocities at different locations, water sections and hydrographs. The post-processing module offers the possibility to export the results to a standard ".avi" format. Moreover, it can also start the analysis before the conclusion of the simulation process with the help of a local area network (LAN).

2. MATHEMATICAL MODEL

The water movement is governed by the basic laws of mass and momentum conservation under the shallow water hypothtesis. This implies a depth average process in the equations and is associated to the assumption of hydrostatic pressure vertical distribution. The formulation takes the form of a set of non linear hyperbolic equations that, in two dimensions, involve the water depth as well as the two depth averaged components of the velocity,

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \mathbf{E} = \frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x}(\mathbf{U}) + \frac{\partial \mathbf{G}}{\partial y}(\mathbf{U}) = \mathbf{S}(x, y, \mathbf{U})$$

$$\mathbf{U} = \left(h, q_x, q_y\right)^T$$

$$\mathbf{F} = \left(q_x, \frac{q_x^2}{h} + \frac{gh^2}{2}, \frac{q_x q_y}{h} \right)^T, \quad \mathbf{G} = \left(q_y, \frac{q_x q_y}{h}, \frac{q_y^2}{h} + \frac{gh^2}{2} \right)^T \tag{1}$$

where $q_x = uh$ and $q_y = vh$ are the unit discharges in the Cartesian directions. The variable h represents the water depth, g is the acceleration of the gravity and (u, v) are the averaged components of the velocity vector $\mathbf{u}$ along the x and y coordinates respectively. The source terms in the momentum equations are the bed slopes Sox, Soy and the friction losses Sfx, Sfy along the two coordinate directions,

$$\mathbf{S} = \left(0, gh(S_{0x} - S_{fx}), gh(S_{0y} - S_{fy})\right)^T \tag{2}$$

where,

$$S_{0x} = -\frac{\partial z}{\partial x}, S_{0y} = -\frac{\partial z}{\partial y} \tag{3}$$

and the friction losses in terms of the Manning's roughness coefficient n, with

$$S_{fx} = n^2 u\sqrt{u^2+v^2}/h^{4/3},\ S_{fy} = n^2 v\sqrt{u^2+v^2}/h^{4/3} \tag{4}$$

3. NUMERICAL METHOD

The physical domain is discretized in grid cells for the numerical resolution. Our method is able to work over cells of different shapes (triangles, quadrilaterals, hexagons) and the choice of the grid is not a trivial step. Structured Cartesian grids are simpler and more efficient whereas unstructured triangular grids offer more geometric adaptability and better refinement possibilities.

A cell-centred finite volume method has been formulated in GUAD2D where all the dependent variables of the system are represented as piece-wise constants (first order). A discrete approximation of Equation (1) is applied in every cell at a given time so that the volume integrals represent integrals over the area of the cell and the surface integrals represent the total flux through the cell boundaries. Denoted by Ui, the average value of the conservative variables over the volume $\mathbf{\Omega}_i$, at a given time, from equation (1), the conservation equation can be written for every cell as follows:

$$\frac{\partial \mathbf{U}_i}{\partial t} A_i + \oint_{\partial\Omega} (\mathbf{E}\cdot\mathbf{n})\,\mathrm{d}s = \int_{\Omega} \mathbf{S}\,\mathrm{d}\Omega \tag{5}$$

where A_i is the area of the cell, n is the unit normal vector pointing outward and $\partial\Omega$ is the contour line. A mesh fixed in time is assumed and the contour integral is approached by a sum over the cell edges. In all of them, the normal flux is approximated via an upwind flux difference splitting technique

$$\oint_{\partial\Omega_i} (\mathbf{E}\cdot\mathbf{n})\,\mathrm{d}s \approx \sum_{k=1}^{NE} (\delta\mathbf{E}_k\cdot\mathbf{n}_k)s_k \tag{6}$$

where, k represents the index of the cell edge, NE is the total number of edges in the cell. The vector nk is the normal outward unit to edge k, sk is the length of the side, and $\delta\mathbf{E}_k\cdot\mathbf{n}_k$ is the numerical flux difference. Upwind schemes are based on the idea of discretizing the spatial derivates so that the information is taken from the side it comes from. When the source terms are present, it has previously been shown that the flux derivates and the source terms have to be discretized in a similar manner (Hubbard et al, 2000, and Brufau, et al, 2002). The evaluation of fluxes and sources at the same local state is important. The mathematical properties of the hyperbolic system of equations (1) include the existence of a Jacobian matrix, $\mathbf{J}_\mathbf{n}$, of the normal flux to a given direction (E n) defined as

$$\mathbf{J}_\mathbf{n} = \frac{\partial(\mathbf{E}\cdot\mathbf{n})}{\partial\mathbf{U}} = \frac{\partial(\mathbf{F})}{\partial\mathbf{U}} n_x + \frac{\partial(\mathbf{G})}{\partial\mathbf{U}} n_y \tag{7}$$

making possible for the generation of an approximate matrix $\mathbf{J}_\mathbf{n}^*$. Using the eigenvalues $\tilde{\lambda}^m$ an eigenvectors $\tilde{\mathbf{e}}^m$ of this matrix it is possible to linearize fluxes and source terms (Brufau et al. 2002 and Hubbard and Garcia-Navarro 2000). The first order formulation of the upwind scheme is as follows (Murillo et al. 2006):

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{k=1}^{NE} \sum_m ((\tilde{\lambda}^{m-} \alpha^m - \beta^m) \tilde{\mathbf{e}}^m)_k^n) \frac{l_k}{A_i} \Delta t \qquad (8)$$

where α^m and β^m are coefficients of the method whose exact expression and meaning has been previously detailed (Murillo et al., 2006). The discretization of the bottom elevation source terms is successfully constructed when it ensures an exact balance between flux gradients and bed variations (Leveque 1998, Hubbard and García-Navarro, 2000). It has been demonstrated that, in first order finite volume schemes, if the upwind technique is applied to the flux and bottom terms, in the case of still water, the equilibrium is maintained for the water level surface (Bermúdez and Vázquez, 1998, Brufau et al, 2002).
The friction term dictates over any other influences in many practical situations, in particular, in wet/dry fronts, characterized by small values of water depth. A explicit upwind discretization of the friction source term that guarantees a perfect discrete balance has been used (Burguete et al., 2006).

4. GUAD 2D

4.1 General Working Issues

GUAD 2D is an application that solves the most popular problems in two-dimensional trade models. Contrary to the package presented here most models have limited simulation capabilities or are developed for flat terrain, and they can not simulate switching flow regimes over levees and embankments correctly. Also this tool is able to deal with huge terrains divided up to 2.000.000 cells. A wide variety of boundary conditions are handled by the model: input hydrographs, constant water levels, dumps, etc., allowing to the user to initiate the simulation in dry/wet condition terrain in any case. To create a simulation, the model requires at least a digital terrain model, single or global rub value for each cell terrain, one or more input/output boundary conditions. The final requirement simulation parameters are the desired time of simulation and result paths to allocate the number of images that the user should previously have defined.

4.2 A Simulation Generation

Starting up GUAD 2D based on a simulation over a modelled terrain, users should follow the steps explained in Figures 4.2.1, 4.2.2, 4.2.3 and 4.2.4. 4.3.

4.3 Simulation Outcomes

This two dimensional model will generate the results in a single extension and a proper format file (*.gds). Those GUADView readable files will contain all the generic information from the GUAD 2D simulation i.e.: modelled simulation terrain, results represented such as water layers showing depth, level, and u & v component of velocity. The user will be able to get this information at any moment in time during the simulation. Therefore, this data will be identified by component values and its effective period of time.

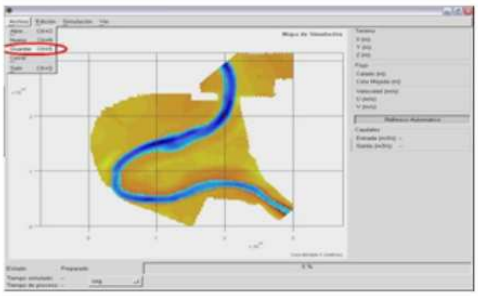


Fig. 4.2.1. Simulation Terrain in GUAD 2D

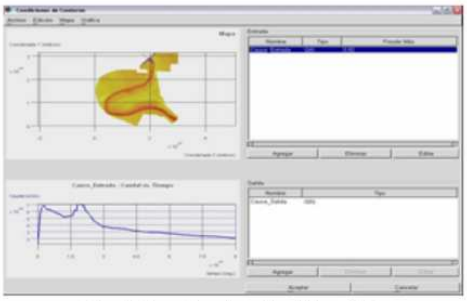


Fig. 4.2.2. Boundary Conditions List

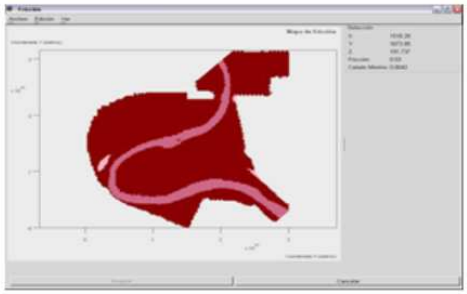


Fig. 4.2.3. Rub Grid for each cell

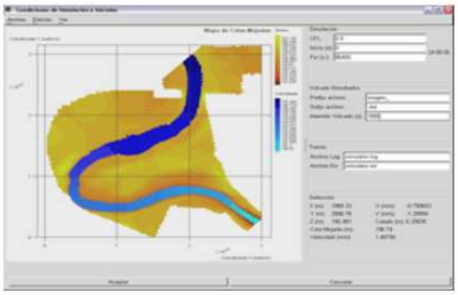


Fig. 4.2.4. Configuration of simulation parameter

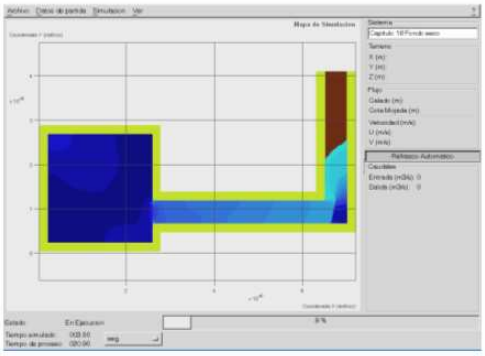


Fig. 4.2.5 Simulation Start-up

5. GUADVIEW

5.1 General Working Issues

GUADView is an analysis application to analyse GUAD 2D results. It can be considered as a GIS itself, with specific analysis capabilities added for this kind of outcomes. It is able to show in different layers data contained on each grid cell. Data is composed by depth, level and module/direction component of velocity. Thanks to the post-processing module capabilities, the analysis of hydraulic information layer is fast and simple (Fig 5.1., Fig 5.2, Fig 5.3 and Fig 5.4). GUADView can also create simulation videos, envelope layers, offers layer control tools, vector and raster layer managing and exports GUAD 2D layer to commercial GIS format.

5.2 Outcomes Analysis

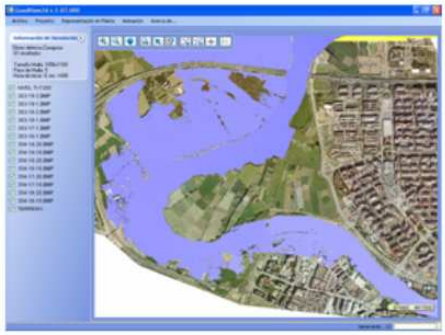


Fig. 5.1. Water Layer Visualization over raster

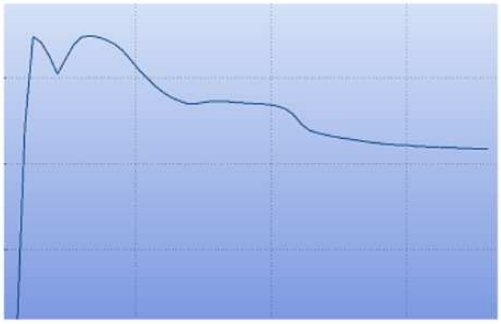


Fig. 5.2. Terrain section hidrographs.



Fig. 5.3 Visualize section details



Fig. 5.4. Video with GUAD 2D water layer

6. PRACTICAL CASE

The case described corresponds to an urban flooding event in Miranda de Ebro (Spain). The topography was obtained with LIDAR technology. With this information a rectangular mesh, with a cell size of 2 meters, was used for the simulation. The flood is generated by the opening of a reservoir located in the north of the city. The flooding wave advances correctly through the streets, from north to south, until it reaches the riverbed. Figure 6 correspond to two calculation instants in which we can depict the evolution of the flood represented on orthophoto using different tones of blue (darker as the depth of the water is greater) for the depth.



Figure 6. Flood evolution.

7. TIME SIMULATION COMPARATION

In order to compare the performance of GUAD 2D respect to other software packages, Sobek and Mike 21, different test cases were defined in Garrote et al., 2008. For the comparison, analytical and real cases were used: a straight channel with a moving

hydraulic jump, a dam break with a 90º bend channel, a channel with an expansion and a contraction and finally a flooding event. The computational times obtained with an Intel Xeon 1.7GH Processor and 1GB RAM memory are shown in Table 1. For more details see Garrote et al., 2008.

	Sobek	Mike 21	GUAD 2D
Numerical Scheme	Finite Differences	Finite Differences	Finite Volumes
Hydraulic jump	111 s	262 s	21 s
Channel contracction	182 s	160 s	44 s
Dam break	9876 s	Inconsistent results	235 s
Flooding event (2200 ha)	119324	Unfeasible computational time	54264 s

8. CONCLUSIONS

The hydraulic tool presented in this work is able to deal with complicated phenomena. The initial and boundary conditions can be easily implemented and the results can be analysed in most commercial GIS applications. Nowadays, GUAD 2D has been used in real cases and tested in several laboratory cases, thereby obtaining results that have been contrasted with present hydraulic situations, offering to hydraulic communities a reliable, fast and functional solution.

9. REFERENCES

Alcrudo F. and García-Navarro P. A high resolution Godunov-type scheme in finite volumes for the 2D shallow water equations. *Int. Journal for Num. Meth. in Fluids*, 16(6), 489-505, 1993.

Beffa C. and Connel R. Two-dimensional flood plain flow. I: Model description, *J. of Hydraulic Engineering*, 6(5):397-405, 2001.

Begnudelli L, Sanders F. Unstructured grid finite-volume algorithm for shallow-water flow and scalar transport with wetting and drying. *Journal of Hydraulic Engineering*, 132(4):371–384. 2006

Bellos C.V., Soulis J.V. and Sakkas J.G. Computation of two-dimensional dam-break induced flows. *Adv. Water Res*., 14(1):31-41, 1991.

Bradford S.F. and Sanders B.F. Finite-volume model for shallow water flooding of arbitrary topography, *Journal of Hydraulic Engineering*, 128(3), 289-298, 2002.

Brufau, P., Vázquez-Cedón, M.E., and García-Navarro, P.. A numerical model for the flooding and drying of irregular domains. *International Journal for Numerical Methods in Fluids*. 39, pp.247-275, 2002.

Brufau P., García-Navarro P. and Vázquez-Cendón M.E., Zero mass error using unsteady wetting/drying conditions in shallow flows over dry irregular topography. *Int. Journal for Num. Meth. in Fluids* ,45, 1047-1082, 2004.

Burguete, J., and García-Navarro, P.. Efficient construction of high-resolution TVD conservative schemes for equations with source terms: application to shallow water flows. *International Journal for Numerical Methods in Fluids*. 37, pp.209-248, 2001.

Burguete J., García-Navarro P. and Murillo J., Analysis of the friction term in one-dimensional shallow water model: application to open channel and river flow. *Journal of Hydraulic Engineering*, (under revision).

Chow, V.T. Open channel flow. MacGraw-Hill Book Co. Inc., 1959.

Cunge J.A., Holly F.M. and Vervey A., Practical aspects of computational river hydraulics. *Pitman* 1980.

Heniche M., Secretan Y., Boudreau P. and Leclerc M. A two-dimensional finite element drying-wetting shallow water model for rivers and estuaries, *Adv. in Water Research*, 23, 359-372, 2000.

Hubbard, M. E. and García-Navarro, P.. Flux Difference Splitting and the Balancing of Source Terms and Flux Gradients. *Journal of Computational Physics*. 165, pp.89-125, 2000.

Kawahara M. and Umetsu T. Finite element method for moving boundary problems in river flow, *Int. Journal for Num. Meth. in Fluids*, 6, 365-386, 1986.

Khan A.A. Modeling flow over an initially dry bed, *Journal of Hydraulic Research*, 38(5), 383-389, 2000.

Murillo J., García-Navarro P., Burguete J. and Brufau P. A conservative 2d model of inundation flow with solute transport over dry bed. *Int. Journal for Numerical Methods in Fluids*, 52:1059–1092 2006.

Murillo J., García-Navarro P., Burguete J. and Brufau P. The influence of source terms on stability, accuracy and conservation in two-dimensional shallow flow simulation using triangular finite volumes. *Int. Journal for Numerical Methods in Fluids*, 54:543–590 2007.

Sleigh P.A., Berzins M., Gaskell P.H. and Wright N.G. An unstructured finite volume algorithm for predicting flow in rivers and estuaries. *Computers and Fluids*, 27(4):479-508 1998.

L. Garrote, F. Laguna and J. Lorenzo, Comparación entre distintos modelos comerciales, *Jornadas Técnicas sobre Hidráulica Fluvial, CEDEX*, p. 203, 2008

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Effect of Mud Rheology on its Inertial Dam-Break Flow

B. Nsom [a] **, K. Bouchlaghem** [b] **& N. Latrache** [a]

[a] *Université Européenne de Bretagne – Université de Brest. LBMS – EA (ENSIETA/ENIB)-UBO - IUT de Brest - LIME – BP 93169. Rue de Kergoat. 29231 BREST Cedex 3, France (*blaise.nsom@univ-brest.fr , noureddine.latrache@univ-brest.fr*)*

[b] *Université de Sousse – ISSAT – Unité Energétique et Environnement. Cité Taffala. SOUSSE, Tunisie (*bouchlaghemkarim@yahoo.fr *)*

Abstract: This paper considers the flow generated by the collapse of a dam retaining argillaceous mud when the inertial effects are dominant against the viscous ones, in the momentum balance equation. This flow configuration was studied by Ritter [1892] for inviscid fluids. The more realistic case where the channel is inclined and rough was considered by Su and Barnes [1970] who extended Whitham's [1955] and Dressler's approximate solution [1952; 1954] by using a perturbation series. In this paper, the mud is modelled using a Herschel-Bulkley rheological equation and the flow develops in a horizontal smooth channel on the basis of the shallow water assumption. Applying the principles of conservation of mass and momentum, an equation of motion is formed in the non-dimensional form without neglecting any of the stress tensor components, and solved analytically in order to point out the effect of the rheological parameters on the dam-break flow characteristics in the inertial regime. This flow configuration is a scenario of initiation of certain geological flows such as muddy flash floods.

Keywords: Dam-break problem; Debris flow; Flash flood; Herschel-Bulkley equation; Inertial regime; Modelling; Argillaceous mud; Shallow water approximation.

INTRODUCTION

Hydraulics represents an important discipline in the engineering school curriculum for environmental science and the dam-break flow has become a main topic for hydro-geological hazards modelling. After long and intense rainfalls in a mountainous region, large quantities of muddy water can be retained behind a natural dam made of cross-linked branches, stones, debris, plants and different other materials obstructing a torrent. If, for some reason, this dam collapses, a dam-break flow is generated. The dam-break flow was studied extensively in the literature with water as fluid. As for any gravity current, the flow description depends on the time scale. Immediately after the dam collapse, the inertial forces are the dominant ones and this flow configuration can model a flash flood. Then, a viscous regime takes place which can model a debris flow (Hunt, 1994; Nsom et al., 2000; Nsom, 2002) where the viscous forces become the dominant ones. Both, the flash floods and the debris flows are responsible for numerous losses of life and for significant damages. Focusing on the inertial regime modelling, the following physical situation is considered (fig. 1): During negative time, water is stored to a depth h_0 between a fixed wall and a vertical plate (model dam), the channel being horizontal and dry downstream. At the initial time (t=0), the dam fails completely and suddenly, and the water is released downstream in the form of a positive surge progressing downstream, while upstream from the dam, a negative wave (depression wave) propagates. For the one-dimensional case, based on the shallow water approximation, the solution of the St Venant equations (De St Venant, 1871)

which describe this problem mathematically was first given analytically by Ritter [1892] and numerically by Faure and Nahas [1961].

It states that the wave front advances with a constant speed $2\sqrt{gh_0}$, while the negative wave moves back with a constant speed $\sqrt{gh_0}$ where g is the gravity. Between these two end points, the average speed and water depth are respectively given by

$$U=\frac{2}{3}\left(\frac{x}{t}+\sqrt{gh_0}\right) \quad (1)$$

$$\sqrt{gh}=\frac{1}{3}\left(2\sqrt{gh_0}-\frac{x}{t}\right) \quad (2)$$

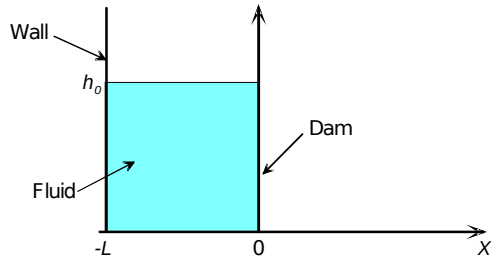


Fig. 1: Configuration of horizontal dam-break flow at negative time

Whitham [1955] provides the following analytical solution for the wave front celerity when the bottom friction affects significantly the propagation of the leading tip:

$$U=\sqrt{gh_0}\left[2-3.452\left(kt\sqrt{g/h_0}\right)^{1/3}\right] \quad (3)$$

where k is a non-dimensional coefficient of friction = g/c_0^2, c_0 denoting the Chezy coefficient. In fact, in the field, fluids are generally muddy, so that Newton's law can no longer describe the fluids rheological behaviour and it is to be expected that the waves' average velocity is affected by these characteristics.

In the second section of the present paper, a brief description of the rheology of argillaceous mud is provided and the equation of motion is given in a non-dimensional form, while the analytical solution is provided in the third section.

2. THE MODEL

2. 1 Rheology

Since the flow is isochoric, it is governed by the conservation of mass and momentum equations written, respectively, in the following form:

$$div\,u=0 \qquad \rho\frac{d\vec{u}}{dt}=\rho\vec{g}-\nabla p+\vec{div}[2\eta D(\vec{u})] \quad (4)$$

where u is the velocity field, p the pressure field, ρ the fluid density, and g gravity. We assume that a relationship exists between the stress and the strain rate. Nguyen and Boger [1983] showed that a debris flow conforms to a viscoplastic behaviour scheme governed by the Herschel-Bulkley rheological equation (Herschel and Bulkley, 1926). The Herschel-Bulkley model is written as follows:

$$\underline{\tau}=2\eta.\underline{D(u)} \quad \text{if} \quad -\tau_{II}>s^2 \quad (5)$$

$$\underline{D(u)}=\underline{0} \quad \text{if} \quad -\tau_{II}\le s^2 \quad (6)$$

where $\underline{\tau}$ denotes the stress tensor, $\underline{D(u)}$ the strain-rate tensor, and η is the apparent viscosity defined as:

$$\eta=k\left(-4D_{II}\right)^{\left(\frac{n-1}{2}\right)}+\frac{s}{\left(-4D_{II}\right)^{1/2}} \quad (7)$$

k being the fluid consistency, s its yield stress and n the power-law index, while $T_{II}=-\frac{tr\left(\underline{T}^2\right)}{2}$ denotes the second invariant of a given tensor $\underline{T}$ (strain-rate tensor and

stress tensors in this case). At this point, some assumptions can be made to simplify the equations (Laigle & Coussot, 1994, Laigle, 1996). However, a more general approach can be formulated without resorting to such assumptions (Piau, 1996).

2. 2 Equations of motion

As generally encountered in nature, the flow is assumed to be two-dimensional in the (x,y) plane of an orthonormal coordinate system (O,x,y,z) in a space domain which is much longer $(x-direction)$ than it is high $(y-direction)$. The following characteristic quantities can be defined: height *H* (e.g.: average fluid depth), velocity U_0 (e.g.: Q/H where Q denotes the average flow rate), length *L* (e.g.: distance along which the flow takes place) and a small parameter ε given by $\varepsilon=\frac{H}{L}$, the following set of normalized variables are defined:

$$X=\frac{x}{L} \qquad Y=\frac{y}{\varepsilon L} \qquad T=\frac{U_0}{L}t \qquad U=\frac{u}{U_0} \tag{8}$$

$$V=\frac{v}{\varepsilon U_0} \qquad P=\frac{p}{\rho U_0^2} \qquad \xi=\frac{\varepsilon L}{U_0}\left(-4D_{II}\right)^{1/2} \tag{9}$$

The equations of conservation of mass and momentum (3) then take the following non-dimensional form (Piau, 1996):

$$U_{,t}+\beta UU_{,x}+\left(g\cos(\alpha)+\frac{s\varphi}{\rho h}\operatorname{sgn}h_{,x}\right)h_{,x}=g\left(\sin\alpha-J\right) \tag{10}$$

$$h_{,t}+Uh_{,x}+hU_{,x}=0 \tag{11}$$

$$U_{,x}dx+U_{,t}dt=dU \qquad h_{,x}dx+h_{,t}dt=dh \tag{12}$$

where α is the sloping angle of the (Oxz) plane to the horizontal plane, β the average form coefficient, φ the exact form coefficient and J the friction slope. The set of equations (10)-(12) constitutes the mathematical model of the inertial dam-break one-dimensional flow of a Herschel-Bulkley fluid along a (Oxz) plane sloping at an angle α to the horizontal plane. The friction slope J appearing in this solution must be determined. Its expression is known only for the power-law fluid (Piau & Debiane, 2005) but not for the Herschel-Bulkley fluid.

2.3 The friction slope

Whatever the rheological equation of the fluid is, the friction slope is defined as $J=-\frac{1}{\rho g}\frac{\partial\tau}{\partial y}$ so, after one quadrature, we get:

$$\tau=-\rho gJ\left(y-h\right) \tag{13}$$

The Herschel-Bulkley fluid is characterized by two zones (Nsom et al., 1999): a rigid zone where the shear stress is less than the yield value, and which is thus characterized by $y \geq h-\frac{s}{\rho gJ}$ and a sheared zone where the shear stress is greater than the yield value, and which is thus characterized by

$$k\frac{\partial\left(\dot{\gamma}\right)^n}{\partial y}=-\rho gJ \tag{14}$$

The two zones are separated by the yield surface of the height $y_0 = h - \frac{s}{\rho g J}$. Both the rigid zone and the yield surface undergo a rigid body motion, while the shearing rate is obtained after an integration of equation (14) from y_0 to y as:

$$\frac{\partial u}{\partial y} = \left(\frac{\rho g J}{k}\right)^{1/n} \left(y_0 - y\right)^{1/n} \tag{15}$$

Integrating equation (15) from 0 to y and putting $m = 1/n$, we can write the velocity profile in a given section in the form:

$$u = u_{\max}\left[1 - \left(1 - \frac{y}{y_0}\right)^{m+1}\right] \qquad \text{for} \qquad y \le y_0 \tag{16}$$

$$u = u_{\max} \qquad \text{for} \qquad y \ge y_0 \tag{17}$$

where

$$u_{\max} = \frac{1}{m+1}\left(\frac{\rho g J}{k}\right)^{m} y_0^{m+1} \tag{18}$$

Based on this solution, the flow rate and the average velocity are derived respectively as:

$$Q = \int_0^{y_0} u \cdot dy + \int_{y_0}^{h} u \cdot dy = \frac{u_{\max}}{m+2}\left[(m+2)h - y_0\right] \tag{19}$$

$$U_1 = \frac{Q}{h} = \frac{\left(\frac{\rho g J}{k}\right)^{m} h^{m+1}\left(1 - \frac{s}{\rho g h J}\right)^{m+1}}{(m+2)(m+1)}\left((m+1) + \frac{s}{\rho g h J}\right) \tag{20}$$

The expression of the friction slope given by Piau and Debiane [2005] for a power-law fluid is recovered by putting $s = 0$ in equation (20), i.e.:

$$J = \frac{k}{\rho g}\left(\frac{m+2}{h^{m+1}}\right)^{1/m} U_1^{1/m} \tag{21}$$

and the following well-known formula for the Newtonian fluid: $s = 0$, $m = 1$, $k = \mu$

$$J = \frac{3\mu}{\rho g h^2} U_1 \tag{22}$$

In the general case of a Herschel-Bulkley fluid ($s \neq 0$, $m \neq 1$), it is difficult to derive an analytical expression of the friction slope from equation (20). Meanwhile, the following approximate expression can be obtained for the limiting case of the inertial regime:

$$J = (m+2)^{1/m} \frac{k}{\rho g h}\left(\frac{U_1}{h}\right)^{1/m} + \frac{s}{\rho g h} \tag{23}$$

3. SOLUTION

3.1 The general solution

Putting $a=\left(gh\cos(\alpha)+\frac{\varphi s}{\rho}\mathrm{sgn}h_{,x}\right)^{1/2}$ if $h\geq\frac{-\varphi s}{\rho g\cos(\alpha)}\mathrm{sgn}h_{,x}$ (24)

and $a=\left(-gh\cos(\alpha)-\frac{\varphi s}{\rho}\mathrm{sgn}h_{,x}\right)^{1/2}$ if $h\leq\frac{-\varphi s}{\rho g\cos(\alpha)}\mathrm{sgn}h_{,x}$ (25)

equation (10) becomes:

$$U_{,t}+\beta UU_{,x}\mp\frac{a^2}{h}h_{,x}=g\left(\sin(\alpha)-J\right) \quad (26)$$

The solution to system (10)-(11)-(12) is obtained if its determinant is zero, i.e.

$$\left(\frac{dx}{dt}\right)^2-U(\beta+1)\frac{dx}{dt}+\left(\beta U^2\pm a^2\right)=0 \quad (227)$$

Provided that $h\geq\frac{-\varphi s}{\rho g\cos(\alpha)}\mathrm{sgn}\left(h_{,x}\right)$, we obtain:

$$\frac{dx}{dt}=U\mp a \quad (28)$$

where we used $\beta=1$, as usual in hydraulics. The system is undetermined if any of the columns of its determinant is replaced by the system's right-hand side. So

$$\frac{dU}{dt}\mp\frac{a}{h}\frac{dh}{dt}=g\left(\sin(\alpha)-J\right) \quad (29)$$

Using

$$h=\frac{a^2}{g\cos(\alpha)}-\frac{\varphi s}{\rho g\cos(\alpha)}\mathrm{sgn}h_{,x} \quad (30)$$

we get

$$\frac{dU}{dt}\mp\frac{2a^2}{a^2-\frac{\varphi s}{\rho}\mathrm{sgn}h_{,x}}\frac{da}{dt}=g\left(\sin(\alpha)-J\right) \quad (31)$$

To solve this differential equation we put $r(a)=a\left(\frac{\rho}{\varphi s}\right)^{1/2}$ and we search a function $f(r)$ such as

$$\mp\frac{2r^2}{r^2-\mathrm{sgn}h_{,x}}\frac{da}{dt}=\frac{d\left(\mp 2a\left(1-r^{-1}f(r)\right)\right)}{dt} \quad (32)$$

so $f(r)$ is the solution of:

$$\frac{-r^2\mathrm{sgn}h_{,x}}{r^2-\mathrm{sgn}h_{,x}}=\frac{df(r)}{dr} \quad (33)$$

If $r=tg(\delta)$, the following solution is obtained after a straightforward algebraic handling:

$$f(r)=arctg(r) \quad \text{or} \quad f(r)=ar\coth(r) \quad (34)$$

according to which $\mathrm{sgn}h_{,x}=-1$ or $\mathrm{sgn}h_{,x}=+1$ respectively and the equation of motion (31) becomes:

$$\frac{d\left[U \mp 2a\left(1-r^{-1}f(r)\right)\right]}{dt}=0 \tag{35}$$

It can be solved using the method of characteristics. As $h_{,x}<0$, it results that

$$a=\left(gh-\frac{\varphi s}{\rho}\right)^{1/2} \qquad \text{and} \qquad r=a\left(\frac{\rho}{\varphi s}\right)^{1/2} \tag{36}$$

Indeed, two characteristics of equation (35) can be defined. The negative characteristic where the gradient dx/dt is constant and equal to $U-a$ and the quantity $U-2a\left(1-r^{-1}f(r)\right)$ remains constant. While the positive characteristic along which the gradient is constant and equal to $U+a$ and the quantity $U+2a\left(1-r^{-1}f(r)\right)$ remains constant.

3.2 The law of propagation of the wave front

Let's consider, at the initial time, the negative characteristic noted C_0^- in such a way that the fluid height is constant and equal to h_0 and the average velocity is zero. Thus, in the $(x-t)$ plane, its graph is a straight line with a slope

$$\frac{dx}{dt}=-a_0 \qquad \text{where} \quad a_0=\left(gh_0-\frac{\varphi s}{\rho}\right)^{1/2} \tag{37}$$

At time $t=\theta$, the negative characteristic C_θ^- is such that the gradient is given by

$$\frac{dx}{dt}=U(\theta)-a(\theta) \tag{38}$$

Along a positive characteristic C^+ intersecting both C_0^- and C_θ^-, the quantity $U+2a\left(1-r^{-1}f(r)\right)$ remains constant, so that

$$U+2a\left(1-r^{-1}arctg(r)\right)=2a_0\left(1-r_0^{-1}arctg(r_0)\right) \tag{39}$$

where $r_0=a_0\left(\frac{\rho}{\varphi s}\right)^{1/2}$. Introducing equation (39) in equation (38) we obtain:

$$\left(\frac{dx_b}{dt}\right)_\theta=2a_0\left(1-r_0^{-1}arctg(r_0)\right)-2a\left(1-r^{-1}arctg(r)\right)-a(\theta) \tag{40}$$

The propagation velocity of the negative wave is derived from equation (40) as:

$$a(\theta)=a_0 \qquad \forall\theta \qquad \text{and} \qquad \left(\frac{dx_b}{dt}\right)_\theta=-a_0 \tag{41}$$

To find the propagation velocity of the wave front in a Newtonian fluid, it is assumed that the fluid height is zero at the front. For a viscoplastic fluid, we assume that at the front, the fluid height is $\varphi s/(\rho g)$ because if this is not the case the set of equations of motion (10)-(11)-(12) is determined and there is no wave propagation. So, applying this condition we get $a=0$ while the equations (38)-(39) give:

$$\left(\frac{dx_f}{dt}\right)_\theta = 2a_0\left(1-r_0^{-1}arctg(r_0)\right) \tag{42}$$

As the gradient along a given characteristic is constant, we can write

$$\left(\frac{dx}{dt}\right)_\theta = \frac{x}{t} \tag{43}$$

So, equation (40) gives:

$$\frac{x}{t} = 2a_0\left(1-r_0^{-1}arctg(r_0)\right) - 2a\left(1-r^{-1}arctg(r)\right) - a(t) \tag{44}$$

where $r = a_0\left(\frac{\rho}{\varphi s}\right)^{1/2}$. Adopting the following set of non-dimensional variables

$$x^* = \frac{x}{H} \qquad h^* = \frac{h}{H} \qquad t^* = t\sqrt{\frac{g}{H}} \tag{45}$$

where H is the reference height. If we note $\phi_c = \frac{\varphi s}{\rho g H}$, equation (44) has the following non-dimensional form:

$$\frac{x^*}{t^*} = \phi_c^{1/2}\left(2r_0 - 2arctg(r_0) + 2arctg(r) - 3r\right) \tag{46}$$

where $r = \left[\frac{h^*}{\phi_c} - 1\right]^{1/2}$ $\qquad \phi_c \le h^* \le 1$ $\qquad r_0 = \left[\frac{1}{\phi_c} - 1\right]^{1/2}$ (47)

and equations (41)-(42) give:

$$\frac{dx_b^*}{dt^*} = \left[1-\phi_c\right]^{1/2} \qquad \frac{dx_f^*}{dt^*} = 2\phi_c^{1/2}\left[2r_0 - arctg(r_0)\right]^{1/2} \tag{48}$$

Note that if we put $s = 0$ in equations (46)-(48), we recover the well-known results for Newtonian fluids, which are:

$$\frac{x^*}{t^*} = \left(2-3h^*\right)^{1/2} \qquad \frac{dx_b^*}{dt^*} = 1 \qquad \frac{dx_f^*}{dt^*} = 2 \tag{49}$$

4CONCLUSION

The flow generated by the collapse of a dam retaining a muddy fluid was considered theoretically. The fluid is assumed to be viscoplastic according to the Herschel-Bulkley

rheological law. The equations of conservation of mass and momentum were written in a non-dimensional form based on the shallow water approximation. Assuming the inertial effects dominant against the viscous ones in order to model a flash flood event, an equation of motion was built. Then, an implicit equation was formed for the friction slope. The equation of motion could then be solved using the method of characteristics. The law of propagation of the front of the positive wave was provided, as well as that of the back wave. The corresponding laws of propagation for the Newtonian fluid and for the power law fluid were recovered when appropriate values of the rheological parameters were used. Piau [1996] gave similar *ad hoc* relations. In this paper, a proof was given of this solution and it is a crucial step for implementing the kinematic wave theory for the muddy fluids (Arattano & Savage, 1994) in our model.

REFERENCES

Arattano, M. & Savage, W.Z., Modelling debris flows as kinematic waves, *Bulletin of the IAEG*, 49, 95-105, 1994

De Saint-Venant, B., Théorie du mouvement non permanent des eaux, *C.R. Acad. Sci*., 73, 147-154 ; 217-225, 1871 (in French)

Dressler, R.F., Hydraulic resistance effects upon the dam-break functions, *J. Res. Nat. Bur. Stand*., 49(3), 217-225, 1952

Dressler, R.F., Comparison of theories and experiments for hydraulic dam-break wave, *Int. Assoc. Sci. Pubs*., 3(28), 319-328, 1954

Faure, J. & Nahas, N., Etude numérique et expérimentale d'intumescences à forte courbure du front, *La Houille Blanche*, 5, 576-587, 1961 (in French)

Herschel, W.H. & Bulkley, R., Über die Viskosität und Elastizität von Solen, *Am. Soc. Test. Mat*., 26, 621-633, 1926 (in German)

Hunt, B., Newtonian fluid mechanics of debris flows and avalanches, *J. Hydr. Eng., ASCE*, 120(12), 1350-1362, 1994

Laigle, D. & Coussot, P., Modélisation numérique des écoulements de laves torrentielles, *La Houille Blanche*, 3, 50-55, 1994 (in French)

Laigle, D., Two-dimensional modelling of debris flow spreading on alluvial fans, in "Hydroinformatics'96", 651-657, Balkema, Rotterdam, The Netherlands, 1996

Nguyen, Q.D. and Boger, D.V., Yield stress measurement for concentrated suspensions, *J. Rheol.*, 27, 321-349, 1983

Nsom, B., Debiane, K., Piau, J.M. & Ayadi, A., Gradually varied flow of unloaded mud in open channel, *Int. Appl. Mech. Eng.* 4(3), 443-470, 1999

Nsom, B., Debiane, K. & Piau, J.M., Bed slope effect in the dam-break problem, *J. Hydr. Res. (IAHR)*, 38(6), 459-464, 2000

Nsom, B., Horizontal Viscous Dam-Break Flow: Experiments and Theory, *J. Hydr. Eng., ASCE*, 128(5), 543-546, 2002

Piau, J.M., Flow of a yield stress fluid in a long domain. Application to flow on an inclined plane, *J. Rheol.*, 40, 711-723, 1996

Piau, J.M. & Debiane, K., Consistometers rheometry of power-law viscous fluids, *J. N.N.F.M.*, 127, 213-224, 2005

Ritter, A., Die Fortpflanzung der Wasserwellen, *Ver Deutsch Ingenieure Zeitschrift*, 36(33), 947-954, 1892 (in German)

Su, S. & Barnes, A.H., Geometric and Frictionnal Effects on Sudden Releases, J. Hydr. Div. ASCE, 96(HY11), 2185-2200, 1970

Whitham, G.B., The effects of hydraulic resistance in the dam-break problem, *Proc. Roy. Soc.*, A(227), 399-407, 1955

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Mud-flow and lava-flow susceptibility and hazard mapping through numerical modelling, GIS techniques, historical and geo-environmental analyses

Giulio Iovine
CNR-IRPI via Cavour, 6 – 87036 Rende (CS) ITALIA (g.iovine@irpi.cnr.it)

Abstract: A method for mapping mud flows and lava flows, and for evaluating related susceptibility and hazard has recently been tested at different study areas of Southern Italy (located in Campania, Calabria, and Sicily), which have repeatedly been affected by damaging events in historical time. The approach is based on numerical models, GIS-techniques, geo-environmental and historical evaluations. Results have been obtained through a statistical approach by simulating a large number of events on a cluster. Employed cellular automata models have first been calibrated, either manually (by trial and error) or by means of Genetic Algorithms, and then validated against past flow events that occurred in the same study areas, or in similar geo-environmental settings. Simulations have been quantitatively evaluated with respect to real cases by means of two distinct functions of fitness based on 1) the affected areas for mudflows, and on 2) affected areas and duration for lava flows. Aiming at susceptibility/hazard assessment, a grid of possible sources has been hypothesised on the basis of historical/geological knowledge and statistics of past events. For each source, a high number of simulations has been planned by adopting combinations of sources' and materials' characteristics. Probabilities of activation, empirically based on past events, have been assigned to each source of the grid by considering its location and geological information. Different probabilities have also been assigned to each "type of event" by taking into account their observed historical frequencies. Two different types of maps were realized in a GIS: 1) a susceptibility map, realised by simply counting the frequencies of flows affecting each site; 2) a hazard map, in which probabilities have been "empirically" attributed to each simulation based on location of sources and types of event. Preliminary results (based on a subset of the overall planned simulations) clearly depict the most susceptible and hazardous sectors.

Keywords: Lava flows; Mud flows; Modelling & Simulation; Susceptibility; Hazard.

1. INTRODUCTION

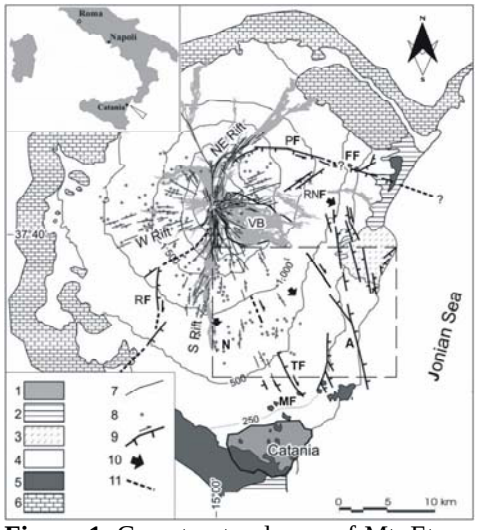
In so far as mud flows and lava flows are considered, hazard mapping procedures require - besides the usual prediction of time of occurrence and of source location - the assessment of the travel distance (run out), or better the area affected by the propagation of the initial event and of its energetic characteristics. For these types of phenomena, hazard mapping procedures are sometimes based on empirical geological reasoning, or they depend on either empirical-statistical evaluations or statistical-probabilistic analyses of past events [cf. D'Ambrosio et al., 2007; Crisci et al., in press - and references therein]. Available approaches are usually distinguished into deterministic and probabilistic methods, or a combination of both techniques.
Modelling and simulation represent a valuable tool for hazard evaluation. By employing such techniques, the timely simulation of the possible flow paths and the evaluation of the effects of planned remedial/control works on the behaviour of future events could profitably be carried out. Following Hungr [1995], any simulation model should be easy to

calibrate through back-analyses against real cases; it should describe the essential characteristics of the phenomenon, allowing also for simulating flow branching and rejoining; it should not suffer from mesh-related problems. As concerns model reliability, one possible approach of evaluation relies on thorough calibration and successive validation by considering a sufficient number of case studies. Sensitivity analyses may further help in evaluating the role of model parameters, mesh geometry and quality of input data on the model behaviour [cf. e.g. D'Ambrosio et al., 2007]. Ideally, a multiphase 3D model of material which moves over a complex topography would be an optimal choice for accurate assessment. However, this type of model generally requires large computational power and a number of parameters which may demand considerable efforts for calibration. Moreover, they generally suffer from instability problems and commonly need site-specific settings. A proper but simplified approach would seem appropriate.
In the past decades a variety of numerical models were successfully developed for simulating various types of gravity flows [e.g. Crisci et al., 1982; Murray and Paola, 1994; Miyamoto and Sasaki, 1997; Iverson et al., 1998; Di Gregorio and Serra, 1999; D'Ambrosio et al., 2001; Iovine and Di Gregorio, 2003; Pitman et al., 2003]. Within the frame of parallel computing, Cellular Automata (CA) models are able to simulate complex systems, which can be described in terms of local interactions. Well known CA examples are Lattice Gas Automata and Lattice Boltzmann models [cf. Succi, 2004], which are particularly suitable for modelling fluid dynamics at a microscopic scale. However, many natural phenomena – and, among these, gravity flows - generally affect quite large areas and need a macroscopic level of description. Standard approaches based on differential equations [cf. McBirney and Murase, 1984; Iverson, 1997 – and many others] may suffer in capturing the essential features of the phenomena to be simulated; employed equations are commonly rather complex and involve adoption of approximated numerical methods of solution. In such cases, Macroscopic Cellular Automata (MCA) represent a suitable alternative [Di Gregorio and Serra, 1999].
Analyses of lava-flow and mud-flow susceptibility/hazard have recently been started in separate study areas of Southern Italy, by employing MCA models combined with GIS techniques [e.g. Iovine et al., 2003a, b; Crisci et al., in press]. Historical and geo-environmental analyses, aimed at quantitatively evaluating return periods of considered event types are presently in progress, as well as part of the overall planned simulations. In the following only preliminary results are therefore briefly described, with reference to 1) lava flows on the SE flank of Mount Etna (Sicily), and to 2) mud-flows on the southern slope of the Pizzo d'Alvano Massif and in the San Martino Valle Caudina-Cervinara area (Campania). Finally, an example of susceptibility zonation is also reported for another sector, located along the southern Tyrrhenian coast of Calabria between the villages of Bagnara and Scilla, not far from the Messina Strait.

2. THE STUDY AREAS

Among the study areas considered for lava-flow hazard analyses, one is located on the SE flank of Mt. Etna, the largest sub-aerial active volcano of Europe (Figure 1). Etna is composed of several nested strato-volcanoes plus scattered eruptive centres, which opened through a basement of tholeiitic/transitional fissural basalts dating back to 500 ky BP. The volcanic basement overlies Middle-Late Pleistocene sediments covering the Maghrebian-Apennine Chain (CNR, 1979). The dominant type of Etnean activity is effusive. In historical times, the eastern flank of the volcano was frequently affected by lava flows originating from flank eruptions [e.g. Romano and Sturiale, 1982], either through lateral fractures or eccentric vents. From a structural point of view, two main "weakness zones" are to be found by the volcano summit: the "N-S Rift" and the "NE Rift". Moreover, a high spatial density of fractures, effusive fissures and pyroclastic cones characterizes the eastern flank of the volcano [Behncke et al., 2005], mainly along a SSE-trending fracture system which crosses the study area between the villages of Nicolosi and Trecastagni [Corazzato and Tibaldi, 2006]. Finally, a dormant deep-seated gravitational sliding (DGSD) towards the Jonian Sea, associated with a larger volcano-tectonic phenomenon, affects the whole south-eastern flank of the volcano [Borgia et al., 1992]. The margins of the DGSD are marked by the mentioned major weakness zones: its right flank crosses the study area and

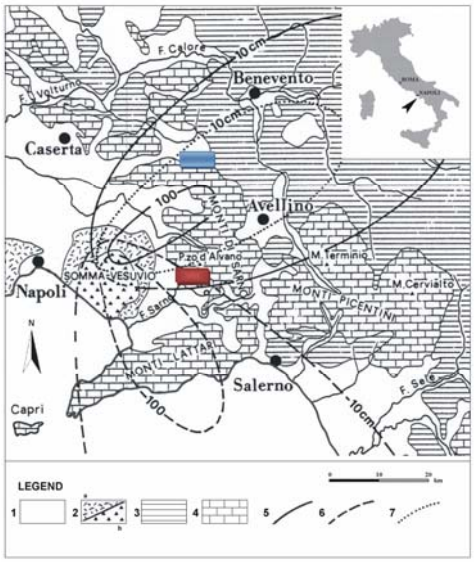
corresponds to the "transitional" weakness zone which connects the N-S Rift to the transtensional fault system marked by the Trecastagni and Mascalucia faults.



In the map on the left, the study area is shown by a dashed rectangle. Symbols and grey tones have the following meanings: 1) 1900-2002 lava flows; 2) recent alluvial deposits; 3) deposits resulting from the emptying of Valle del Bove; 4) Etnean volcanics; 5) Early Quaternary clays; 6) Pre-Quaternary sediments; 7) main eruptive fractures; 8) pyroclastic cones; 9) main faults; 10) direction of displacement of either the DGSD or the volcano-tectonic phenomenon; 11) boundaries of the unstable sector; FF) Fiumefreddo Fault; MF) Mascalucia Fault; PF) Pernicana Fault; RF) Ragalna Fault; RNF) Ripe della Naca Fault; TF) Trecastagni Fault; A) Acireale; N) Nicolosi; VB) Valle del Bove.

Figure 1. Geo-structural map of Mt. Etna - after Acocella and Neri [2003], mod.

As far as "mudflows" [*sensu* United Nations, 1997] are concerned, the first study area is located in the Caudina Valley on the northern slope of Mt. Pizzone, which belongs to the Partenio-Monti di Avella massif (blue rectangle in Figure 2). Still in the same region, the second study area is located about 15 km east of Mt. Vesuvius, on the southern slope of Mt. Pizzo d'Alvano, belonging to the Monti di Sarno massif (red rectangle in Figure 2).



Mesozoic carbonate rocks constitute the backbone of both the massifs, which are bounded by steep slopes cut along major tectonic elements. The bedrock is overlain by a regolith of residual and colluvial soils, mainly derived from Quaternary pyroclastic materials of the Somma–Vesuvius volcanic complex and of the Phlegrean Fields. The thickness of the regolith, and its geotechnical properties vary widely from site to site. At the base of the massifs, thick colluvial and alluvial deposits mark the transition to the alluvial plains; here, at the mouth of the main drainages, voluminous debris fans are to be found. Several villages with their related infrastructures are to be found in the transitional sector at the base of the massifs.

Figure 2. Geological map of the study areas in Campania - after Del Prete et al. [1998], mod. Key: (1) alluvial sediments; (2) pyroclastic terrains (a) and lavas (b); (3) clay, marl and sandstone; (4) carbonate; (5–7) isopleths (in cm) of the Somma-Vesuvius air-fall deposits of year 1800 B.C. (5), 79 A.D. (6) and 472 A.D. (7).

At Pizzo d'Alvano, hundreds of soil slip-debris flows were triggered by heavy rains on 5–6 May 1998 [Del Prete et al., 1998], severely inundating the villages of Sarno, Siano, Bracigliano and Quindici (and causing 161 deaths). In the Caudina Valley, the villages of Cervinara and San Martino V.C. were particularly affected by the soil slip-debris flows triggered by the meteoric event of 15–16 December 1999 [Vittori et al., 2000].

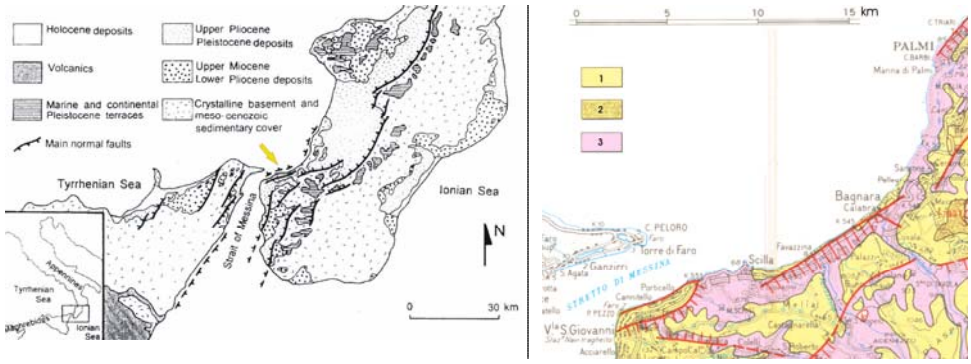


Figure 3. Geological sketches of the Calabrian study area. a) On the left, a geological map (location marked by the yellow arrow) - after Tortorici et al. [1995], mod. b) On the right, a litho-structural map. Key: 1) terraced sand, clay, pebble (Quaternary); 2) sand, clay, and marl, with subordinate evaporite (Upper Miocene-Pliocene); 3) acid intrusive and middle-high grade metamorphic rocks of the Alpine continental nappe (Palaeozoic). In red, recent normal faults - after Sorriso-Valvo & Tansi [1996], mod.

Finally, the third study area for mud-flow susceptibility analyses is located in Calabria, between the villages of Bagnara Calabra and Scilla (Figure 3). This narrow strip of land along the southern Tyrrhenian coast is characterized by: notable slope energy in a terraced geomorphologic setting; short ephemeral streams with small watershed; high longitudinal gradients. Palaeozoic intrusive and metamorphic rocks that crop out along the coast and the main streams are overlain by Upper Miocene-Pliocene pelitic sediments and evaporites, which are in turn followed by Quaternary sediments. A recent NE-SW trending normal fault, which occurs along the coast, has a notable morphological evidence.
In the area, heavy rains frequently trigger soil slips on the steep slopes facing the sea; in the last 40 years, one event of single/multiple landslide activation occurred on the average every three years. The mobilized materials commonly became fluidized and ran towards the base of the slopes, damaging the villages and the major infrastructures (highway, state road SS.18, railway) located thereby. Among the recentmost damaging events, those that occurred on 12 May 2001 and on 31 March 2005 struck the hamlet of Favazzina and its surroundings, severely damaging (and interrupting) the cited infrastructures.

3. THE METHOD

Details concerning the models employed for simulating mudflows and lava flows can be found in D'Ambrosio et al. [2003a] and in Crisci et al. [2004], respectively, and will not be repeated here; the main characteristics of both models are listed in the Appendix. Likewise, the approach for mapping mud-flow susceptibility was first applied by Iovine et al. [2003a, b; 2007]; the method for lava-flow susceptibility and hazard mapping can be found in Crisci et al. [in press].
The model for simulating lava flows belongs to the family SCIARA of deterministic MCA models [Crisci et al., 1982]. The two-dimensional version *fv* [Crisci et al., 2004], based on a hexagonal tessellation of space, was used in this study. In *fv*, lava feeding is handled by specifying the cell/s that will act as vent/s; for this purpose an effusion-rate function has to be assigned to each vent. Lava flows among the cells are computed by applying the "minimisation algorithm of the differences" (cf. Appendix), proposed by Di Gregorio and Serra [1999]. Changes of lava temperature are modelled by means of the radiation equation, and lava viscosity varies accordingly. As a consequence of viscosity, the amount of lava which cannot flow out of the cell is computed (in terms of adherence). Finally, the halting of the lava flow depends on temperature: when it drops below a fixed threshold the thickness of lava is added to the cell elevation. The elementary processes included in the transition function σ of the model are: σ_1) lava outflows, σ_2) variation of lava temperature, and σ_3) lava halting.
Calibration was performed against the Etnean 2001 Mt. Calcarazzi eruption by employing parallel Genetic Algorithms (GA). Simulations were compared to real cases by means of the fitness function $e_{1t} = \sqrt{(e_1)^{t_1}(e_1)^{t_2}}$, defined in the interval [0,1] (where 1 corresponds

to perfect overlap), which takes into account both the areal extent of the real and simulated events, and their temporal duration. Note that, $e_1 = \sqrt{(R \cap S)/(R \cup S)}$ is the "basic" fitness function, which only compares the affected areas of the real (R) and simulated (S) events. Moreover, t_1 represents the duration of the real event, while $t_2 = t_1 + \Delta t$, where Δt is an arbitrary time increment. Validation was then performed against the 2002 Linguaglossa, and the 1991-93 Valle del Bove events, which are characterised by different durations, lava volumes, and effusion rates [cf. D'Ambrosio et al., 2006].

For susceptibility/hazard mapping purposes, a regular grid of hundreds of "possible" vents, located at 500 m intervals and uniformly covering the area, was hypothesised (Figure 4a). Aiming at pre-processing the simulation results, a sub-set of n.88 vents (located at 1 km intervals) was employed, allowing for the preliminary considerations discussed in the present paper. For each vent, a statistically-significant number of simulations was planned by adopting appropriate combinations of durations and lava volumes (Table 1), and effusion-rate functions (e.g. Figure 4c), which were selected on the basis of volcanological/historical knowledge. Accordingly, it was implicitly assumed that the volcanic style will not significantly change in the near future.

Performed simulations were stored in a GIS for susceptibility and hazard analyses. Maps for susceptibility evaluations were obtained by simply overlapping all the simulations; in other words, the same probability of occurrence was assigned to all types of events. By normalizing the obtained frequencies of events, five equal-interval classes of increasing "relative susceptibility" were distinguished and mapped by means of grey-tones (adopted ranges: 1=[0-0.2], 2=]0.2-0.4], 3=]0.4-0.6], 4=]0.6-0.8], 5=]0.8-1.0]).

As for hazard evaluations, historic and volcanological information on past eruptions was also taken into account. Probabilities of activation ($h_a = h_{elev} \cdot h_{wz} \cdot h_{df}$), empirically based on past behaviour of the volcano, were assigned to each vent of the grid by considering i) elevation (h_{elev} - Figure 4b), and ii) location with respect to the volcanic edifice (cf. Figure 5), and particularly the proximity to the main weakness zones (h_{wz}) and the spatial densities of eruptive fissures (h_{df}). Different probabilities ($h_e = h_{dv} \cdot h_{\gamma}$) were also assigned to the simulated eruption types, depending on combinations of durations and lava volumes (h_{dv}), and effusion-rate functions (h_{γ}). At any place affected by a given simulation (s_i), a value of hazard ($h_i = h_a \cdot h_e$) could be ascribed, depending on the conditioned probabilities of occurrence of that simulation. The overall hazard ($H_t = \sum h_i$) could be obtained by simply overlapping all the simulations in the GIS and by adding the related probabilities. Depending also on the number (n) of performed simulations, computed values ranged between two extremes; they were normalised and ranked the into five classes of "relative hazard" (H_r) - mapped in grey-tones (same ranges as for susceptibility).

Based on adopted criteria for probability evaluation, different maps of lava-flow hazard could be compiled [Crisci et al., in press]. In the example shown in the next section, a significant role was ascribed to the SSE-trending weakness zone on the right flank of the cited DSGD (cf. Figures 1 and 4). In more detail, h_{elev} was assigned by assuming a cubic relationship with vent elevation (cf. tendency line in Figure 4b, formula in caption), according to historical data [after Behncke et al., 2005]; h_{df} was set to 0.1, 0.3, and 0.6, respectively, depending on spatial density (ρ, number of vents per square km) of observed eruptive fissures and pyroclastic cones, which according to the same Authors can be distinguished into three classes ($\rho \leq 0.5$, $0.5 < \rho \leq 1$, and $\rho > 1$); h_{wz} assumed an "emphasising" role, by increasing by 30% values of activation probabilities of the vents located within its influence buffer (i.e. the weakness zone). Depending on possible combinations of h_{df} and h_{wz}, six classes of normalised activation probabilities resulted (see description of Figure 5).

Frequencies of historical eruptions were ranked into classes by durations and lava volumes, based on volcanological and historical data. Normalised probabilities (h_{dv}) for the types of event were derived from such frequencies (cf. Table 1, where only the types of eruption with shortest durations - considered in the present study - are listed). Finally, one single type of effusion-rate function was adopted for preliminary analyses (i.e. h_{γ} was ignored), characterized by a peak discharge at 1/3 of the eruption duration (cf. Figure 4c).

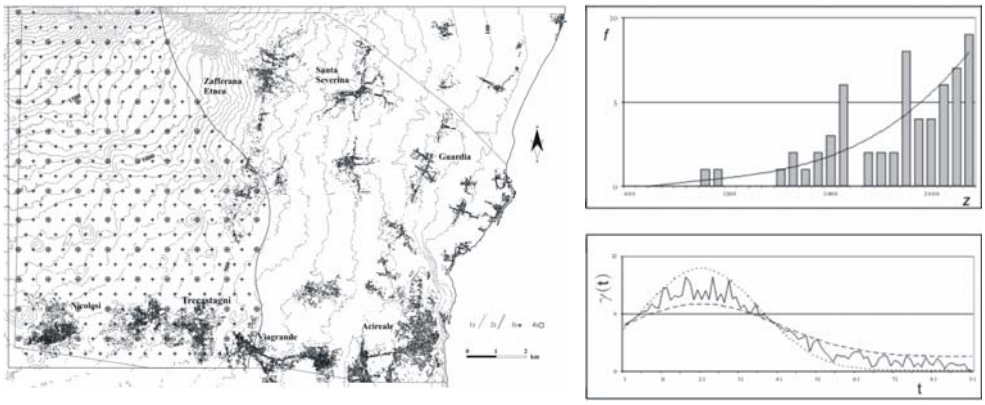
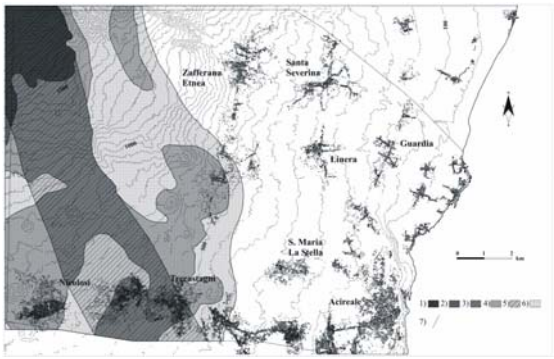


Figure 4. SE flank of Mt. Etna. a) On the left, grid of hypothesised vents. Key: 1) boundary of the study area; 2) eastern limit of the vents; 3) set of n.340 vents, located at 500 m intervals. 4) set of n.88 vents, at 1 km intervals. b) On top-right, relationship between number of historical lava flows (*f*) and elevation of the vents (*z*, in m a.s.l.). Tendency line is expressed by: $h_{elev} \propto 8 \cdot 10^{-10} \cdot z^3 + -2 \cdot 10^{-6} \cdot z^2 + 1.6 \cdot 10^{-3} \cdot z + 3.009 \cdot 10^{-1}$. c) On bottom-right, example of randomly-generated effusion-rate function: dotted lines define the variation range in which values were computed – after Crisci et al. [in press], mod.



In the map on the left, the major SSE-trending weakness zone, which marks the right margin of the DGSD (cf. Figure 1), is shown by a dashed pattern. As far as locations of the vents are considered, the six classes of assumed probabilities ($h_{df} \cdot h_{wz}$) were normalised as follows: 1=0.34, 2=0.26, 3=0.17, 4=0.13, 5=0.06, 6=0.04.

Figure 5. SE flank of Mt. Etna. Activation probability of the vents, as a function of spatial densities (ρ) of historical eruptive fissures and pyroclastic cones, and of proximity to the main weakness zone. Key: 1-6) classes of probability (see text); 7) boundary of the study area – after Crisci et al. [in press], mod.

Table 1. Statistics on lava-flow events at Mt. Etna in the last 400 years (data from Behncke et al., 2005). Historical events are distinguished in classes by duration (in days) and emitted volume of lava (in 10^6 m^3): observed frequencies and, in parentheses, related probability (in %) are listed. In Italics: interpolated values. Asterisks: unrealistic types.

duration \ *volume*	] 0÷32]	] 32÷64]	] 64÷96]	] 96÷128]	] 128÷160]
] 0÷15]	19,0 (20,4)	3,0 (3,2)	*4,2* (*4,5*)	0,0 *	0,0 *
] 15÷30]	6,0 (6,4)	3,0 (3,2)	1,0 (1,1)	*2,3* (*2,5*)	0,0 *
] 30÷60]	3,0 (3,2)	*2,9* (*3,1*)	1,0 (1,1)	*1,3* (*1,4*)	1,0 (1,1)

The model employed for simulating mud flows also belongs to a family of deterministic MCA models [cf. D'Ambrosio et al., 2003a, b] based on the cited algorithm of Di Gregorio and Serra [1999]. In this study, the simplified release $S3_{\text{-}hex}$ of SCIDDICA, first described with reference to a square tessellation by Iovine et al. [2002], was employed. Four elementary processes constitute the transition function σ of $S3_{\text{-}hex}$ [D'Ambrosio et al., 2003a]: σ_1) entrainment of regolith, according to the energy of the flow; σ_2) debris outflows (according to pressure gradients across the cells and to material properties); σ_3) update of debris thickness and energy; σ_4) energy loss and run-up.
Statistics of the May 1998 and December 1999 disasters in Campania, and of the May 2001 and March 2005 events in Calabria, were obtained through interpretation of air-photographs and geologic-geomorphologic field surveying. The amount of regolith available for entrainment was also surveyed in the field, together with other evidence

related to triggering mechanisms, flow development and rheology, erosive character, etc. The "erodability" of the regolith by the flowing masses was empirically estimated by means of visual inspection and appropriate (expeditious) cementation tests. Together with historic information on previous activity, and other geological knowledge, such data allowed us to estimate the characteristics of future phenomena in the same areas. Recent planning instruments (and several unpublished technical reports), produced by local governmental agencies for the same study areas (cf. Figures 6 and 7) in accordance with "emergency" laws (ex DL.180/98, L.267/98, L.226/99, OMI.3029/99), permitted us to imagine a set of likely sources among the sectors exposed to highest landslide potential.

For the two study areas in Campania, calibration was first carried out manually (by trial and error) on a standard platform by employing the fitness function e_1, and later refined with GA. The 1998 Curti mud flow on the southern slope of Pizzo d'Alvano and the 1999 Vallicelle mud flow on the northern slope of Mt. Pizzone were taken as reference cases. Validation was performed against similar mud-flows that occurred during the same disasters in the vicinity of the reference cases [cf. Iovine et al., 2003a, b]. In Calabria, the mud-flow events triggered on the northern slope of Mt. Serro Indice, which stroke the hamlet of Favazzina on 12 May 2001 and 31 March 2005, were selected for calibration and subsequent cross validation. Calibration was first performed manually, and only later by means of sequential GA (again, by employing the fitness function e_1). Incidentally, an improved release of the model (S_{4c}), which is substantially similar to *S3*$_{\text{-}hex}$, was employed for this study area – for details on the model, see D'Ambrosio et al. [2007].

As for lava-flow susceptibility/hazard analyses, a regular grid of hundreds of "possible" sources, located at ca. 50 m intervals along the uppermost border of each hazardous sector, was assumed. For each source, a statistically-significant number of simulations was planned by adopting appropriate combinations of source extent and model parameters. These were selected on the basis of geomorphologic/historical knowledge, statistics of past events, and of calibration results (cf. Table 2). Accordingly, a scenario of soil-slip activation (and related mud-flows), which is comparable with the recentmost disasters occurred in the same areas, was assumed. Nevertheless, as detailed historical and field analyses of past events are still in progress, to date only simplified susceptibility maps could be produced. Therefore, a rather limited number of sources were hypothesized within sectors exposed to highest hazard, using available literature and administrative plans. In particular, a sub-set of n.12 sources was selected for each of the Campanian study areas [Iovine et al., 2003a, b], corresponding to locations with the highest probability of soil-slip triggering - as shown in the maps created after the 1998 and 1999 disasters by Caruso et al. [2005], and by the Basin Authority of the rivers Liri, Garigliano and Volturno [2000], respectively (cf. Figure 6). A regularly spaced set of sources was instead selected for the Calabrian study area, along the uppermost unstable sectors of the slopes, as classified in the "Plan of Hydro-geological Asset" (PAI) landslide maps of Bagnara Calabra, Scilla and Favazzina (cf. Figure 7) by the Calabrian Basin Authority [2001]. In all these cases, source extents were chosen according to a precautionary principle, by selecting maximum values per each combination of slope angle and elevation, from available statistics of past events.

Finally, mudflow susceptibility maps were constructed through GIS overlapping and successive frequency classification, by following (roughly) the same approach as described above for lava flows; resulting frequencies were mapped through grey- or colour-scales.

For each study area, required information for the application of the models (*fv* or *S3*$_{\text{-}hex}$) were arranged as input matrices of hexagonal cells, containing: 1) elevation data (*fv* and *S3*$_{\text{-}hex}$); 2) thickness of erodable regolith (only for *S3*$_{\text{-}hex}$); 3) location and extent of the vent/s (*fv*) or landslide source/s (*S3*$_{\text{-}hex}$); and 4) real events (*fv* and *S3*$_{\text{-}hex}$). Model parameters were initialised by considering the characteristics of events to be simulated, in addition to implementation constraints. For the model *fv*, the parameters are: p_a) cell apothem; p_s) duration of a CA step; p_{Tv}) temperature of lava at the vent; p_{Ts}) temperature of lava at stopping; p_{adv}) adherence at the vent; p_{ads}) adherence at stopping; p_{cool}) cooling parameter. For *S3*$_{\text{-}hex}$, the parameters are: p_a) cell apothem; p_{adh}) adherence; p_f) threshold for outflows; p_r) relaxation rate; p_{ri}) run-up loss; p_{mt}) threshold for erosion; p_{pef}) erosion factor. In Table 2, "optimal" values obtained through calibration are listed.

In all the performed experiments, mud and lava flows were simulated one at a time, results being only later superimposed in the GIS. In other experiments (here not described), flows were allowed to start together, in order to appreciate their mutual interference.

Table 2. Parameters and optimal values for the models *fv* (on the left), after Crisci et al. [in press], and *S3-hex* (on the right), after Iovine et al. [2003a, b]. In the last column, double values refer to the Pizzo d'Alvano (left) and Vallicelle (right) study areas.

parameter	*measure*	*opt. value*
p_a	m	5
p_s	s	155.29
p_{Tv}	K	1373
p_{Ts}	K	1165.35
p_{adv}	m	0.7
p_{ads}	m	12
p_{cool}	$m \cdot t^{-1} \cdot K^{-3}$	$2.9 \cdot 10^{-14}$

parameter	*measure*	*opt. value*	
p_a	m	1.25	
p_{adh}	m	0.001	
p_f	m	0.1	
p_r	-	1	
p_{rl}	m	0.6	1.5
p_{mt}	m^2	3.5	2.0
p_{pef}	-	0.015	0.065

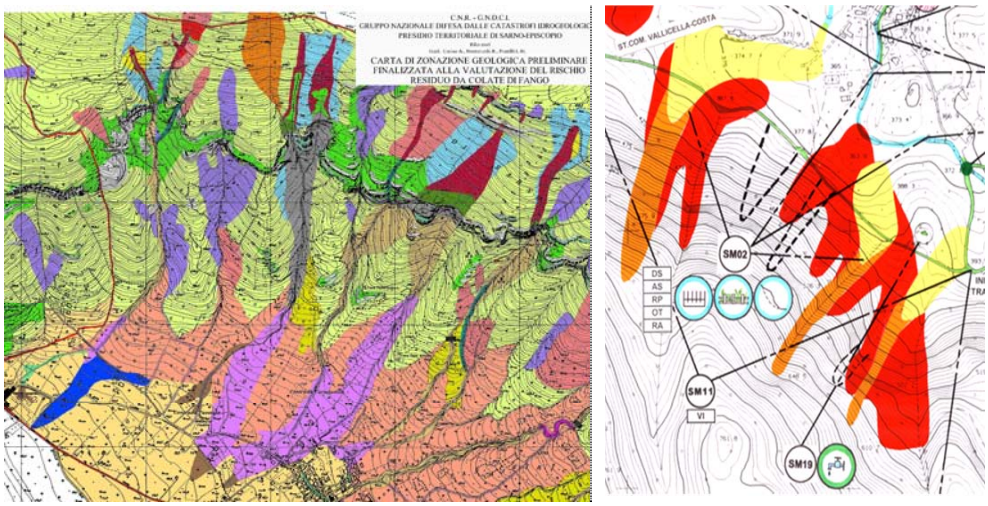


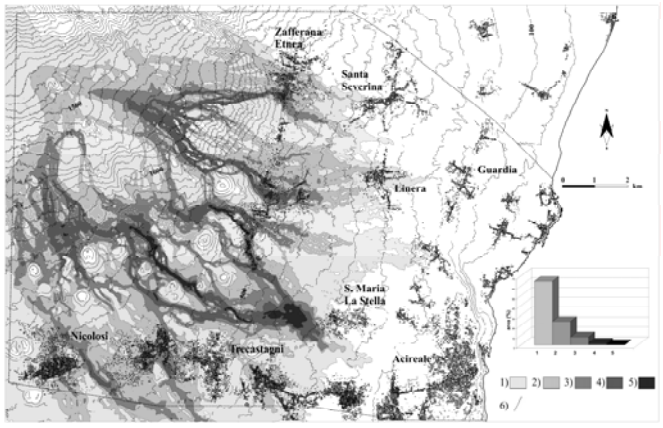
Figure 6. a) On the left, extract of the "Mudflow residual risk map" of the southern slope of Mt. Pizzo d'Alvano. Highest probabilities of soil-slip activation are shown in red tones, orange, dark-violet and in light-blue. Bare rock outcrops are in green, while zones exposed to mud-flow deposition are in light violet and dark-pink - after Caruso et al. [2005]. b) On the right, extract of the "Plan for Removal of the Highest Risk Conditions" of the northern slope of Mt. Pizzone. Highest landslide risk levels are in red, while mud-flows that occurred in December 1999 are either in orange (source-track) or in yellow (deposition zone) - after Basin Authority of the rivers Liri, Garigliano and Volturno [2000].



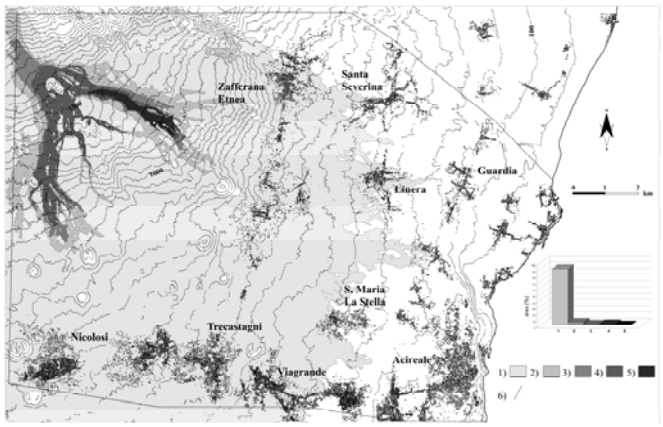
Figure 7. Extracts of the "Hydro-geological Setting Plan" (PAI) for Scilla (left), Favazzina (middle), and Bagnara Calabra (right). Key: red or black polygons delimit recent/active or dormant/inactive landslides, respectively; dashed polygons delimit landslide-zones; blue lines follow main drainages. Landslide types are distinguished by means of symbols inside (V: slide, U: flow) – after Calabrian Basin Authority [2001].

4. RESULTS

Examples of lava-flow susceptibility and hazard map for the SE flank of Mt. Etna (based on the cited sub-sample of n.1056 simulations) are shown in Figure 8. In the susceptibility map, the main drainages appear mostly exposed to invasion from lava flows, which may originate within the assumed "source zone" (cf. Figures 4 and 5) - a rather similar result may be obtained by merely evidencing the drainages on a contour map. Most of the poorly-exposed area, between Zafferana and Trecastagni, is presently diffusely urbanized. In contrast, in the hazard map highest hazard values are limited to the NW portion of the study area – i.e. at highest elevation a.s.l., and closer to the right flank of the DSGD.



In the susceptibility map (on the left), sites not affected by any simulation amount to ca. 41.5% of the considered area. At the other extreme, the maximum number of simulations affecting a given site is 952. The histogram shows area percentages with respect to the overall affected sites, per each class of relative susceptibility.



The hazard map (on the left) is based on the same set of simulations employed for elaborating the susceptibility map shown above. The histogram shows the percentages, with respect to the overall affected area, per each class of relative hazard (values: 90.3, 5.2, 1.7, 2.4, and 0.5, respectively).

Figure 8. SE flank of Mt. Etna. Lava-flow susceptibility map (a, on top), and hazard map (b, on bottom). Key: 1-5) classes of relative susceptibility/hazard (see text); 6) boundary of the study area - after Crisci et al. [in press], mod.

Figure 9 shows examples of mud-flow susceptibility zonation for the Vallicelle and the southern slope of Pizzo d'Alvano Massif. The zones affected by one/more simulated phenomena are distinguished by means of different colours. Depending also on location of selected sources, some of the drainages appear to be mostly exposed to mud-flow events. At the base of the massifs, the sectors characterized by the worst conditions coincide fairly well with those severely stricken during the recent 1998 and 1999 disasters.

Figure 10 illustrates examples of mud-flow susceptibility zonation in proximity of Bagnara Calabra, Favazzina and Scilla, along the Tyrrhenian coast of Calabria [La Manna, in prep.]. Also in these maps, zones affected by a different number of simulated events are shown by means of red-tones. In this case, the sectors which appear to be most exposed to flow-invasion are mainly located along the main drainages and the coastal plain. Many of the simulated cases directly affect the urbanized areas as well as threaten the existing infrastructures - as actually occurred in several of the historical events.

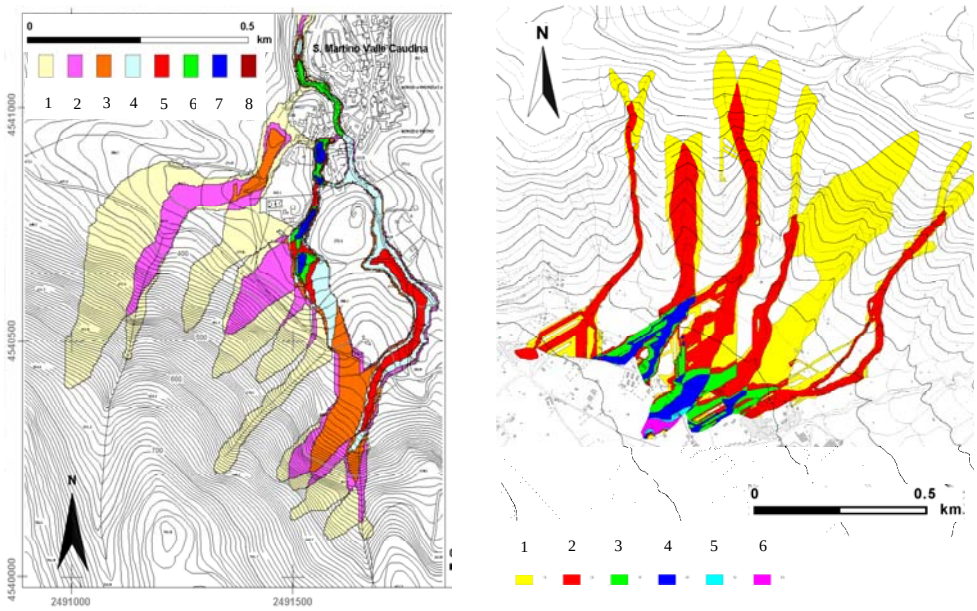


Figure 9. Preliminary mud-flow susceptibility maps for the study areas of: a) Vallicelle area (on left), and b) southern slope of Pizzo d'Alvano Massif (on right). The number of simulated cases which affect a given sector is graphically shown with different colours (ranging from 1 to either 6 or 8, according to the legends on bottom of the maps).

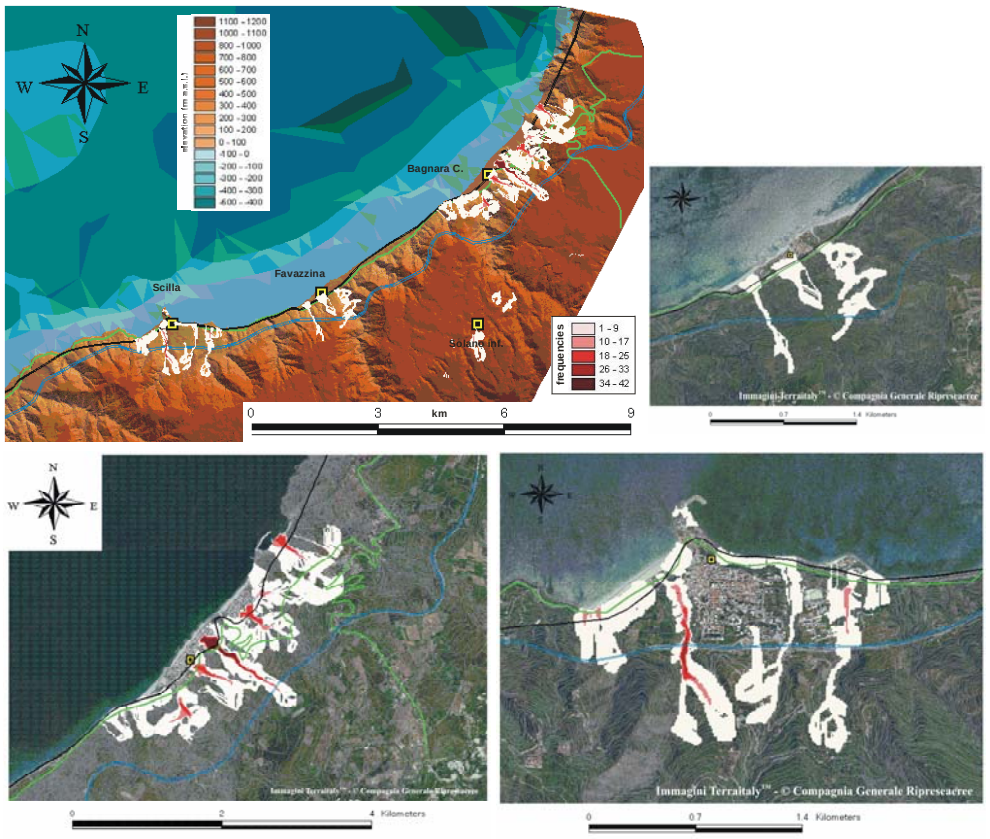


Figure 10. Preliminary mud-flow susceptibility maps for the Calabrian study area (only the sectors in proximity of the villages of Bagnara Calabra, Favazzina and Scilla were analysed). On the top-left, overall view of the study area; elevation ranges are shown by brown-blue tones. Blow-ups: Bagnara Calabra (bottom, left), Favazzina (top-right) and Scilla (bottom right). In all the maps, the number of events (frequency) affecting each zone is shown by means of red-tones (cf. legend in the overall map). Black, blue and green lines mark the railway, the highway, and the state road SS.18, respectively.

5. CONCLUSIONS

The methodological approach for mud-flow and lava-flow susceptibility/hazard mapping, briefly illustrated by examples in the present paper, is based on CA-modelling and GIS-analyses, combined with geological *l.s.* knowledge and historical information in order to empirically assign probabilities to the events of each considered scenario. Source locations and characteristics of the simulated events (e.g. source size for soil slips, effusion functions for vents) were in fact selected also on the basis of statistics of past events that occurred in the same study areas. By storing the performed simulation into a GIS environment and by successively overlapping and weighting the simulations (based on criteria which may be refined in time), the zonation of the study areas into sectors variously exposed to the considered dangers resulted to be a quite fast and effective task.
The CA-modelling approach surmounted several difficulties, which commonly hamper classic numerical methods. Its inherent limits, stemming from the necessity of a "discrete" description of the phenomena, are indeed well balanced by other appealing strong points. Among the greatest advantages of CA: they have the great capability to derive high performance from parallel computers: they are quite robust and do not generally show instability problems; they can fully incorporate even very-dense DEMs.
In the examples described, two simplified releases of CA-models were employed. A number of real cases – that occurred in the same study areas or within similar settings – were employed for calibration and validation (greatly facilitated by use of GA). Incidentally, a strongly innovative derivation of the model for mudflows (FLOW-S*) is presently undergoing final tests and calibration against real cases (data from both field real cases, and flume experiments performed at the CNR-IRPI laboratory). This new release is by far more physically-based with respect to previous ones. It is inspired by the well-known "equivalent fluid" approach, adapted to the discrete space-time CA viewpoint, and is characterized by a subset of physical parameters, which can be tuned depending on the material actually involved in the flow [cf. Iovine & Mangraviti, 2008]. Similarly, a more "physical" version of the model for lava flows is being derived, aiming at better capturing some specific behaviours of the considered real cases (e.g. tunnelling). Finally, thanks to parallel GA, libraries of parameters for both the models are presently being implemented by considering a set of "representative" types of study cases. This should provide a faster implementation of the described approach even to other (but comparable) areas of interest.
The above-presented preliminary results of mud-flow and lava-flow susceptibility/hazard maps, although based only on a sub-set of the overall planned simulations, clearly depict the most hazardous sectors of the considered study areas. The same zones were the most severely affected by the recent disasters that occurred in the past ten years.

ACKNOWLEDGMENTS

The approach described has been fine-tuned in the past ten years within a framework of multidisciplinary cooperation among researchers of the Italian National Research Council and colleagues of different universities. Unpublished, preliminary results for the Calabrian study area, anticipated herein, are being developed by Stefano La Manna for his degree thesis at the University of Calabria (under the direction of G. Iovine, V. Lupiano and R. Rongo). All the GIS elaborations were performed by V. Lupiano. Extracts of orthophotos of the Calabrian coast belong to the 1998 flight "Immagini Terraitaly™", courtesy of Compagnia Generale Ripreseaeree S.p.A., Parma (www.terraitaly.it).

REFERENCES

Acocella, V., and M. Neri, What makes flank eruptions? The 2001 Etna eruption and its possible triggering mechanisms. *Bulletin of Volcanology*, *65*, 517-529, 2003.

Basin Authority of the rivers Liri, Garigliano and Volturno, Plan for Removal of the Highest Risk Conditions – Map of the areas exposed to highest landslide risk levels, Tav.4.2 in scale 1/5,000 (*in Italian*), Technical Report, Genio Civile di Avellino, 2000.

Behncke, B., M. Neri, and A. Nagay, Lava flow hazard at Mount Etna (Italy): New data from a GIS-based study. In: Manga, M., Ventura, G. (Eds.), Kinematics and dynamics of lava flows. *Geological Society of America Special Paper*, *396*, 189-208, 2005.

Borgia, A., L. Ferrari, and G. Pasquarè, Importance of gravitational spreading in the tectonic and volcanic evolution of Mount Etna. *Nature*, *357*, 231-235, 1992.

Calabrian Basin Authority, Hydro-geological Setting Plan (PAI) (*in Italian*), Regione Calabria, Councillorship for Public Works, Catanzaro, 2001 - available on line at http://www.autoritadibacinocalabria.it/PAI/Home/html/Pai_home.htm

Caruso, A., R. Monteverde, and L.M. Puzzilli, Preliminary geological zonation map for the evaluation of the residual mud-flow risk (in scale 1,5000). CNR-GNDCI, Presidio Territoriale di Sarno-Episcopio (*in Italian*), Technical Report. In: Cascini, L., D. Guida, and G. Sorbino, The territorial presidium, a field experience, Chapter 6, Rubettino, Soveria Mannelli, *GNDCI publ. 2857*(B), 2005 - available on-line at http://www.commissario2994.it/originali/Zonazione_Episcopio.jpg

Corazzato, C., and A. Tibaldi, Fracture control on type, morphology and distribution of parasitic volcanic cones: an example from Mt. Etna, Italy. *Journal of Volcanology and Geothermal Research*, *158*, 177-194, 2006.

Crisci, G.M., S. Di Gregorio, and G.A. Ranieri, A cellular space model of basaltic lava flow. In: Mesnard, G. (Ed.), Proc. Int. AMSE Conf. on Modelling and Simulation, *11*, 65-67, AMSE, Paris, 1982.

Crisci, G.M., R. Rongo, S. Di Gregorio, and W. Spataro, The simulation model SCIARA: the 1991 and 2001 lava flows at Mount Etna, *Journal of Volcanology and Geothermal Research*, *132*, 253–267, 2004.

Crisci, G.M., G. Iovine, S. Di Gregorio, and V. Lupiano, Lava-flow hazard on the SE flank of Mt. Etna (Southern Italy), *Journal of Volcanology and Geothermal Research*, in press.

D'Ambrosio, D., S. Di Gregorio, S. Gabriele, and R. Gaudio, A Cellular Automata model for soil erosion by water, *Physics and Chemistry of the Earth*, *26*(B), 33-40, 2001.

D'Ambrosio, D., S. Di Gregorio, and G. Iovine, Simulating debris flows through a hexagonal Cellular Automata model: Sciddica $S_{3\text{-}hex}$. *Natural Hazards and Earth System Sciences*, *3*, 545–559, 2003a.

D'Ambrosio, D., S. Di Gregorio, G. Iovine, V. Lupiano, R. Rongo, and W. Spataro, First simulations of the Sarno debris flows through Cellular Automata modelling, *Geomorphology*, *54*, 91-117, 2003b.

D'Ambrosio, D., Rongo, R., Spataro, W., Avolio, M.V., Lupiano, V., Lava invasion susceptibility hazard mapping through cellular automata. In: El Yacoubi, S., B. Chopard, and S. Bandini (Eds.), ACRI 2006, LNCS *4173*, 452-461, 2006.

D'Ambrosio, D., G. Iovine, W. Spataro, and H. Miyamoto, A macroscopic collisional model for debris-flows simulation, *Environmental Modelling and Software*, *22*(10), 1417-1436, 2007.

Del Prete, M., F.M. Guadagno, and A.B. Hawkins, Preliminary report on the landslides of 5 May 1998, Campania, southern Italy, *Bulletin of Engineering Geology and Environment*, *57*, 113–129, 1998.

Di Gregorio, S., and R. Serra, An empirical method for modelling and simulating some complex macroscopic phenomena by cellular automata, *Future Generation Computer Systems*, *16*, 259–271, 1999.

Hungr, O., A model for the runout analysis of rapid flow slides, debris flows, and avalanches. *Canadian Geotechnical Journal*, *32*, 610-623, 1995.

Iovine G., S. Di Gregorio, D. D'Ambrosio, and V. Lupiano, Debris Flows and Cellular Automata: an example of simulation from the 1998 disaster of Sarno (Italy). In: Delahaye D., F. Levoy, and O. Maquaire (Eds.), Geomorphology: from expert opinion to modelling. A tribute to Professor Jean-Claude Flageollet. Proc. Int. Symp. "Geomorphology: from expert opinion to modelling", Strasbourg, 26-27 April 2002, SODIMPAL, Rouen, 55-64, 2002.

Iovine, G., and S. Di Gregorio, Simulating mud flows by means of Cellular Automata: considerations on the model SCIDDICA (release S2) (*in Italian*). *Bollettino della Società Geologica Italiana*, *122*, 63-84, 2003.

Iovine, G., S. Di Gregorio, and V. Lupiano, Assessing debris-flow susceptibility through cellular automata modelling: an example from the May 1998 disaster at Pizzo d'Alvano (Campania, Southern Italy). In: Rickenmann, D., Chen, C.L. (Eds.), Proc. 3rd Int. Conf.

on Debris-Flow Hazards Mitigation: Mechanics, Prediction, and Assessment, Davos, Switzerland, 1. Millpress, Rotterdam, 623–634, 2003a.

Iovine, G., S. Di Gregorio, and V. Lupiano, Debris-flow susceptibility assessment through cellular automata modeling: an example from the 15-16 December 1999 disaster at Cervinara and San Martino Valle Caudina (Campania, southern Italy). *Natural Hazards and Earth System Sciences 3*, 457–468, 2003b.

Iovine, G., S. Di Gregorio, D. D'Ambrosio, V. Lupiano, L. Merenda and G. Nardi, Evaluating and mapping mud-flow susceptibility by applying a Cellular Automata model (*in Italian*). In: Proc. "Mudflow risk mitigation at Sarno and other villages affected by the May 1998 disaster". Napoli, 2-3 May 2005, 175-185, 2007.

Iovine G., and P. Mangraviti, The model FLOW-S* for flow-type landslides. Preliminary considerations on model performances. Proc. AOGS 5th Annual General Meeting, 16-20 June, Busan-Korea, Abstract Volume in CD-ROM, 2008.

Iverson, R.M., The physics of debris flows, *Reviews in Geophysics*, *35*, 245-296, 1997.

Iverson, R.M., S.P. Schilling, and J.W. Vallance, Objective delineation of lahar-inundation hazard zones, *Geological Society of America Bulletin*, *110*(8), 972-984, 1998.

La Manna, S., Mudflow susceptibility mapping through CA-modelling and GIS-techniques along the southern Tyrrhenian coast of Calabria – The sector between Scilla and Bagnara Calabra. Degree thesis at the University of Calabria, in prep.

McBirney, A.R., and T. Murase, Rheological properties of magmas, *Annual Review of Earth Planetary Sciences*, *12*, 337–357, 1984.

Miyamoto, H., and S. Sasaki, Simulating lava flows by an improved cellular automata method, *Computers & Geosciences*, *23*, 283-292, 1997.

Murray, A.B., and C. Paola, A cellular model of braided rivers, *Nature*, *371*, 54-57, 1994.

Pitman, E.B., A. Patra, A. Bauer, M.F. Sheridan, and M. Bursik, Modeling and Computing Geophysical Mass Flows. In: Hou, T., and E. Tadmor (Eds.), Hyperbolic Problems: Theory, Numerics, Applications, Springer, New York, 807-818, 2003.

Romano, R., and C. Sturiale, The historical eruptions of Mt. Etna (volcanological data). *Memorie della Società Geologica Italiana*, *23*, 75-97, 1982.

Sorriso-Valvo, G.M., and C. Tansi, Large-scale landslides and deep-seated gravitational slope deformations in Calabria (*in Italian*). Geografia Fisica e Dinamica Quaternaria, *19*, 395-408, 1996.

Succi, S., The Lattice Boltzmann Equation for Fluid Dynamics and Beyond. Oxford University Press, pp.368, Oxford USA, 2004.

Tortorici L., C. Monaco, C. Tansi, and O. Cocina, Recent and active tectonics in the Calabrian arc (Southern Italy). *Tectonophysics*, *243*, 37-55, 1995.

United Nations, Mudflows. Experience and Lessons Learned from the Management of Major Disasters. Department of Humanitarian Affairs, U.N., Geneva, 1997.

Vittori, E., F. Fumanti, D. Ligato, and A. Triglia, Technical report on the surveying in the area affected by the hydrogeological crisis of 15–16 December 1999 in the Caudina Valley (Avellino) (*in Italian*), ANPA and ARPA Campania, 14 January, 2000.

APPENDIX

A) The "minimisation algorithm of differences" [after Di Gregorio and Serra, 1999]

At step t, two quantities are identified in the central cell: the fixed part $q(0)$ and the mobile part p of its height. The mobile part can be distributed to the adjacent cells (e.g. thickness of landslide debris or lava flow); the fixed part cannot change during the simulation (e.g. elevation related to the unerodable bedrock). The height of the central cell is given by the sum of two terms $p+q(0)$. The height of the i^{-th} adjacent cell of the neighbourhood is: $q(i)$, $1 \le i \le m-1$. At step $t+1$, the outflow from the central cell to the i^{-th} neighbouring cell is denoted by $f(i)$, $0 \le i \le m$, where $f(0)$ is the part of p which is not distributed. Let $q_0(i)=q(i)+f(i)$, $0 \le i \le m-1$ be the sum of the content of a neighbouring cell (at step t), plus the flow coming from the central cell, and let $q_{0\min}$ be the minimum value for $q_0(i)$. The determination of the outflows from the central cell to the adjacent cells is based on the local minimisation of the differences in "height", as described by the following expression:

$$\sum_{i=0}^{m-1}(q'(i) - q'_{\min}) \qquad (1)$$

B) The model S_{3-hex} [after D'Ambrosio et al., 2003]

Formally, it is: $S_{3-hex} = \langle R, X, Q, P, \sigma \rangle$, where

- $R=\{(x,y)\in Z^2 \mid -l_x \le x \le l_x, -l_y \le y \le l_y\}$ is the cellular space, where Z is the set of integers.
- X={central, NNW, NNE, E, SSE, SSW, W} is the neighbourhood, given by the "central" cell and its six adjacent cells.
- $Q=Q_a \times Q_{th} \times Q_e \times Q_d \times Q^6_o$ is the finite set of states of the *finite automaton*, given by the Cartesian product of the sets of the considered substates: Q_a = average cell elevation a.s.l. (bedrock + regolith); Q_{th} = thickness of landslide debris; Q_e = energy of landslide debris; Q_d = depth of erodable regolith; Q_o = debris outflow from the central cell toward any cell of the neighbourhood. The value of the substate x in the cell is expressed by $q_x \in Q_x$.
- $P=\{p_a, p_t, p_{adh}, p_f, p_r, p_{rl}, p_{mt}, p_{er}\}$ is the set of the model parameters (cf. Table 2).
- $\sigma=Q^7 \to Q$ is the transition function, constituted by the elementary processes $\sigma_1, \sigma_2, \sigma_3, \sigma_4$.
- σ_1: $Q_a \times Q_e \times Q_{th} \times Q_d \times p_{er} \times p_{mt} \to Q_a \times Q_e \times Q_{th} \times Q_d$ is the process of "entrainment". The erosion condition is: $q_e(0) > p_{mt}$. If $\Delta_d < q_d(0)$, the eroded quantity of regolith is $\Delta_d = (q_e(0) - p_{mt}) p_{er}$, else $\Delta_d = q_d(0)$.
- σ_2: $Q^7_a \times Q^7_{th} \times Q_e \times p_{adh} \times p_r \times p_f \to Q^6_o$ is the process of "debris outflows".
- σ_3: $(Q_{th} \times Q_e \times Q^6_o)^7 \to Q_{th} \times Q_e$ is the process of "update" of debris thickness (q_{th}) and energy (q_e).
- σ_4: $Q_e \times Q_{th} \times p_{rl} \to Q_e$ is the process of "energy loss and run-up". It is modelled, by reducing the run-up (r) to a value – not lower than $q_{th}(0)$ – by the parameter p_{rl}. If $[k \cdot q_e(0)/q_{th}(0) - p_{rl}] > q_{th}(0)$, the loss of run-up is $\Delta_r = p_{rl}$, else $\Delta_r = k \cdot q_e(0)/q_{th}(0) - q_{th}(0)$, where $k=2/\rho g A$, ρ=density of the material, g=gravitational acceleration, A=area of the cell.

C) The model *fv* [after Crisci et al., 2004]

Formally, it is: $fv = \langle R, L, X, Q, P, \sigma, \Gamma \rangle$, where

- R and X = see above.
- $L \subset R$=specifies the vents.
- $Q=Q_a \times Q_t \times Q_T \times Q_f^6$ is the finite set of states of the *finite automaton*, given by the Cartesian product of the sets of the considered substates: Q_a is the average cell elevation a.s.l.; Q_t is the thickness of lava in the cell; Q_T is the average lava temperature in the cell; Q_f^6 is the lava outflow from the central cell toward any cell of the neighbourhood.
- $P = \{ p_a, p_s, p_{Tv}, p_{Ts}, p_{adv}, p_{ads}, p_{cool} \}$ is the set of the model parameters (cf. Table 2).
- $\sigma = Q^7 \to Q$ is the transition function, constituted by the elementary processes $\sigma_1, \sigma_2, \sigma_3$.
- σ_1: is the process of "lava outflows". The minimisation algorithm of differences (see above) is applied to the amount of lava exceeding the adherence (v), i.e. the aliquot of lava which remains in the cell, due to viscosity. It is:

$$\nu = k_1 e^{-k_2 T} \tag{2}$$

in which $T \in Q_T$ is the lava temperature; k_1 and k_2 depend on lava rheological properties and are computed as follows:

$$\begin{cases} p_{adv} = k_1 e^{-k_2 p_{Tv}} \\ p_{ads} = k_1 e^{-k_2 p_{Ts}} \end{cases} \tag{3}$$

- σ_2: is the process of "variation of lava temperature". It is computed by means of a two-step procedure, as temperature can change due either to a) mixing of masses at different temperatures, or b) surface radiation. Accordingly, in the first step:

$$T_{av} = \left(t_r \times T(0) + \sum_{i=1}^{6} f(i,0) \times T(i) \right) \Big/ \left(t_r + \sum_{i=1}^{6} f(i,0) \right) \tag{4}$$

is the weighted average temperature of the residual lava inside the cell plus the lava inflows from neighbouring cells, where $t_r \in Q_t$ is the residual lava thickness inside the central cell (after outflows distribution), $T \in Q_T$ is the lava temperature, and $f(i,0)$ is the lava inflow from the i^{th} neighbouring cell (as computed by the minimisation algorithm).
In the second step, the loss of thermal energy due to surface radiation is computed as:

$$T = T_{av} \Big/ \sqrt[3]{1 + \left(T_{av}^3 C \Delta t A / V \right)} \tag{5}$$

where $C=p_{cool}$ depends on lava rheology, $\Delta t = p_s$, A is the exposed area of the lava, and V the lava volume.

- σ_3: is the process of "lava halting". If the temperature of the lava inside the cell drops below a given threshold, p_{Ts}, the mass of lava comes to a halt, due to an increase of viscosity, thus becoming part of the topography. The elevation of the cell is increased by an amount equal to lava thickness, and lava thickness is set to zero.
- $\Gamma = \{ \gamma_1, \gamma_2, \ldots \gamma_i \ldots \gamma_v \}$ is the set of effusion-rate functions assigned to the vents, where v is the number of considered vents, and $\gamma_i : Q_t \times \mathbf{N} \to Q_t$ is the effusion-rate function (in time), which specifies the amount of lava emitted from the i^{-th} vent during the simulation (expressed in terms of thickness, in meters) at each step $s \in \mathbf{N}$, with $\mathbf{N}$ = set of natural numbers.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Early warning of fast landslides triggering based on instrumented slope data analysis

Emilia Damiano[a], Roberto Greco[a,b], Andrea Guida[a] and Lucio Olivares[a,b]

[a] Dipartimento di Ingegneria Civile, Seconda Università di Napoli, via Roma 29 – 81031 Aversa (CE), Italy (roberto.greco@unina2.it)

[b] CIRIAM – Centro Interdipartimentale di Ricerca in Ingegneria Ambientale, Seconda Università di Napoli, via Roma 29 – 81031 Aversa (CE), Italy

Abstract: The results of laboratory infiltration experiments on instrumented small scale model slopes are presented. Loose granular pyroclastic soils from the mountainous area north-eastern of Napoli have been examined. The experiments aimed to reproduce the simple case of a homogeneous indefinite slope, for a better understanding of the hydraulic processes leading to slope failure. With this respect, useful information was provided by the results of coupled measurements of soil suction and volumetric water content, respectively carried out by minitensiometers and Time Domain Reflectometry. The obtained results indeed showed how soil wetting during rainfall infiltration took place under matric suction values around 5kPa smaller than what could be predicted by equilibrium water retention curves estimated in laboratory. Since steep slopes equilibrium is often guaranteed by cohesion increment due to suction under unsaturated conditions, the obtained result indicates that effective early warning of fast landslides should rely on direct monitoring of soil suction and water content in the field.

Keywords: Flowslides; Pyroclastic granular soils; Physical modelling; TDR; Early warning.

1. INTRODUCTION

Occurrence of fast landslides has become more and more dangerous during the last decades, due to the increased density of settlements, industrial plants and infrastructures. Such problem is particularly worrying in Campania (Southern Italy), where the fast population growth led a diffuse building activity without planning: indeed, recent flowslides caused hundreds of victims and heavy damages to buildings, roads and other infrastructures.

Large mountainous areas in Campania are mantled by loose pyroclastic granular soils up to a depth of a few meters from top soil surface. These soils have usually a grain size that falls in the domain of silty sands, including pumice interbeds (gravelly sands), with saturated hydraulic conductivities up to the order of 10^{-1}cm/min. Such deposits often cover steep slopes, which stability is guaranteed by the apparent cohesion due to suction under unsaturated conditions, that are the most common conditions for these slopes [Olivares and Picarelli, 2003]. Whereas rainfall infiltration causes soil to approach saturation, suction vanishes and slope failure may occur.

Besides soil physical properties, landslide triggering is influenced by several factors, such as rainfall intensity, soil initial moisture and suction, slope inclination, boundary conditions. Whereas slope failure occurs with soil close to being saturated, landslide may develop in form of fast and destructive flowslide.

Calibration of reliable mathematical models of such a complex phenomenon requires availability of experimental observations of the major variables of interest, such as soil

moisture and suction, soil deformation and displacements, pore water pressure, during the entire process of infiltration until slope failure. Due to the sudden trigger and extremely rapid propagation of such type of landslides, such data sets are rarely available for natural slopes where flowslides occurred. As a consequence landslide risk assessment and early warning in Campania rely on simple empirical models [Sirangelo et al., 2003] based on correlation between some features of rainfall records (cumulated height, duration, season etc.) and the correspondent observed landslides. Laboratory experiments on instrumented small scale slope models represent an effective way to provide such data sets [Olivares et al., 2008] useful for building up more complex models of landslide triggering prediction.

In this paper, a series of laboratory tests carried out on model slopes made by a single layer of granular pyroclastic soil taken in the mountainous area north-eastern of Napoli, are presented. Among the responses of the various sensors installed, coupled measurements of water content profiles by TDR and matric suction by tensiometers are presented. The obtained results show how instrumented model slope infiltration tests allow to estimate effective water retention curves experienced by the soil during infiltration.

2. MATERIALS AND METHODS

At the Geotechnical Laboratory of CIRIAM an instrumented flume to investigate on the mechanics of rainfall-induced landslides in initially unsaturated deposits of granular soils is available [Olivares et al. 2003a]. In the flume, a homogeneous model slope is reconstituted by a moist-tamping technique: with the flume in horizontal position, the soil is lain in thin layers having a thickness around 0.5cm in order to get a uniform distribution of porosity. After 48 hours, before inclining the flume at the desired angle, soil is gently wet with an artificial rainfall of weak intensity for few minutes, in order to prevent detachment of small aggregates from soil surface. All the tests have been carried out with uniform and constant rainfall until slope failure. Rain intensity, variable between 10mm/h and 110mm/h, is reproduced by an artificial system, through several spray nozzles installed above the lateral sides of the flume. In order to prevent erosion of the soil surface, the adopted nozzles produce nebulized water particles 0.1mm in diameter. During the infiltration process, slope behaviour is observed. In particular, matric suction at different positions and depths is measured by insertion of mini-tensiometers within the slope, while positive pore pressure is monitored by miniature transducers installed at the base of the model. Soil water content profiles are measured by Time Domain Reflectometry (TDR) metallic probes [Topp et al.,1980]. At the ground surface, the displacement field is observed through laser sensors, installed in different zones of the slope with the optical axis perpendicular to the soil surface in order to monitor settlements, and high definition video-cameras to obtain, through a Particle Image Velocimetry technique (PIV), the overall horizontal displacement field. Rainfall intensity and air temperature are measured respectively by a rain gauge located at the foot of the slope and a standard thermocouple. A sketch of the flume is given in figure 1; more details about all the devices installed are reported in Olivares et al. [2008].

For the presented applications in granular soils conventional jet-fill type tensiometers (Soil-Moisture Equipment), equipped with a porous ceramic cup with a nominal bubble pressure of 100kPa and a standard Bourdon-type vacuum gauge, have been used. The range of operation of the unit is 0-85kPa, sufficient for laboratory testing in sandy soils, with a sensitivity of about 1kPa. The response of the devices has been experimentally tested, resulting fast enough to follow the expected changes in suction during infiltration in the model slope [Olivares et al., 2008]. Prior to installation, all tensiometers are filled and tested with de-aired water to ensure that the ceramic tips are free from cracks and all connections are properly sealed.

Soil water content measurements have been carried out with a recently proposed method allowing to retrieve moisture distribution along a TDR metallic probe buried into the soil [Greco, 2006; Greco and Guida, 2008]. Soil moisture measurements by TDR rely on the relationship between bulk soil dielectric permittivity, ε_r, and its volumetric water content, θ

[Topp et al., 1980]. TDR experimental apparatus consists of a Tektronix 1502C cable tester connected to a three rods metallic probe installed in the slope orthogonal to soil surface, as shown in figure 1.

In all the presented tests the infinite slope scheme (ratio thickness/length smaller than 1/10) has been investigated, since it represents the most frequent case in the concerned area. The experiments have been conducted on air-fall unsaturated cohesionless pyroclastic soils involved in catastrophic flowslides, taken at the experimental fields of Cervinara [Olivares et al., 2003b] and Monteforte Irpino [Nicotera and Papa, 2007], North-East of Napoli. Figure 2 reports grain size curves of the investigated materials: in all cases, the soil is a silty sand with a high sandy component but also a significant amount of non-plastic silt.

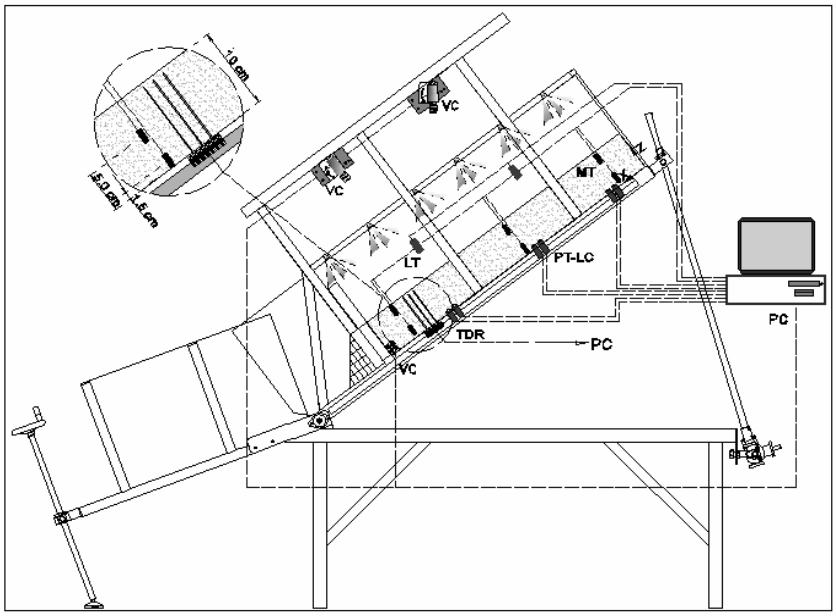


Figure 1. Sketch of the instrumented flume: LT = laser transducer; MT = minitensiometer; PT = pressure transducer; LC = load cell; TDR = TDR probe; VC = digital video-camera.

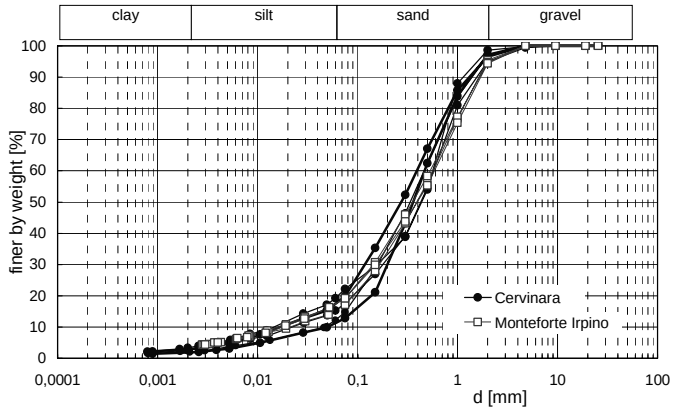


Figure 2. Grain size distributions of volcanic ashes.

Mechanical and hydraulic characterization has been carried out on undisturbed specimens recovered from pits dug along the slopes: Olivares and Picarelli [2003] and Nicotera and Papa [2007] respectively investigated the main soil properties of Cervinara and Monteforte Irpino ashes, which are summarized in Table 1. The volcanic ashes show quite similar properties: the specific weight is about 25kN/m^3; the porosity is quite high, ranging between 68% and 74%; the saturation degree, depending on environmental conditions, varies within a wide range but is generally much less than one; the saturated hydraulic conductivity of natural samples, measured by constant head tests, ranges between 1.0E-07 and 7.0E-07m/s; the saturated shear strength, measured through both CID and CIU triaxial

tests on saturated natural samples, is characterised by a friction angle of about 37° and a nil value of cohesion.

Since the unsaturated condition is the most frequent condition for the examined soils, their unsaturated properties have been investigated too. The water retention curves, shown in figure 3, have been evaluated through conventional laboratory tests: evaporation tests (using minitensiometer to explore suction between 0 and 80kPa and pressure plate for higher values of suction) on Monteforte Irpino ash [Nicotera and Papa, 2007]; evaporation and infiltration tests by use of minitensiometer and tests in suction controlled triaxial apparatus on natural and remoulded specimens of Cervinara ash [Olivares et al., 2003b]. The experimental results have been fitted through the expression proposed by van Genuchten [1980]: the correspondent values of parameters are reported in Table 1. A comparison between the retention curves of both soils seems to highlight that, in the range of suction of interest (0÷80kPa), the slope of the curve related to Cervinara ash is higher than the others obtained for Monteforte Irpino ash.

Table 1. Physical, mechanical and hydraulic properties of investigated soils.

Parameters	Cervinara	Monteforte Irpino
Specific weight (kN m^{-3})	25.4÷25.9	25.2
Porosity	0.71÷0.74	0.71
Degree of saturation	0.48÷0.71	0.71
Saturated hydraulic conductivity k_{sat} (m s^{-1})	$1 \cdot 10^{-7} \div 5 \cdot 10^{-7}$	$7 \cdot 10^{-7}$
Effective friction angle ϕ' (°)	38	37
Cohesion c' (kPa)	0	0
Residual volumetric water content θ_r	0.25	0.16
Saturated volumetric water content θ_{sat}	0.69	0.66
α (m^{-1}) van Genuchten	1.8	1.2
n van Genuchten	7.1	1.5
m van Genuchten	0.15	0.33

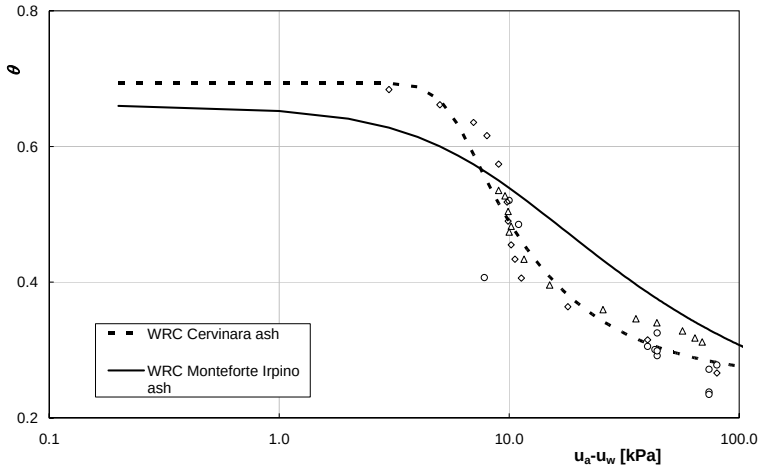


Figure 3. Water retention curves of volcanic ash from Monteforte Irpino [Nicotera and Papa, 2007] and Cervinara.

The unsaturated shear strength of Cervinara ashes has been measured by means of suction controlled triaxial tests using a mean net stress ($p-u_a$) between 20 and 70kPa and a suction (u_a-u_w) between 10 and 80kPa which are quite consistent with those existing in site [Olivares e Picarelli, 2003]. Such data show the significant role of suction, even for small values of (u_a-u_w): in particular, for a suction of 4-8kPa, the cohesion is 2-6kPa. Similar

results have been obtained by Nicotera and Papa [2007] for Monteforte Irpino ash.

For the application of TDR technique, the $\theta(\varepsilon_r)$ relationship has been experimentally determined for Cervinara soil over reconstituted samples. The obtained experimental points, plotted in figure 4, fall very close to $\theta(\varepsilon_r)$ relationships for low density volcanic soils available in literature [Tomer et al., 1999; Regalado et al., 2003], while application of the so-called 'universal calibration relationship' by Topp et al. [1980] would lead to soil water content underestimation.

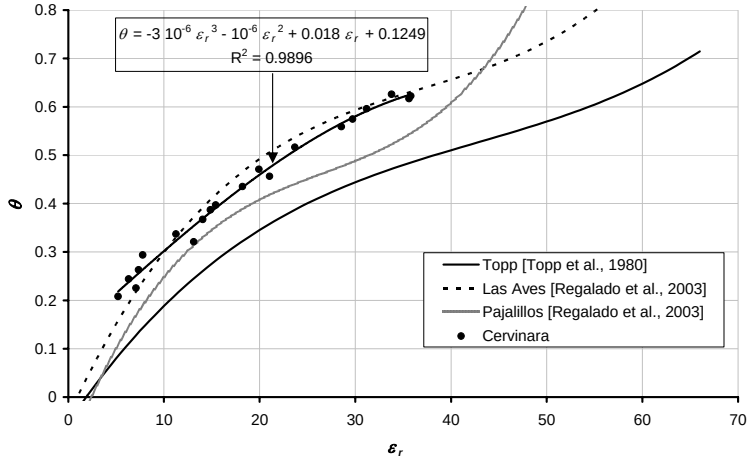


Figure 4. Dielectric permittivity vs. water content relationship for Cervinara ash compared with TDR 'universal calibration relationship' and specific relationship available in literature for volcanic ashes.

Table 2. Main characteristics of the tests

Test	D3	D4	D7
Investigated soil	Cervinara ash	Cervinara ash	Monteforte Irpino ash
Rainfall intensity (mm/h)	55	56	54
Duration of test (min)	36	30	95
Initial porosity n_0	0.75	0.75	0.69

The experiments have been carried out on 10cm thick, 50cm wide and 1.10m long slopes, with inclination of 40°. The bottom of the flume was impervious, while at the toe of the slope a geotextile drain was placed. A 9.5cm long TDR probe was placed at the bottom of the flume. Minitensiometers were installed at various positions within the slope. In particular two of them were placed nearby the TDR probe, respectively 1.5cm and 5cm above the bottom of the flume. The characteristics of each tests discussed in the following are given in Table 2.

3. RESULTS AND DISCUSSION

All the infiltration experiments lasted until slope failure was observed. Although the applied rainfall intensity was in all cases much higher than soil saturated hydraulic conductivity, no significant surface runoff was observed during the experiments, denoting that soil potential infiltration was never exceeded.

The response of the minitensiometers placed at various positions within the slope indicates that the infinite slope conditions were satisfactorily reproduced in all the tests: differences in suction values measured at the same depth inside the soil layer were always smaller than 1.5kPa. This result allows to consider water content distribution measured by the single TDR probe installed as representative of the entire slope behaviour.

Figure 5, in which z- axis is orthogonal to soil surface and directed from flume bottom (z=0) towards soil surface, shows some of the retrieved volumetric water content profiles at various times during experiment D3. The shape of the profiles is consistent with the expected infiltration process, while the time growth of cumulated infiltration height, h, estimated by integrating the retrieved profiles, is in good agreement with the applied artificial rainfall intensity, as it is showed in figure 6. In fact, the slope of h(t) relationship results 44.0mm/h, roughly equal to the projection of the applied rainfall intensity along the inclined slope, until 18min after the beginning of the test. The following slope increase was due to the reduction of soil layer thickness, witnessed by observed surface settlement, w, also given in figure 6.

The last retrieved profile shown in figure 5 was acquired few seconds before slope failure, when soil saturation was reached. The departure of saturated water content from soil initial porosity value increases with depth. This result indicates that wet soil experiences volumetric collapse due to overburden load.

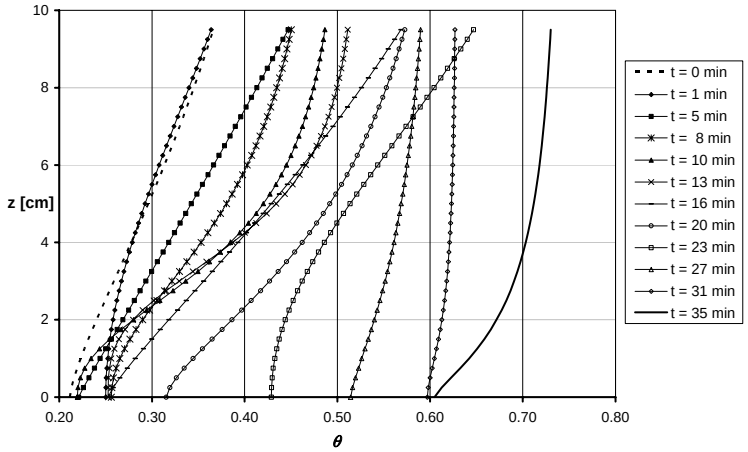


Figure 5. Volumetric water content profiles retrieved by TDR during experiment D3.

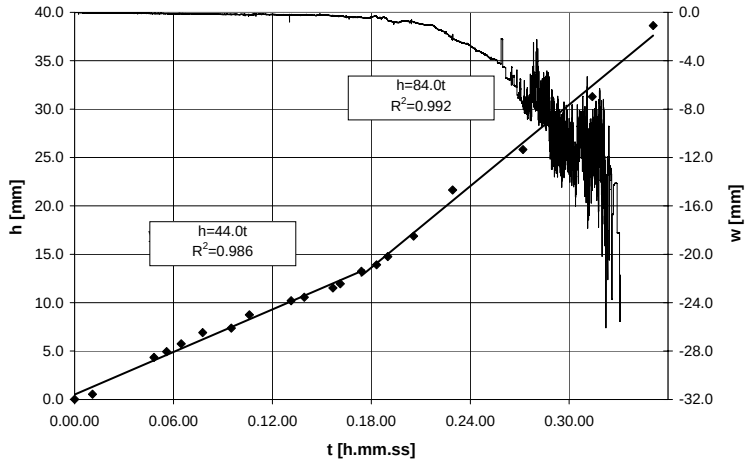


Figure 6. Cumulated infiltration height, h, and mean soil surface settlement, w, measured in experiment D3.

Figure 7 shows the time history of suction measured during experiments D4 and D7 by two minitensiometers buried at 1.5cm and 5.0cm above flume bottom, compared with the temporal evolution of volumetric water content at the same depths as given by water content profiles retrieved by TDR. It is worth to note that neither matric suction nor water content at 1.5cm changed before 30min after the start of experiment D7. A quite different response was observed at the same depth in experiment D4, during which the observed variables at the same depth started changing only 12min after the beginning of the test, although nearly the same artificial rainfall intensity was applied. Such result denotes that in

experiment D7 the component of water flow parallel to slope was much larger than during test D4.

The (θ, u_a-u_w) points obtained by coupling suction values measured by minitensiometers at 1.5cm and 5cm above flume bottom, with soil water contents read at the same depths along the profiles retrieved by TDR, are given in Figure 8. Although some dispersion is present, nearly all the experimental points fall below water retention curve estimated in laboratory, especially when soil saturation is approached. This result indicates that during the infiltration process reproduced in the model slope, soil experienced matric suction around 5kPa smaller than what could be expected from estimated water retention curve. As a consequence, soil suction contribution to slope stability resulted significantly smaller than expected. Since infiltration rate and boundary conditions applied to the slope during flume experiments resemble real slope conditions, it is expected that this result would be qualitatively confirmed also by field observations.

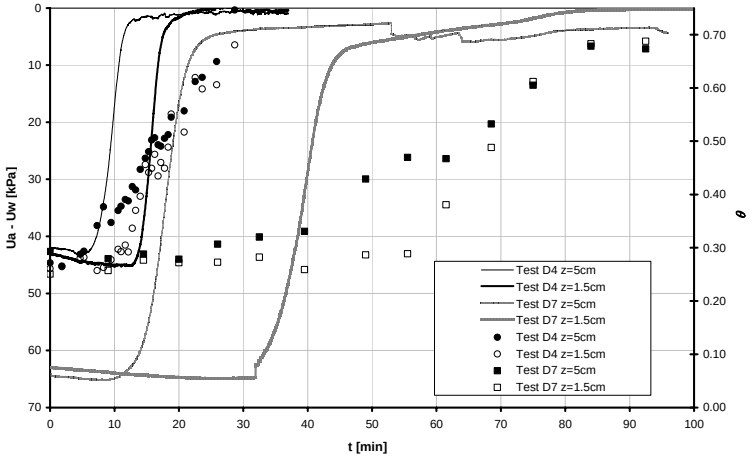


Figure 7. Matric suction (continuous and dashed lines) and volumetric water content (dots) time histories during experiments D4 and D7.

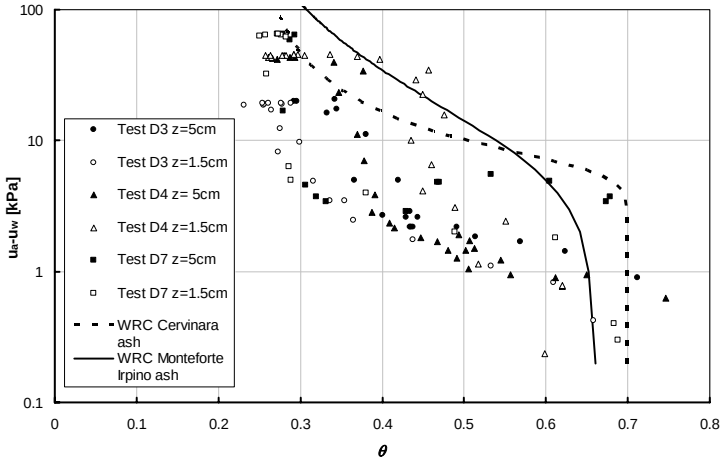


Figure 8. (θ, u_a-u_w) points measured during experiments D2, D4 and D7 compared with water retention curves estimated in laboratory for Cervinara and Monteforte Irpino ashes.

4. CONCLUSIONS

The results of infiltration tests in loose granular pyroclastic soil carried out on model slopes reproduced in an experimental flume in laboratory are presented. The measurement instruments installed in the flume allowed to monitor the entire phenomenon from the beginning of infiltration until slope failure. In particular, very useful information is

provided by the coupled measurements of soil matric suction and volumetric water content at various depths, acquired, respectively, with minitensiometers and a TDR metallic probe. The observed (θ, u_a-u_w) points, in fact, show that during an infiltration process in a steep slope, with inclination angle comparable with soil internal friction angle, soil wetting process takes place under smaller suction than what could be predicted by soil water retention curves estimated in laboratory. Slope equilibrium conditions may therefore be significantly affected by such behaviour, especially for the case of very steep slopes, with inclination larger than internal friction angle, where equilibrium is mainly guaranteed by cohesion component due to suction. The obtained results indicate that field measurements of soil water content by TDR can fruitfully supplement suction measurements for the implementation of effective early warning systems.

ACKNOWLEDGEMENTS

The research was co-financed by the Italian Ministry of University, by means of the PRIN 2006 PRIN program, within the research project entitled 'Definition of critical rainfall thresholds for destructive landslides for civil protection purposes'. The authors gratefully aknowledge prof. G Urciuoli of University of Naples 'Federico II' for providing soil samples and data of Monteforte Irpino.

REFERENCES

Greco, R., 2006, Soil water content inverse profiling from single TDR waveforms, *Journal Hydrol.*, 317, 325-339.

Greco, R. and Guida, A., 2008, Field measurements of topsoil moisture profiles by vertical TDR probes. *Journal Hydrol.*, 348, 442-451.

Nicotera, M. and Papa, 2007, Comportamento idraulico e meccanico della serie piroclastica di Monteforte Irpino. In *Piattaforme Evolute di Telecomunicazioni e di Information Technology per l'Offerta di Servizi al settore Ambiente PETIT-OSA a cura di C. Nunziata Aracne editrice srl Roma ISBN 978-88-548-1184-3.*

Olivares, L. and Picarelli, L., 2003, Shallow flowslides triggered by intense rainfalls on natural slopes covered by loose unsaturated pyroclastic soils, *Géotechnique*, 53(2), 283-288.

Olivares, L., Damiano, E. and Picarelli, L., 2003a, Wetting and flume tests on a volcanic ash, *International Conference on Fast Slope Movements - Prediction and Prevention for Risk Mitigation*, Naples, Italy, 399-404.

Olivares, L., Andreozzi, L., Damiano, E., Avolio, B., Picarelli, L., 2003b, Hydrological response of a steep slope in pyroclastic unsaturated soils, *International Conference on Fast Slope Movements - Prediction and Prevention for Risk Mitigation*, Naples, Italy, 391-397.

Olivares, L., Damiano, E., Greco, R., Zeni, L., Picarelli, L., Minardo, A., Guida, A. and Bernini, R., 2008, An instrumented flume for investigation of the mechanics of rainfall-induced landslides in unsaturated granular soils. *Submitted for publication in ASTM Geotechnical Testing Journal.*

Regalado, C.M., Munoz Carena, R., Socorro, A.R. and Hernandez Moreno, J.M., 2003, Time domain reflectometry models as a tool to understand the dielectric response of volcanic soils, *Geoderma*, 117, 313-330.

Sirangelo, B., Versace P. and Capparelli, G., 2003, Forewarning model for landslides triggered by rainfall based on the analysis of historical data file. *IAHS Red Book*, 278, 298-304.

Tomer, M.D., Clothier, B.E., Vogeler, I. and Green, S., 1999, A dielectric-water content relationship for sandy volcanic soils in New Zealand, *Soil Sc. Soc. Am. J.*, 63, 777-781.

Topp, G.C., Davis, J.L., and Annan, A.P., 1980, Electromagnetic determination of soil water content: measurement in coaxial transmission lines, *Water Resour. Res.*, 16, 574-582.

van Genuchten, M.Th., 1980, A closed-form equation for predicting the hydraulic conductivity of unsaturated soil, *Soil Sci. Soc. Am. J.*, 44, 615-628.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Landslide Data Analysis with Gaussian Mixture Model

V. Timonin[a], S.B. Bai[b], J. Wang[b], M. Kanevski[a] and **A. Pozdnoukhov[a]**

[a] *Institute of Geomatics and Analysis of Risk, University of Lausanne, Switzerland (Alexei.Pozdnoukhov@unil.ch)*

[b] *Nanjing Normal University, Nanjing, China*

Abstract: In this paper we present the approach for the analysis and modeling of landslide data using the Gaussian Mixture Model. We model the probability density of the landslide events in the high-dimensional space of parameters, conventionally used for predicting the landslide susceptibility. This work describes the application of the method for the area of Bailongjiang River, in northwest China. The available information includes the digital elevation model of the region, geological map and different GIS layers including land cover data obtained from satellite imagery. The landslides were observed with aerial imagery and documented during the field studies.

Keywords: Susceptibility mapping; Landslides; Gaussian Mixture Model.

1. INTRODUCTION

In this paper we present an approach to the analysis and modeling of landslide data using one of the baseline statistical and machine learning algorithms known as Gaussian Mixture Model. The information available for modeling includes the digital elevation model of the region with resolution of 30 meters, the geological map, the climatological information obtained as an output of large scale physical model, and several different GIS layers including land cover data obtained from satellite imagery. The preceding landslides were observed during the field studies, providing the coordinates and the shapes for 144 landslides.

Traditionally, the logistic regression approaches are commonly used in landslide susceptibility mapping. However, complex relations of the landslide phenomena with different factors ask for modeling with more adequate tools. These are, for example, the robust non-linear non-parametric data-driven algorithms of machine learning. The application of machine learning algorithms [Hastie et al., 2001, Haykin 1998] for the analysis and modeling of landslides, in particular for the landslides susceptibility analysis, have already demonstrated their predictive capability and robustness [Ermini et al., 2005; Melchiore et al., 2008; Yao et al., 2008].

Another important concern deals with the setting of the problem of landslide susceptibility mapping. Traditionally, it has been considered as a classification problem aiming to discriminate the dangerous zones from the safe ones. While it is relatively straightforward to define the dangerous zones as the ones where landslides has happened, the selection of the safe ones which are both reliably safe and brings useful discriminative information is much less evident. Having said that, we will study the setting of the susceptibility mapping as a probability density modeling problem, that is, the task of this study is the modeling of the probability density of the landslides in the space of input parameters.

A nonparametric approach based on Gaussian Mixture Models (GMM) is applied for nonlinear mapping of probability densities of landslide areas. Input information contains eleven features and tuning of GMM model was based on Expectation-Maximization algorithm.

2. GEOLOGICAL ENVIRONMENT OF THE STUDIED AREA

The study area covers 1361 km^2 mainly on the Bailongjiang River, in northwest China. The site includes 6 sub-basins and the landslide area covers an area of 19.67 km^2. The site lies in the middle south of the west wing of Qinling orogen. The area is formed by Qinghai-Tibet tectonic belt and Wudu arc structure and is affected by unlift of the Qinghai-Tibet plateau. When structure movement shift its forms from main horizontal movement to main vertical movement, it is a very profound effect to the wings of Qinling in this location. The unbalanced vertical movement creates the extreme development of folds, crushs, faults and joints in the location. Lithology is mainly phyllite, schist, slate, carbonate rocks and all kinds of clastic rocks. Today in China tectonic units can be divided into three teconic systems: Alps - Himalaya tectonic system, shortly as Tethys tectonic domain; the Pacific tectonic system and the new generation of Central Xibailie tectonic system (or ancient Asia tectonic domain). Two tectonic belts: Helan-Sichuan-Yunnan north and south of the belt and Kunlun-Qilian-Qinling-Dabieshan east and west of tectonic belt (central orogenic system).

2.1 Lithology of stratum

The case study area possesses much of loose soil layer due to all kinds of causes, such as the Silurian, Devonian, Carboniferous, Permian and Triassic in Paleozoic and the Triassic of Mesozoic and Quaternary. The oldest stratum is the Silurian in the area which mainly distributes in the two shores of Bailongjiang River and forms multiple anticlines. The stratum constructed mainly by shallow marine sedimentary metamorphic rocks and carbonate rocks. From east to west, the sediment thickness changes from thin to thick, sediment is from coarse to fine and carbon capacity is from low to high. Lower Silurian is composed by carbonate rocks and clastic rock and is the nuclear of multiple anticline of Bailongjiang river, at the same time, it lies in the south of Bailongjiang river. Between middle and upper Silurian is called Bailongjiang group maked up of clastic and carbonate rocks. It mainly distributes in the Bailongjiang northern shore of the east of two estuaries and southern shore from the west of two estuaries to Zhouqu. Devonian develops greatly in the area and mainly located in two wings of Bailongjiang multiple anticlines. Middle Devonian and upper Devonian have a close connection with landslide, while as we know Middle Devonian mainly distributes in the north wing of Bailongjiang multiple anticlines.

2.2 Geological Structure

The study area is located in the intersection of the new regional structure north-south and east-west tectonic activity zone in the north edge of the Qinghai-Tibet Plateau. In the tectonic, it is the West Qinling orogen ofa Qinling micro-plot .With the long-term impact of tectonic activity , displayed extrusion zone along in the western ;It is along a EW trend in the eastern ,including folds and faults,while the stratums also distribute to the NW-NWW as a band Since the Cenozoic, Indian plate and Eurasian plate collision developed mountains, the Qinghai-Tibet Plateau uplift ,and large-scale strike-slip of altiplano crust result in stress-strain field and tectonic activity exceptionally complex. Strong seismic activities and water system development of Bailong River drainage are the main reasons for the landslide and damming disasters.

2.3 Landform

In the landform, the study area is located in the southern of the Qinling Mountain, and has tall upright mountains, steepness terrain, deep valleys, fast-flowing river and presents V-shaped valley or canyon terrain features. It is the erosion middle and high mountain landform.

2.4 Hydrology and climate characteristics

The study area is the drainage of Bailong River, which is a tributary of the Changjiang River. The water system of Bailong River is plume-shaped elongated at NW-SE. The mainstream of Gansu is 475 km long and the both sides of river are asymmetric, South being wide and North narrow. Bailong River has many tributaries. The most lengthy tributaries are Min River, Beiyu River and Baishui River.

The rainfall is not balanced in the time and space. In time, the rainfalls mainly concentrates in the 4-10 months, from 6-9 months, Dangchang rainfall is 59.9 percent of the annual rainfall, Wudu is 65.5%, Wenxian is 62% and Zhouqu is 61.7%. Short and heavy rainfalls are typical for the region.

In the spatial distribution, rainfall trends to decrease from south to north such as from 900mm of the Wudu southeast to 500mm of the Wudu northwest. The rainfall intensity tends to increase with elevation. The average annual rainfall is 400-500mm below the elevation of 1500m, 500-600mm between 1500m and 2000m and more than 600 mm above 2000m.

Different climatic zones differentiated by elevation are easily distinguished in this region. The temperature decreases with altitude, providing the annual average temperature to be below 5°C at the elevations of more than 2500m. There are very severe differences between the extreme yearly maximum and minimum temperatures.

3. GAUSSIAN MIXTURE MODEL

3.1 Model description

First let us consider briefly some theoretical questions concerning Gaussian Mixture Models and mapping via density modeling. Usually Mixture Models are used for the ***density estimation*** of the data. Density estimation is the construction of an estimate, based on observed data, of an unobservable underlying ***probability density function (p.d.f.)***. Mixture Model estimates density distribution in a form of a linear combination of some simple functions (called components, units, or kernels):

$$p(x) = \sum_{j=1}^{m} p(x \mid j) P(j) \tag{1}$$

Such representation of p.d.f. is called a ***mixture distribution*** [Titterington ***et al.***, 1985]. ***P(j)*** are mixing coefficients. In a Bayesian framework, ***P(j)*** can be considered as ***prior*** probabilities of any data point having been generated from component ***j*** of the mixture. These priors, like any probability, should satisfy the following constraints:

$$\sum_{j=1}^{m} P(j) = 1, \qquad 0 \le P(j) \le 1 \tag{2}$$

The component density functions ***p(x|j)*** are also normalised such that

$$\int p(x \mid j) dx = 1 \tag{3}$$

In a Bayesian framework, one can introduce the corresponding ***posterior*** probabilities, which can be expressed using theorem of Bayes in the following form:

$$P(j \mid x) = \frac{P(x \mid j) P(j)}{p(x)}, \quad \sum_{j=1}^{m} P(j \mid x) = 1, \qquad 0 \le P(j \mid x) \le 1 \tag{4}$$

The value of ***P(j|x)*** – the ***posterior*** probability - represents the probability that a particular component ***j*** was responsible for generating data point ***x***.

The most important property of such model is that it can approximate any continuous density with an arbitrary accuracy (if the number of components is large enough and the parameters of the model are chosen correctly, see, for example, [Silverman, 1986]).

The most often choice of the function for the component is a Gaussian one:

$$p(x \mid j) = \frac{1}{(2\pi\sigma_j^2)^{d/2}} \exp\left\{ -\frac{\left\| x - \mu_j \right\|^2}{2\sigma_j^2} \right\} \tag{5}$$

Here the σ^2 parameter for each j from $1...m$ interval is a scalar and we deal with a model of m *isotropic* components. In more general case, parameter σ^2 may be presented as a full covariance matrix Σ. Covariance matrix is a squared symmetrical matrix of dimension d by d and the number of parameters is $d(d+1)/2$. In this case the (5) can be rewritten as:

$$p(x \mid j) = \frac{1}{(2\pi)^{d/2}(\det \Sigma_j)^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1}(x-\mu_j)\right\} \quad (6)$$

where *det* denotes determinant and Σ^{-1} is an inverse of Σ matrix.

A Mixture Model with Gaussian components is called Gaussian Mixture Model.

A model with *anisotropic* components is much more flexible but less stable and requres more calculations and patemeters tuning. It is vital especially in case of complex multidimensional data. So in the current study isotropic model will be used.

3.1 Tuning of parameters. Expectation-Maximisation algorithm

Various procedures have been developed for determining the parameters of the GMM from a given data set. One of the most famous is an *expectation-maximisation (EM)* algorithm [Dempster *et al.*, 1977].

At the first step, *E-step*, algorithm tries to detect which component is responsible for generating each point of the data set. At the next step, *M-step*, parameters of the mixture taking into account the results of the first step (by maximising the expected likelihood) are tuned.

Updating of parameters of the mixture model by EM algorithms is given by the following iterative formulas:

$$\mu_j^{new} = \frac{\sum_n P^{old}(j \mid x^n) x^n}{\sum_n P^{old}(j \mid x^n)} \quad (7)$$

$$(\sigma_j^{new})^2 = \frac{\sum_n P^{old}(j \mid x^n) \left\| x^n - \mu_j^{new} \right\|^2}{\sum_n P^{old}(j \mid x^n)} \quad (8)$$

$$P(j)^{new} = \frac{1}{N} \sum_n P^{old}(j \mid x^n) \quad (9)$$

where "*old*" means values from the previous step of the algorithm.

4. CASE STUDY

The GMM will be used to model the probability density function which, being applied to the previously unseen territories, can be used for mapping the landslide susceptibility, i.e. the potential areas where landslides can occur. As a training data set, the areas (scars and masses) where landslides were observed are used. The training data set is formed pixel-wise from inside the landslides polygons, which were outlined by experts. To validate the quality of mapping, the studied area was divided in two parts: west (76 polygons, training part) and east (68 polygons, validation one). In Figure 1 the digital elevation model (DEM) of the studied area with 144 polygons of observed landslides and data division into training and validation subsets are presented.

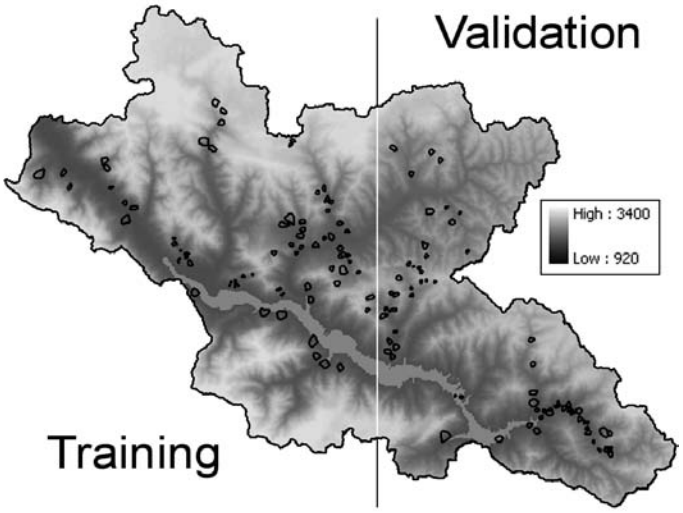


Figure 1. Digital elevation model (DEM) of the studied area. Polygons are the areas (scars and masses) of the observed landslides outlined by experts (total 144). Left (76 polygons, west) part is a training data set, right (68 polygons, east) is a validation one.

Eleven different features were presented as input information: DEM, slope, aspect, curvature (profile and plan), distance to inhabited locality, distance to river, distance to road, fall distance, lithofacies, and land use types. Note that last 2 features are categorical and cannot be presented to the model based on the distance calculation. So it is necessary to present such type of information in some encoded manner. The widely used schema for this purpose is to present m-category feature as an m-dimensional attribute with a single non-zero component corresponding to the category number. For example, the landuse type feature has six categories. So, it is necessary to define six-dimensional attribute as input information for the model. Category 1 will be presented as (1,0,0,0,0,0) vector, category 2 as (0,1,0,0,0,0), and so on. The lithofacies feature has seven categories, and the seven-dimensional attribute was defined. As a result, the original 11 features of information are presented for the model as 22-dimensional input vector.

Two GMM models with 12 and 40 isotropic Gaussian kernels (5) were explored. The density distribution (1) was calculated with EM-algorithm until convergence. The resulting model is combined as a model with the maximum likelihood from 50 tries with different initial locations of the units.

Both models are flexible enough to reproduce the training data, that is, they provide high probabilities at the areas where the landslides have occurred. The next step in model evaluation is to calculate the probability density values for all studied area. The validation of the model will be carried out using the areas of validation landslides. Given sufficient performance, the model could be used to explore the areas where high susceptibility values are predicted but historically the landslides have not yet occurred (or were not yet registered by experts).

5. RESULTS

The results as the maps of the probability density values obtained with two models are presented in Figure 2. Both maps are very similar instead of a different number of kernels (number of free, tuned parameters) in the models. The model with 12 kernels is somewhat smoother than the 40-kernels one. Both models cover the areas outlined by the training polygons in the validation (east) part of the area. Also there are some well-defined areas with high density values in the areas without training polygons. These areas can be treated as potentially landslides dangerous (where historically landslides were not occurred or not registered by experts).

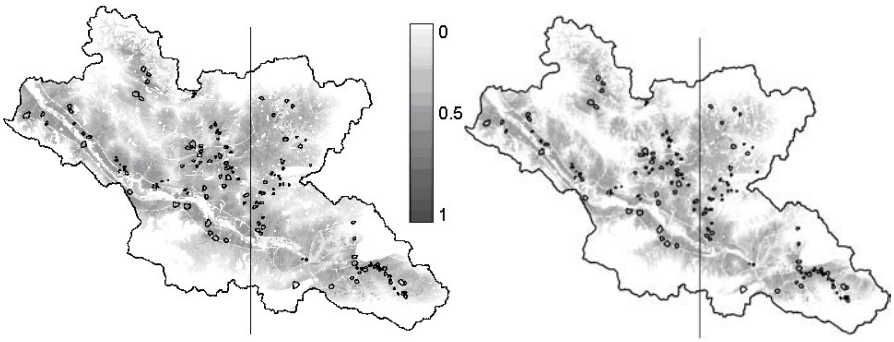


Figure 2. Map of mixture density estimation of the landslide susceptibility areas by 12-kernels model (left), and 40-kernels model (right. Darker colours present pixels with higher density values. Training and validation areas are outlined with polygons.

The analysis of the outputs can also be done by considering the obtained landslide susceptibility model with respect to the input physical parameters. The histogram of susceptibility values and slope inclination is illustrated in Figure 3 (left). It suggests that most landslides in the area are observed at the inclinations of 30-35 degrees.

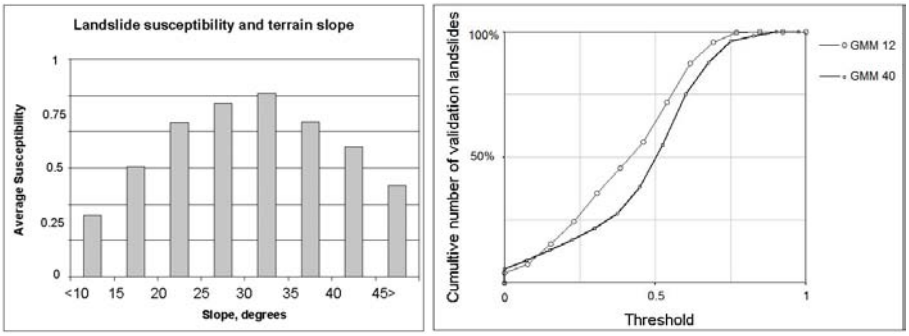


Figure 3. Left: Average susceptibility suggested by GMM model with respect to slope inclination. Right: cumulative number of landslides with respect to the selected threshold, computed pixel-wise on the validation dataset.

Figure 3 (right) illustrates the cumulative number of landslide cells (calculated pixel-wise) when changing the susceptibility threshold. One can notice that at the threshold level of 0.8 the model reproduced all of the validation landslides. By further taking several thresholds on the probability density values, one may select the conventional categorical hazard levels.

6. DISCUSSION AND CONCLUSIONS

The current study shows the feasibility of the GMM modeling as a tool for landslides predictions or susceptibility mapping. Model does not only reproduce the areas where landslides were occurred (both in the training and validation parts of the studied area) but also predict potentially hazardous areas. By selecting the thresholds for the continuous output of GMM model, one may delineate the conventional hazard class zones. While the results on validation dataset were quite promising in terms of landslide detection, more analysis is needed to analyse the false alarm rate. This would in turn require finding a non-trivial dataset of safe areas, which can be obtained with long and intensive field studies.

The same problem appears when trying to apply the conventional susceptibility mapping technique such as logistic regression. GMM avoids the latter obstacle by directly modeling the probability of the dangerous class in the input space of parameters. The results are robust from the point of view of model stability, that is, with respect to the number of units in the mixture. Though, GMM model is based on pair-wise distances in the input space of high dimension thus may be prone to the so-called “curse of dimensionality”.

The next steps in this research may be directed towards the selection of the most informative input features or their (generally, nonlinear) combinations for improving the quality of mapping. Other machine learning algorithms, including the one-class Support Vector Machines could also find interesting applications in landslide susceptibility mapping.

ACKNOWLEDGEMENTS

The authors wish to thank Swiss National Science Foundation (SNF) for the partial financial support of the projects “GeoKernels”(200021-113944).

REFERENCES

Dempster, A.P., N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society*, B 39 (1), 1-38, 1977.

Ermini L., Catani F., and Casagli N. Artificial neural networks applied to landslide susceptibility assessment. *Geomorphology*, 66, 327-343, 2005.

Hastie T., Tibshirani R. and Friedman J. (2001). The elements of Statistical Learning. Springer.

Haykin S. Neural Networks. A Comprehensive Foundation. (1998). Prentice Hall.

Melchiore C., Matteucci M., Azzoni A., and Zanchi A. Artificial neural networks and cluster analysis in landslide susceptibility zonation. *Geomorphology*, 94, 379-400, 2008.

Titterington, D.M., A.F.M. Smith, and U.E. Makov. Statistical Analysis of Finite Mixture Distributions. New York: John Wiley, 1985.

Silverman, B.W. Density Estimation for Statistics and Data Analysis. Chapman and Hall, New York, 1986.

Yao X., Tham L.G., and Dai F.C. Landslide susceptibility mapping based on support vector machines: A case study on natural slopes of Hong Kong, China. Geomorphology, doi: 10.1016/j.geomorph.2008.02.011, 2008.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Modelling the Portalet landslide mobility (Formigal, Spain)

J.A.Fernández-Merodo[a], G.Herrera[b], P.Mira[c], J.Mulas[b] and M.Pastor[c], L.Noferini[d], D.Mecatti[d], G.Luzi[d]

[a] *Universidad Rey Juan Carlos, Dept. Arquitectura de Comp. y CC. de la Comp., Móstoles (jose.fernandez@urjc.es)*

[b] *IGME, Unidad de Riesgos Geológicos, Ríos Rosas 23, 28003 Madrid (g.herrera@igme.es)*

[c] *CEDEX, Dept. Ingeniería Computacional, Alfonso XII 3 y 5, 28014 Madrid, (mpastor@cedex.es)*

[d] *University of Florence, Dept. of Electronics and Telecommunications (DET), Firenze – Italy, (linhsia.noferini@unifi.it).*

Abstract:

The Portalet landslide is a large, active slope failure located on the upper part of the Gállego River valley, in the Central Spanish Pyrenees (Sallent de Gállego, Huesca). The landslide is located on a southwest-facing hillside of Petrasos Peak, close to the ski resort of Formigal. The dimensions of the landslide are 500 m long and up to 700 m wide at the toe extended over an area of 0.35 km2 that shows superficial cracking and distinct ground displacements. Displacements are quite important (several cm in a month) from summer 2004 when the construction of a parking area was carried out by digging the foot of the slope. Field observations show a close relation between ground movement and rainfall intensity. Monitoring and predicting the landslide behaviour is necessary to guarantee the safety of this ski resort area. Monitoring was performed with four differential GPS (DGPS) campaigns between May and November 2006 and installing a novel ground based SAR (GBSAR) monitoring system. GBSAR measurements were recorded continuously for 47 days, from the 5th of October 2006 to the 21st of November 2006, providing up-to-date displacement maps of the landslide area at a rate of 1 per hour. Two numerical models have been proposed to forecast the recorded movements directly from rainfall intensity. A simple 1D infinite slope model solving the momentum equation gives a first approximation. The model includes a viscoplastic behaviour and takes into account directly the known daily rainfall intensity and the dissipation of the excess pore-fluid through a simple consolidation equation. A fully coupled hydro-mechanical finite element analysis in plane strain conditions using an elastoplastic constitutive model has also been performed. The results show that prediction of slope behaviour is feasible if data parameters are well documented.

Keywords: GBSAR monitoring; 1D infinite slope model, viscoplastic model, consolidation model, coupled hydro-mechanical finite element modelling

1. Introduction

Two broad categories of models can be pointed to predict landslide mobility: the phenomenological models and the physically based models. The former employ empirical relationships, statistical or probabilistic approaches and artificial neural networks to relate soil movements and their causes; the latter provide relationships taking into account the

mechanical behavior of the soil. A summary of the different methods proposed in the literature can be found in Federico et al. (2004).

Physically based models have been mostly used in practical cases to estimate landslide occurrence and stability conditions for a given scenario through a stability factor based on limit equilibrium analysis. In the present paper, a simple 1D infinite slope model assuming a pre-existing slip surface where the shear strength is at residual conditions and including a viscoplastic behaviour (Vulliet and Hutter 1988) is used for predicting the movements of Portalet landslide. The movements are concentrated within a relatively narrow shear zone above which the sliding mass moves essentially as a rigid body. This model was also used by other authors (Van Ash and Van Genuchten 1990, Corominas et al. 2005). In their studies the hydrologic contribution is directly given from groundwater level changes measured by piezometers. Here, no precise information about the position of the groundwater level was available, a simple improvement is proposed to overcome this limitation and make prediction directly from the recorded daily rainfall intensity. Leaning on the observed velocity given by the monitoring data a dissipation model of the excess pore-fluid through a simple consolidation equation is proposed. The main advantage of such a procedure lies in its relatively simple applicability.

Not to mention earthquake study, time dependent analysis is requested when hydrological conditions change as in the case of rainfall, when resistant parameters reduce as in the case of strain softening and weathering processes or when creep behaviour is taken into account. Complex numerical models as the finite element method can provide a good understanding of the mechanism of failure in these cases as they can correctly reproduce the fully coupled hydrogeological and mechanical behaviour, moreover, advanced constitutive laws, complex 3D geometries and spatial variations in soil properties can be used. Many numerical modelling case studies of landslides mechanics have already been published focussing on strain localization (Dounias et al. 1988, Troncone 2002), visoelastoplasticity (Desai et al. 1995, Vulliet 2000) or hydromechanical coupling in large landslides (François et al. 2007).

The aim of the article is to use the monitoring data obtained by the continuous GBSAR measurements to calibrate the proposed models and to predict the Portalet landslide mobility directly from the recorded rainfall intensity.

2. Portalet landslide description

The landslide subject of this paper is located in Central Spanish Pyrenees (Sallent de Gállego, Huesca). This slope is composed mainly by slates and shales, all of which are of Carboniferous and Devenian age, characterised by an intense weathering and a high plasticity.

The landside, which extends over an area of 0.35 km2, presents a steep cliff that corresponds to the main scarp and it is of the rotational kind with unidirectional movement from a global point of view (zone a in Figure 1). The mass movement was triggered by the excavation of a parking area damaging a connexion road to France. Constructive solutions were carried out to minimize the slope displacement rate. The



Figure 1. Geomorphological sketch and geological cross-sections of the Portalet landslide. Location of total station points, inclinometers and DGPS ground control points.

landslide has generated several blocks sliding between each other along pre-existing slip surfaces (zone d in Figure 1); section AA' has been chosen for the modellisation and previous stability analysis exhibited a nearly translational behaviour of this slide.

A total of four DGPS campaigns were performed in the Portalet landslide between May and November 2006 involving around 100 benchmarks. Overall, during this period maximum displacements of 22 cm were observed in 180 days.

A GBSAR sensor (Nofirini et al. 2007) was installed at a distance of about 600 m from the main body of the landslide on a hill overlooking the parking area, providing automatically one displacement map per hour continuously along 47 days (over 1000 maps) involving around one million pixels with a resolution of 2 x 2 m. During this monitoring campaign a maximum displacement of 14 cm was estimated with an error in the range of millimetres.

The analysis of the monitoring data measured with conventional techniques in different time periods, and particularly of the continuous GBSAR displacement temporal evolution has allowed understanding some of the characteristics of the Portalet landslide mechanism. It has been established that the rate of displacement is very sensible to rainfall. From this we infer that there is a rapid, almost instantaneous response of the ground water level to rainfall events.

3. Modelling using a simple viscoplastic sliding-consolidation model

Several authors (Van Ash and Van Genuchten 1990, Corominas et al. 2005) have used a simple 1D infinite slope model to predict slow landslides mobility directly from groundwater level changes. The model includes a viscoplastic behaviour (Vulliet and Hutter 1988) of the landslide but seems to be too simplistic since neither 2D and 3D effects are taken into account nor elastic deformation is predicted. Nevertheless, the model can give a first approximation of the landslide kinematics and can be used to reproduce the ground displacements and velocities measured with the GBSAR and DGPS techniques.

The landslide is assumed to be as a translational infinite slide with constant depth h and constant slope α, in this case the equilibrium over a unit surface can be described with the simple equation:

$$\mathbf{F} - \mathbf{F}_r = \mathbf{F}_i + \mathbf{F}_v$$

Where F represents the destabilising force equal to the weight of the considered volume, F_r is the resisting force that can be estimated with a Mohr–Coulomb criterion, F_i is the inertial force and F_v represents a viscous force usually dependent of the strain rate of the shear zone that can be evaluated using a Bingham model. The momentum balance equation can be written over the slope direction as:

$$\tau - \left[c + \left(\sigma_n - p_w(t)\right)\tan\phi\right] = m\,a(t) + \frac{\eta}{d}v(t) \qquad \text{or}$$

$$m\,a(t) + \frac{\eta}{d}v(t) = mg\cdot\sin\alpha\cos\alpha - \left[c + \left(mg\cos^2\alpha - p_w(t)\right)\tan\phi\right] \tag{1}$$

where τ is the destabilising shear stress, σ is the normal stress, ϕ is the friction angle, c is the cohesion, m is the mass $= \rho \cdot h$, ρ being soil density and h the height of the landslide, η is the viscosity, d is the thickness of the shear zone, $p_w(t)$ is the pore water pressure $p_w(t) = z(t)\cdot\gamma_w\cdot\cos^2\alpha$ assuming a parallel flow to the slope surface, z is the position of the groundwater level, γ_w is the specific weight of water, a(t)is the acceleration and v(t) is the velocity.

Prediction of ground motion is obtained by solving equation (1) given the geometry of the landslide (h, α and d parameters), material properties (ρ,ϕ, c, η parameters) and initial and time variation of the ground water level $z(t)$. The parameters are obtained by field observations, in-situ tests and laboratory tests. If one of these parameters is not available it can be estimated by back analysis in a fixed period of time, in this case prediction must be

made in a different period of time. If more than one parameter is not available, the prediction is analyzed for given scenarios or providing reasonable estimates. In the present study 2 parameters are not determined: the viscosity η and the position of the ground water level $z(t)$.

Section AA' defined in Figure 1 has been chosen for the modelisation, the average inclination of the section is α=14º. Information from inclinometer I1 and borehole S1, located on the section, indicates that the depth of the failure surface is located at h=12m and the thickness of the shear zone is d=0.4 m. Tests on undisturbed samples (ICOG 2005) assign an average density of the material ρ=2310 kg/m^3 and an estimate shear strength ϕ=18º, c=0 t/m^2 for the fragmented slate material located in the shear band.

No precise information about the position of the groundwater level was available, due to environmental constraints no piezometer was installed. The equilibrium state (zero velocity and zero acceleration) corresponding to a zero value of the right hand side of equation (1) is achieved when the groundwater level is at z=6.499 m. An initial state $z(0)$=6.5 m is postulated. The results are very sensitive to the initial conditions, small variations of this value have a strong effect on the final response of the landslide mobility.

A simple model is now proposed to make prediction directly from the known daily rainfall intensity and not from groundwater level changes. In this case, the effect of the snow melting during the spring season cannot be considered.

R.M. Iverson 2000 presented a complete mathematical model based on Richards's equation to evaluate the hydrological process during rain infiltration in saturated material. In the present study, measured displacements show that the landslide reacts almost immediately to rainfall inputs. This rapid response is likely due to the presence of superficial cracks and preferential drainage pathways. Changes in groundwater level have been taken directly proportional to the rainfall intensity through:

$$dz = \frac{I_{\text{rainfall}} / 1000}{n}$$

where the rainfall intensity is recorded in mm/m^2/day and n is the material porosity supposed to be constant equal to 0.15. No runoff is contemplated when rainfall intensity exceeds the infiltration capacity.

Following the work done by Hutchinson 1986 and Pastor et al. 2002, describing the variation of excess pore-fluid pressure during flow slides, an attempt was made to reproduce slow consolidation process in the Portalet landslide. Assuming a parallel flow to the slope surface, the dissipation of the excess pore-fluid pressure in the saturated layer of length h_s is governed by the Terzaghi's one dimensional consolidation theory, one solution is given by:

$$ep_w(t) = ep_{w0} \cdot e^{\frac{-t}{T_v}}$$

where ep_{w0} is the initial excess pore pressure, T_v is a time factor defined by $T_v = \frac{4h_s^2}{\pi^2 c_v}$ and c_v is the consolidation coefficient. T_v controls the dissipation time of the excess pore-fluid pressure, two values have been considered, T_v=1s corresponding to an instantaneous dissipation and T_v =5.e6s corresponding to a slow consolidation process.

The viscosity has been determined by back analysis during the period of time from October, 5th to November, 21st 2006 corresponding to the GBSAR campaign. Table 1 summarizes the set of values taken in the different computations done.

Set of parameters	$z(0)$ in m	T_v in s	η in Pa.s
Computed 1	6.5	1.	7.5 e9
Computed 2	6.5	5.e6	6.5 e10

Table 1. Values taken in the different computations done

Figure 2 presents the comparison between GBSAR measured and computed displacements and velocities for the "computed 1 and 2" set of parameters from the 5th of October to the 21st of November 2006. For the "computed 1" analysis, the velocity is directly proportional to the rainfall intensity and it is constant equal to 0.07 cm/day during the dry periods. In the "computed 2" analysis slow consolidation process is taken into account, the velocity pattern shows a smoother response of the landslide to rainfall input.

Figure 3 shows the relation between the predicted water table height evolution and rainfall intensity. In the case of "computed 1" the water level is constant in dry periods and varies instantaneously when it rains. On the contrary, the "computed 2" water table height is higher than "computed 1", because the excess pore-fluid pressure has no time to dissipate between two following rain events.

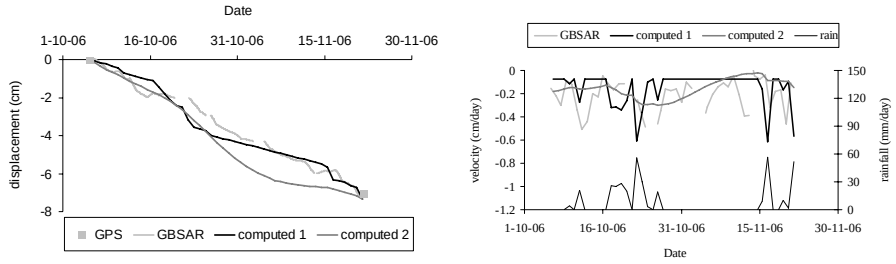


Figure 2. Measured and computed displacement and velocity from the 5th of October to the 21st of November 2006

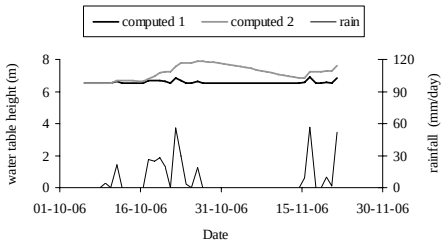


Figure 3. Predicted water table heights and rainfall intensity from the 5th of October to the 21st of November 2006

4. Modelling using a 2D coupled hydro-mechanical finite element model

A fully coupled hydro-mechanical plane strain finite element analysis has been performed to overcome limitations of the simple viscous model. The finite element code GeHoMadrid (Pastor et al. 1998) for transient Biot's equation resolution has been used.

The geological structure of the landslide has been obtained by means of geological, geomorphological assessment and drilling data (ICOG 2005). Section AA' has been discretized using quadratic triangular elements (6 nodes for displacement degree of freedom and 3 nodes for pore pressure degree of freedom). The mesh is composed by 2151 triangles and 3975 nodes, Figure 4.

Material behaviour has been represented with a Drucker-Prager constitutive model using material parameters defined in Table 2. Saturated condition is taken in all the materials.

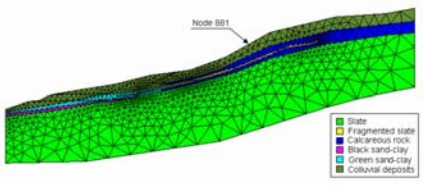


Figure 4. Finite element mesh and geological structure

Material	γ (Kg/m^3)	E (N/m^2)	ν	φ (°)	c (N/m^2)
Colluvial deposit	2100.	15.e6	0.2	29.	0.
Green sand-clay	2040.	14.e6	0.2	18.	0.
Black sand-clay	2220.	25.e6	0.2	30.	0.
Calcareous rock	2170.	40.e6	0.2	38.	0.
Fragmented rock	2170.	20.e6	0.2	25.	20.
Slate rock	2770.	4000.e6	0.2	31	640

Table 2. *Material parameters*

Two important factors have an influence on the slope stability: initial stress condition and ground water level position. Initial conditions have been treated in a first attempt modelling the construction process of the parking area described in section 2. Excavation process induced local instability near the toe of the excavated slope. The local collapse was also observed on the field and an immediate repair solution was carried out by replacing the excavated material. Consequently the final executed profile has been chosen as the studied geometry.

No information about the position of the ground water level was available. The initial hydraulic condition has been chosen performing an initial static stability analysis. The transient response taking into account rain condition has been performed in a second analysis.

4.1. Stability analysis

Stability analysis is performed in two phases: in the first one, gravity is progressively applied using a high value of suction on the upper boundary keeping the slope in a stable state. This condition is equivalent to imposing a water table parallel to the surface at a depth of 10 metres. In the second one, suction on the upper boundary is progressively reduced performing successive static equilibrium analysis, the water table is progressively raised to the surface. Plastic deformation develops along a shear surface until failure occurs for a given boundary suction condition.

A nearly linear behaviour is observed during the first phase corresponding to the gravity activation. A more marked plastic behaviour can be observed during the second phase corresponding to the decreasing suction boundary condition. Collapse is reached when the suction is equal to 34105. Pa, it is equivalent to a water table located at 3.48 m below the surface, plastic deformation and displacement contours at this moment are presented in Figure 5. The computed location of the rupture surface agrees very well with field observations. The observed scarps, Figure 1, are located at the same position on the upper part of the slope. Besides, rupture surface located at 12 m given by borehole 1 fits in with the computed rupture surface.

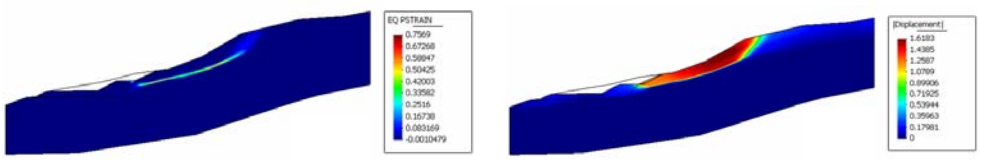


Figure 5. Equivalent plastic deformation contour and displacement contour (in m) at failure

4.2. Transient analysis with hydrological changes

The initial condition of the transient analysis is chosen from the stability analysis. The equilibrium state corresponding to a ground water level located 6 meters below the surface has been chosen. The choice of this initial state has an influence on the transient analysis,

states corresponding to the highest values of the ground water level can induce sudden collapse for small variations of pore pressure.

Rain is modelled using an input flow condition on the upper boundary. This condition supersedes the initial given hydraulic prescription, to do this an equivalent flow condition must replace the suction prescribed condition.

Different values of permeability (k_w=1.e-5m/s, k_w=2.e-5m/s, k_w=3.e-5m/s) constant over the domain, have been postulated. Figure 6 presents a comparison between GBSAR measured and computed displacement and velocity at node 881 during the autumn 2006 campaign. Permeability has an important influence on time response, further investigations must be carried out to correctly determine this parameter in each material as it can give also a different mechanical response in different zones of the slope profile.

It can be seen how motion is governed by rain infiltration. The model allows prediction of swelling process during no rain periods corresponding to the excess water pressure dissipation. The highest values of computed velocity coincide with long time rain periods, when excess pore pressure has no time to be dissipated. The observed discrepancy: the difference in the gradient of the measured and computed displacement curves and the zero computed velocity during no rain period can be reduced with a viscoplatic model.

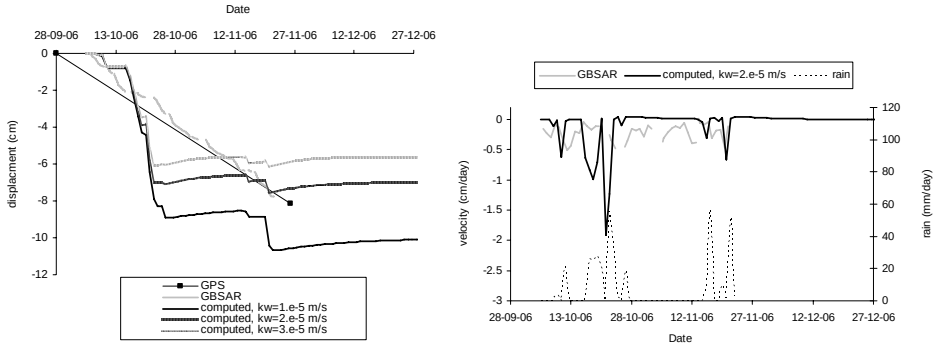


Figure 6. Measured and predicted displacement and velocity at node 881

5. Conclusions

An attempt was made to model the slope mobility directly from rainfall intensity trough a simple 1D infinite model. The model includes a viscoplastic behaviour and takes into account directly the recorded daily rainfall intensity and the dissipation of the excess pore-fluid through a simple consolidation equation. In spite of the simplistic hypothesis done and undetermined parameters, the model can give a first approximation of the landslide kinematics.

A fully coupled hydro-mechanical finite element model overcomes limitations of the 1D model as it includes multidirectional behaviour, elastic response and delayed answer due to consolidation. A forthcoming analysis, taking into account viscoplastic behaviour and using the GBSAR data that are being collected will give a better approximation within an extended period of time.

After calibrating the unknown model parameters by back analysis in a fixed period of time, these models can be used to predict the landslide mobility in another period of time using scenarios of rainfall events with different return periods, or snow melt events.

Such local scale models can also be used to define danger criteria, for example when computed velocities reach unacceptable values or when velocities increase more and more for a same hydraulic solicitation. These models are useful tools for quantitative risk assessment and management strategies for landslides.

ACKNOWLEDGEMENTS

This work has been partially funded by the project GALAHAD, "Advanced Remote Monitoring Techniques for Glaciers, Avalanches and Landslides Hazard Mitigation" of the EU 6FP (Global Change and Ecosystems, Contract N. 018409) and by the Spanish Geological and Mining Institute (IGME) within project 2004051. The authors recognize the valuable support of Javier Andrés Álvarez and Jose Antonio Villacampa from ARAMON S.L, and Francisco Torrijo Echarri from ICOG, that have provided the geotechnical data used in this work and have also provided logistical support in the monitoring experiments. The rainfall series used in this work have been provided by the Spanish Meteorological National Institute (INM).

REFERENCES

Corominas J., Moya j., Ledesma A., Lloret A., Gili J.A. (2005) Landslides 2, 83-96.

Desai C.S., Samtani N.C. and Vulliet L. (1995) Constitutive modeling and analyzing of creeping slopes. Journal of Geotechnical Engineering, ASCE, 1221(1), pp. 155-175.

Dounias G.T., Potts D.M. and Vaughan P.R. (1988). Finite element analysis of progressive failure: two case studies. Computers and Geotechnics, 6, pp 155-175.

Federico A., Popescu M., Fidelibus C., Interno G (2004) On the prediction of the time of occurrence of a slope failure: a review. In: Proceeding 9th International Symposium on Landslides. Rio de Janeiro, Taylor and Francis, London, vol. 2, pp 979-983.

François B., Tacher L., Bonnard Ch., Lalui L., Triguero V. (2007) Numerrical modeling of the hydrological and geomechanical behavior of large slope movement: the Triesenberg landslide (Liechtenstein). Can. Geotech. J. 44, pp 840-857.

ICOG, 2005. Informe sobre las inestabilidades detectadas en una ladera en las inmediaciones del Puerto de Portalet (Sallent de Gállego, Huesca), donde se está ejecutando la construcción de un futuro aparcamiento. Technical Report, Ilustre Colegio Oficial de Geólogos de Aragón, Zaragoza, Spain.

Pastor M., Li T.C., Mira P. and Fernández-Merodo JA. (1998) GeHoMadrid a Finite Element Program for geotechnical applications. Internal report CEDEX.

Hutchinson JN. (1986) A sliding-consolidation model for flow slides. Can. Geotech. J. Vol. 23, pp. 115-126.

Iverson R.M. (2000) Landslides triggering by rain infiltration, Water Resources Research vol. 36, nº7, pp. 1897-1910.

Noferini L., Pieraccini M., Mecatti D., Macaluso G., Atzeni C., Mantovani M., Marcato G., Pasuto A., Silvano S., Tagliavini F. (2007) Using GB-SAR technique to monitor slow moving landslide, Engineering Geology 95, pp 88-98.

Pastor M., Quecedo M., Fernández-Merodo J.A., Herreros M.I. (2002). Modelling tailing dams and mine waste dumps failures, Geotechnique, vol. 52, nº 8.

Troncone A. 2002. Numerical analysis of a landslide in soiuls with strain-softening behaviour. Geotechnique 55 (8), pp. 585-596.

Van Ash Th.W.J. and Van Genuchten P.M.B. (1990) A comparaison between theoretical and measured creep profiles of landslides. Geomorphology 3, pp. 45-55.

Vulliet L., Hutter K. (1988) Viscous type sliding laws for landslides. Can. Geotech. J. 25, 467-477.

Vulliet L. (2000) Natural slopes in slow movement. In Modelling in geomechanics. Edited by Zaman M., Booker J.R. and Gioda G.. John Wiley & Sons, Chister, UK. Pp. 654-676.

The role played by mountain tracks on rainfall-induced shallow landslides: a case study

L. Cascini[a], **S. Cuomo**[a], **M. Pastor**[b]

[a] *University of Salerno, Department of Civil Engineering, Italy (l.cascini@unisa.it, scuomo@unisa.it)*
[b] *Centro de Estudios y Experimentación de Obras Públicas (CEDEX) and ETS de Ingenieros de Caminos, Madrid, Spain (mpastor@cedex.es)*

Abstract: Mountain tracks often play a key role for landslide occurrence inside shallow soil deposits affecting both superficial and sub-superficial groundwater circulation. This is essentially related to local changes of topography and soil hydraulic conductivity. Significant examples of tracks-related landslides were observed during the event that affected the Campania region (Southern Italy) in May 1998. These phenomena occurred inside source areas characterised by complex landforms, stratigraphy and groundwater regime. Referring to this event, the paper addresses the role played by the stratigraphy on the first-failure stage by using limit equilibrium and stress-strain analyses. The numerical results outline some important aspects that must be carefully considered in analyses aimed at an advanced characterisation of landslide source areas.

Keywords: Rainfall, Shallow landslides, Flow-type, Mountain tracks, Cuts.

1. INTRODUCTION

Rainfall is the most frequent triggering factor for shallow landslides in a variety of unsaturated soils [Alonso et al., 1995]. For these landslides, the infiltration process and consequent failure stage are primarily controlled by rainfall patterns (i.e. intensity and duration) [Rahardjo et al., 2001], soil initial conditions [Tsaparas et al., 2002], and local natural factors such as slope angle breaks, thinning of deposits and close-ended layers.

Slope instability phenomena can also be induced by anthropogenic factors such as mountain tracks and cut slopes, especially when acting on singular points of the hillslope. Mountain tracks are generally associated with changes in topography, discontinuities of stratigraphy and lower soil hydraulic conductivity values compared to those of the neighbouring zones [Gasmo et al., 2000; Luce and Black, 2001]. As a consequence, along the tracks, rainwater can generate surface runoff, ponding conditions and erosion processes that often produce significant rills and gullies, as reported for well documented cases of Thailand [Ziegler et al., 2004], Australia [Croke and Mockler, 2001], US Virgin Islands [McDonald et al., 2001] and Malaysia [Sidle et al., 2004]. Cut slopes, commonly associated to mountain tracks, can also cause increased erosion processes and volumes of water runoff. Finally, lateral redistribution of water runoff can occur at the bends of tracks affecting the adjacent zones [Ziegler et al., 2004].

The effects of anthropogenic factors on superficial and sub-superficial groundwater circulation are particularly relevant in some areas of the Campania region (Southern Italy) where pyroclastic deposits overlie carbonate massifs. A paradigm example is provided by the May 1998 Sarno-Quindici event when some destructive landslides systematically occurred inside source areas characterised by complex landforms, stratigraphy and groundwater regime. A comprehensive analysis of the occurred phenomena requires a careful investigation of the previously mentioned factors. Among these, the paper focuses on the soil stratigraphy taking into account the typical in-situ conditions and local boundary conditions related to mountain tracks.

2. CASE STUDY

2.1 The landslides occurred in May 1998

In May 1998, heavy rainfall [Cascini et al., 2000; Fiorillo and Wilson, 2004] triggered several destructive landslides of flow-type in unsaturated pyroclastic deposits overlying the Pizzo d'Alvano limestone massif (Sarno-Quindici event) [Cascini, 2004; Cuomo, 2006] caused by different and complex triggering mechanisms. Cascini et al. [2008] recognised that the 16% of the 133 analysed landslide source areas were related to the presence of mountain tracks (Fig. 1a, b), along the hillslopes facing the town of Quindici (Fig. 1c). These slope failures mostly occurred along the flanks of gullies that were extensively affected by anthropogenic modifications especially in the last decades [Calcaterra et al., 2003]. Particularly, the failures occurred at the bends of tracks (scheme M3a of fig. 1b), later propagating downwards and upwards along the hillslopes. Adjacent zones were also systematically affected by landslides (scheme M3b of fig. 1b) due to lateral enlargement of the original unstable areas. In some cases, source areas with compound landforms were observed also due to the combination of the above anthropogenic factors and natural discontinuities of the pyroclastic deposits (i.e. bedrock scarps up to tens of meters tall) (Fig. 2).

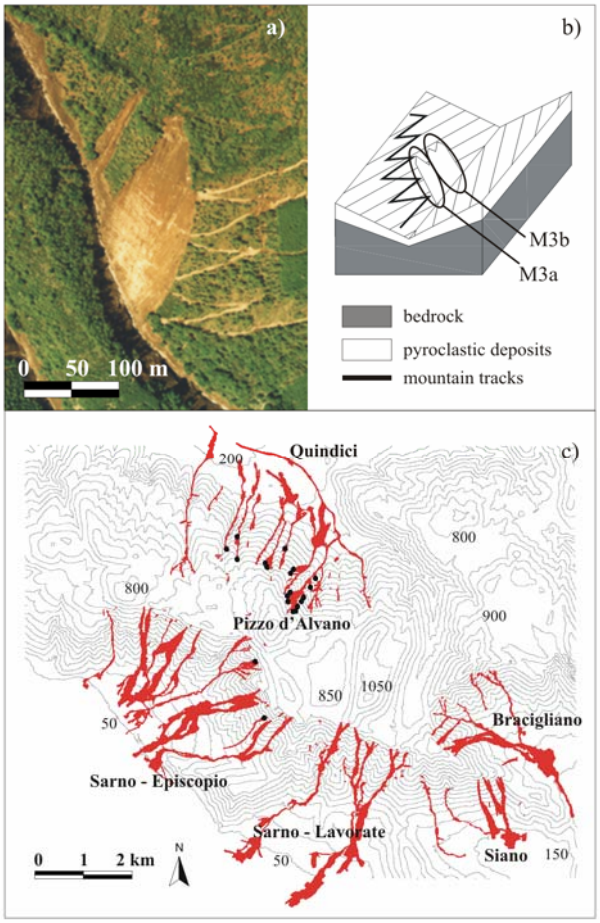


Figure 1. Shallow landslides of flow-type connected to mountain tracks during the May 1998 event.

The correspondence of some landslide source areas with mountain tracks is widely emphasised in the current literature. Di Crescenzo and Santo [2005] outline that 44% of these areas were connected to man-made cuts and tracks, located at distances lower than 10 m in most cases. Guadagno et al. [2005] recognise that 49% of the analysed failures occurred above man-made cuts and 12% involved fills downslope the tracks. These authors interpreted the failures as debris avalanches phenomena (i.e. characterised by triangular shaped source



Figure 2. In-situ evidences for some landslides occurred along the hillslopes facing the town of Quindici.

areas) and performed statistical analyses to correlate the height of cuts to other morphometrical parameters (i.e. apex angle of source areas and slope angle). However, poor correlations were obtained essentially due to the assumed triggering mechanism and because some key triggering factors were not included in the analysis (i.e. pore water pressures, local boundary conditions).

By modelling the first-failure stage of landslides induced by mountain tracks, Guadagno et al. (2003), referring to steady-state saturated flow conditions, outline that natural or anthropogenic discontinuities of pyroclastic deposits strongly increase the simulated displacements for upslope portions of the slope up to failure. The role of the above discontinuities for landslide occurrence is also emphasised by Crosta and Dal Negro (2003), on the basis of analyses that model the transient groundwater regime during the critical rainfall event of 4-5 May 1998.

These contributions, while providing useful insights on the topic, do not exhaustively analyse the relationships between tracks and landslides occurrence, neglecting some relevant aspects such as, for instance, the superficial water circulation and the role played by local stratigraphy. These aspects are addressed in the following sections that provide a preliminary modelling of the first-failure stage inside the landslide source areas.

2.2 Proposed approach, in-situ conditions and soil properties

Rigorous modelling of superficial and sub-superficial water circulation should take into account several factors, i.e. rainfall intensity, soil unsaturated conditions, soil hydraulic conductivity, topography, stratigraphy as well as anthropogenic factors that, in singular points and/or zones of the hillslope, can produce runoff phenomena and concentrated superficial water fluxes. To this aim, the physically-based model proposed by Savage et al. [2003] can be used. However, the lack of an accurate DTM can notably reduce the potentialities of these models. In such a case, a preliminary insight in the superficial water circulation can be obtained referring only to topography and computing the contributing areas along the hillslope (Pack et al., 1998; Tucker and Bras, 1998). The latter ones represent, at each point, the upslope catchment areas and the highest values can be reasonably assumed as corresponding to the zones where superficial rainwater concentrates. Once the hydraulic boundary conditions are outlined for different zones of the hillslope, pore water pressures can be computed with an infiltration model also considering soil unsaturated conditions through the numerical integration of the Richards' equation [Richards, 1931]. The failure stage can be thus simulated through either limit equilibrium or stress-strain analyses, adopting a Mohr-Coulomb criterion extended to unsaturated conditions [Fredlund et al., 1978].

This simplified approach is applied to the May 1998 event, using the available data set that outlines the presence of three main soil classes differently arranged along the Pizzo d'Alvano massif (Fig. 1c). Considering the aim of the study, parametric analyses are performed, referring to simplified schemes (Fig. 3) that reproduce typical stratigraphical settings observed all over the massif. The schemes include finer deep ashy soil (Class A) and/or coarser superficial ashy soil (Class B) also

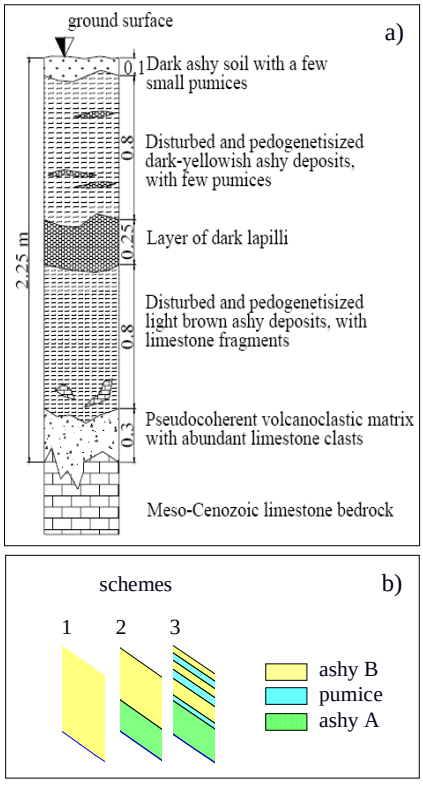


Figure 3. Typical stratigraphy of pyroclastic deposits in the study area.

eventually including continuous pumice soil layers. The mechanical properties of these soils in saturated and unsaturated conditions are provided by the scientific literature (Sorbino and Foresta, 2002; Bilotta et al., 2005; Cascini et al., 2005, among others) and summarised in table 1.

Table 1. Soil mechanical properties.

	γ_d (kN/m^3)	γ_{sat} (kN/m^3)	n (-)	k_{sat} (m/s)	c' (kPa)	φ' (°)	φ^b (°)	ν (-)	E (MPa)	ψ (°)
Class B ashy soil	7.30	13.1	0.58	10^{-5}	0 ÷ 5	36 ÷ 41	20	0.26	5 ÷ 7	10 ÷ 20
Pumice soil	6.20	13.1	0.69	10^{-4}	0	37	20	-	-	-
Class A ashy soil	9.10	15.7	0.66	10^{-6}	5 ÷ 15	32 ÷ 35	20	0.30	1 ÷ 3	10 ÷ 20

γ_d, (γ_{sat}): dry (saturated) unit weight of soil, n: porosity, k_{sat}: saturated soil conductivity, c': effective cohesion, φ': friction angle, φ^b: rate of increase in shear strength due to suction, ν: Poisson's ratio, E: Young's modulus, ψ: dilation angle.

3. GEOMECHANICAL MODELLING

3.1 Groundwater modelling

The modifications of superficial water circulation induced by mountain tracks were firstly evaluated, referring to a schematic steep (30°) hillslope, crossed by a steep track (14°) with one or more bends (Fig. 4a, 4b). The contributing areas were computed at each point of the hillslope through a detailed 0.5m×0.5m digital elevation model. The highest contributing areas concentrate at the bends of tracks (scheme M3a in Fig. 1 and 4) and complex scenarios are simulated for the scheme including more than two bends. Particularly, the zones adjacent to the bends are strongly affected by lateral redistribution of runoff (scheme M3b in Fig. 1 and 4). Concisely, both zones near and adjacent to the bends are outlined as prone to runoff phenomena and similar analyses can be extended to real cases if detailed topography (> scale 1:1000) and field data on the anthropogenic factors are available.

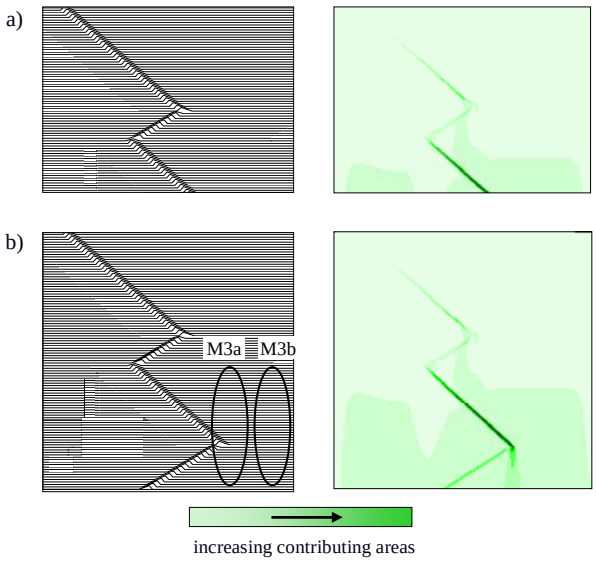


Figure 4. Contributing areas along an open slope for different mountain tracks.

Based on these findings, the groundwater regime was modelled for different zones of the hillslope, i.e. near and adjacent to the bends, considering typical in-situ conditions and the stratigraphical schemes in figure 3. Pore water pressures induced by the 4-5 May 1998 rainfall storm were computed through the Seep/W Finite Element code (Geoslope, 2004) using the computational schemes of figure 5 that

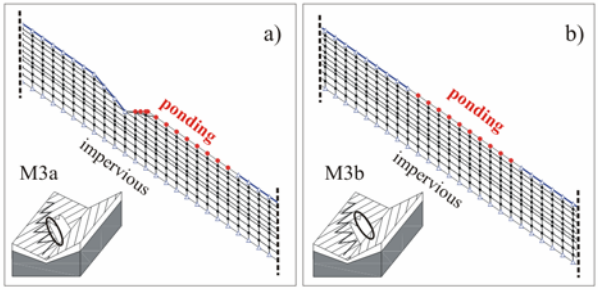


Figure 5. Computational schemes corresponding to different portions of the hillslope (schemes M3a and M3b of figure 4c).

reproduce the schemes M3a and M3b of figures 1 and 4. Concerning scheme M3a, the rainfall values provided by Cascini et al. [2003] were assumed as boundary condition at the ground surface while a local ponding was considered downslope of the bend. For scheme M3b (Fig. 5b), the same ponding condition was assumed taking into account the possible lateral redistribution of surface runoff highlighted in figure 4b. For scheme M3a, the obtained results highlight that the ponding condition assumed at the bend of the track and downslope produce different effects depending on the local stratigraphy (Fig. 6). Particularly, if only ashy B soils are present (scheme 1 of Fig. 6) a local and slight increase of pore water pressures can be simulated; strong variations are conversely computed due to presence of both ashy A and B soils (scheme 2 of Fig. 6). Finally, for slopes with continuous pumice soil layers (scheme 3 of Fig. 6), pore water pressures increase involving larger portions of the slopes than in the previous cases. Similar results are obtained for scheme M3b (i.e. corresponding to zones adjacent to the bends), thus highlighting that the groundwater regime is primarily controlled by both the assumed boundary conditions at the ground surface and the stratigraphy rather than by local geometrical discontinuities.

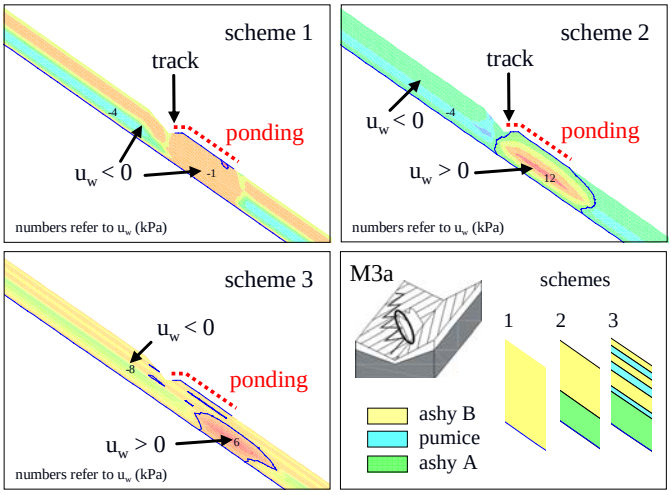


Figure 6. Pore water pressure induced by rainfall and a local ponding condition for the scheme M3a of figure 5a with different stratigraphies.

3.2 Slope stability analyses

For the scheme M3a, limit equilibrium analyses (Fig. 7) were performed through the Slope/W code (Geoslope, 2004) assuming as input the computed pore water pressures. The obtained results show that for scheme 1 of figure 6 (only ashy B soils) a local ponding condition does not cause any failure, as the slope safety factor (FS) is greater than one (Fig. 7a). Conversely, for the scheme 2 of figure 6 (ashy A and B soils) failure is simulated along slip surfaces involving zones both downslope and upslope the cut slope (Fig. 7b). Similar results are obtained for the scheme 3 of figure 6 with a slip surface involving both the upslope and downslope zones of the track.

Similar limit equilibrium analyses were performed for scheme M3b. Figure 8 provides the results obtained, at the same time instant of figure 7, for different stratigraphical settings assuming an equal local ponding condition. Particularly, for scheme 1 (only ashy B soils) failure conditions are not simulated (Fig. 8a); conversely these conditions involve different zones of the slope in the case of scheme 2 (only ashy A

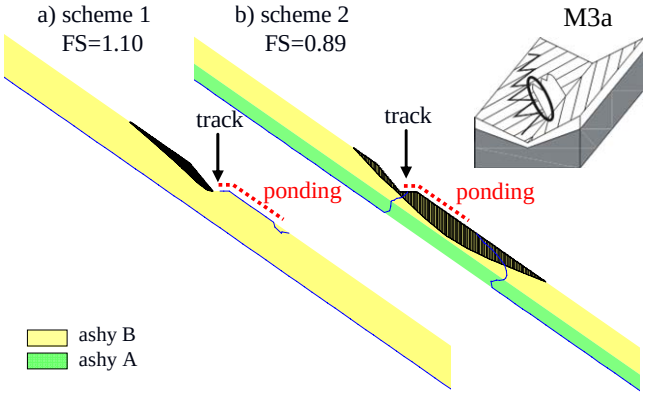


Figure 7. Limit equilibrium analyses for the scheme M3a of figure 5a.

soils) and scheme 3 (ashy A and B and pumice soils) (Fig. 8c). The obtained results clearly outline the role played by the stratigraphy on the failure onset. Moreover, comparing scheme 2 of figure 7 with scheme 2 of figure 8 it derives that the slope safety factors are about the same (respectively FS=0.89 and FS=0.88) independent of the presence of a geometrical discontinuity inside the pyroclastic deposit.

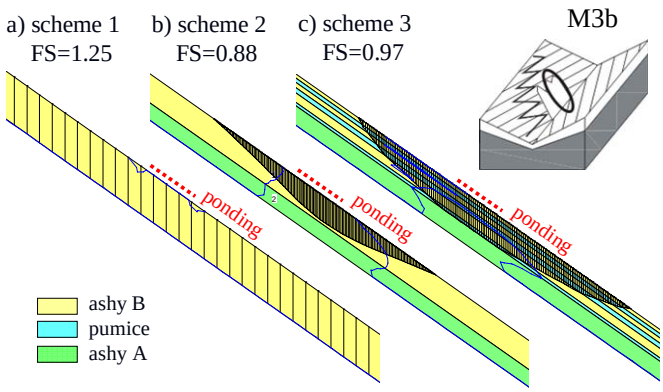


Figure 8. Limit equilibrium analyses for zones of hillslope far away from bends of tracks.

To validate the previous results, pore water pressures were used as input for uncoupled and coupled stress-strain analyses. For scheme M3b, Figure 9 shows the displacements computed for an homogeneous soil deposit (scheme 1 of Fig. 6). It is worth noting that in the case of rainfall only as boundary condition, the failure onset is not simulated. On the contrary, including a ponding condition, the simulated displacements are characterised by quite high gradients leading to failure conditions after about 24 hours (Fig. 9).

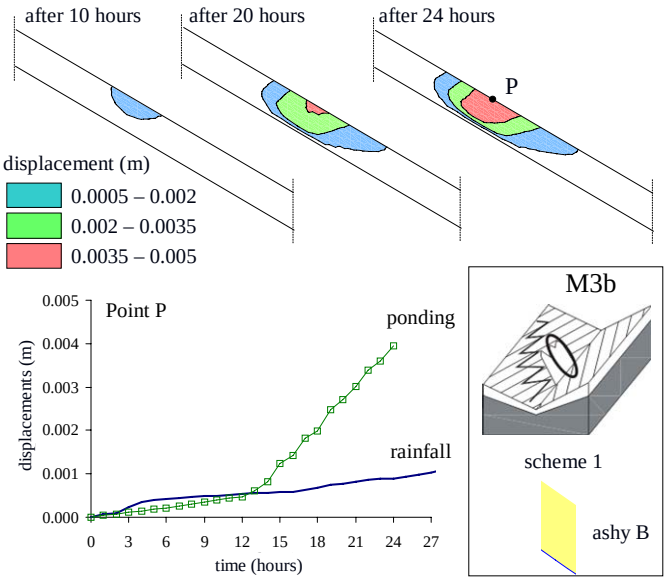


Figure 9. Stress-strain analyses for zones far away from the bends (scheme 1 of Figure 8).

4. CONCLUSION

Anthropogenic factors often influence the superficial groundwater circulation during rainfall. Particularly, mountain tracks and cut slopes cause erosion processes and runoff that can locally trigger slope instability phenomena. Sometimes, multiple failures develop inside large areas according to complex mechanisms not easy to recognise and/or to model.

Examples of significant landslide source areas related to anthropogenic factors were recorded in the Campania region (Southern Italy) in May 1998, when several and multiple landslides affected shallow unsaturated pyroclastic deposits covering a carbonate massif. Considering the complexity of the occurred phenomena, few contributions in the scientific literature address the topic. In order to obtain a preliminary insight in the role played by different factors on the triggering mechanisms, groundwater and slope instability analyses were performed, in a parametric form, referring to simplified and bi-dimensional schemes of the landslide source areas.

Groundwater modelling highlights that ponding conditions strongly affect the transient

pore water pressures, mostly depending on the local stratigraphical setting. Limit equilibrium and stress-strain analyses outline that geometrical discontinuities act as an important aggravating factor for slope stability conditions. However, stratigraphy is confirmed as the key factors for the triggering of large instability phenomena for all the landslide source areas. Consequently, more advanced numerical models than those used in the present paper should take stratigraphy into account together with more elaborate methodologies such as, for instance, the three-dimensional analysis of slope stability conditions.

REFERENCES

Alonso, E., Gens, A., Lloret, A., Delage, C., Effect of rain infiltration on the stability of slopes. Proc. 1th Int. Conference on Unsaturated Soils, Alonso&Delage eds, Paris, 241-249, 1995.

Bilotta, E., Cascini, L., Foresta, V., Sorbino, G., Geotechnical characterization of pyroclastic soils involved in huge flowslides. *Geotechnical and Geological Engineering, 23*, 365-402, 2005.

Calcaterra, D., Parise, M., Palma, B., Combining historical and geological data for the assessment of the landslide hazard: a case study from Campania, Italy. *Natural Hazard and Earth System Sciences, 3*, 13-16, 2003.

Cascini, L., The flowslides of May 1998 in the Campania region, Italy: the scientific emergency management. *Italian Geotechnical Journal, 2*, 11-44, 2004.

Cascini, L., Cuomo, S., Sorbino, G., Flow-like mass movements in pyroclastic soils: remarks on the modelling of triggering mechanisms. *Italian Geotechnical Journal, 4*, 11-31, 2005.

Cascini, L., Cuomo, S., Guida, D., Typical source areas of May 1998 flow-like mass movements in the Campania region, Southern Italy, *Engineering Geology, 96*, 107-125, 2008.

Cascini, L., Guida, D., Nocera, N., Romanzi, G., Sorbino, G., A preliminary model for the landslides of May 1998 in Campania Region. *Proc 2nd Int. Symposium on Geotechnics of Hard Soil-Soft Rock – Napoli, Balkema, 3*, 1623-1649, 2000.

Cascini, L., Sorbino, G., Soil suction measurement over large areas: a case study. *Proc. 3rd Int. Conference on Unsaturated Soils, Recife (Brasil), Balkema 2*, 829-834, 2002.

Cascini, L., Sorbino, G., Cuomo, S., Modelling of flowslides triggering in pyroclastic soils. *Proc. Int. Conference on "Fast Slope Movements - Prediction and Prevention for Risk Mitigation", Napoli, Patron Ed. 1*, 93-100, 2003.

Croke,J., Mockler, S., Gully initiation and road-to-stream linkage in a forested catchment, Southeastern Australia. *Earth Surface Processes and Landforms, 26(2)*, 205-217, 2001.

Crosta, G.B., Dal Negro, P., Observations and modelling of soil slip-debris flow initiation processes in pyroclastic deposits: the Sarno 1998 event. *Natural Hazards and Earth System Science, 3*, 53-69, 2003.

Cuomo, S., Geomechanical modelling of triggering mechanisms for flow-like mass movements in pyroclastic soils. *PhD dissertation at the University of Salerno, Italy*, pp. 274, 2006.

Di Crescenzo, G., Santo, A., Debris slides–rapid earth flows in the carbonate massifs of the Campania region (Southern Italy): morphological and morphometric data for evaluating triggering susceptibility. *Geomorphology, 66*, 255-276, 2005.

Fiorillo, F., Wilson, R.C., Rainfall induced debris flows in pyroclastic deposits, Campania (Southern Italy). *Engineering Geology, 75*, 263- 289, 2004.

Fredlund, D.G., Morgenstern, N.R., Widger, R.A., The shear strength of unsaturated soils. *Canadian Geotechnical Journal, 15*, 313-321, 1978.

Gasmo, M., Rahardjo, H., Leong, E.C., Infiltration effects on stability of a residual soil slope. *Computers and Geotechnics 26 (2)*, 145-165, 2000.

Geoslope, User's guide. GeoStudio 2004, Version 6.13. Geo-Slope Int. Ltd. Calgary, Canada, 2004.

Guadagno, F.M., Martino, S., Scarascia Mugnozza, G., Influence of man-made cuts on the stability of pyroclastic covers (Campania-Southern Italy): a numerical modelling approach. *Environmental Geology, 43*, 371-384, 2003.

Guadagno, F.M., Forte, R., Revellino, P., Fiorillo, F., Focareta, M., Some aspects of the initiation of debris avalanches in the Campania Region: the role of morphological slope discontinuities and the development of failure. *Geomorphology, 66,* 237-254, 2005.

Luce, C.H., Black, T.A., Spatial and temporal patterns in erosion from forest roads. Influence of Urban and Forest Land Uses on the Hydrologic-Geomorphic Responses of Watersheds. *M.S. Wigmosta and S.J. Burges, Editors. American Geophysical Union, Washington, D.C.*, 165-178, 2001.

MacDonald, L.H., Sampson R.W., Anderson, D.M., Runoff and road erosion at the plot and road segment scales, St John, US Virgin Islands. *Earth Surface Processes and Landforms, 26*, 251-272, 2001.

Montgomery, DR, Dietrich, WE., A physically based model for the topographic control on shallow landsliding. *Water Resources Research,* 30(4), 1153-1171, 1994.

Pack, R.T., Tarboton, D.G., Goodwin, C.N., The SINMAP Approach to Terrain Stability Mapping. *8th Congress of the International Association of Engineering Geology, Vancouver, British Columbia, Canada*, 1998.

Rahardjo, H., Li, X.W., Toll, D.G., Leong, E.C., The effect of antecedent rainfall on slope stability. *Geotechnical and geological Engineering [19]*, 371-399, 2001.

Richards, L.A., Capillary conductions of liquids through porous mediums. *Journal of Applied Physics* 1, 318-333, 1931.

Savage W.Z., Godt J.W., Baum R.L., A model for spatially and temporally distributed shallow landslide initiation by rainfall infiltration. in press. Proc. 3rd Int. Conf. on Debris Flow Hazards Mitigation: Mechanics, Prediction, and Assessment, Davos, Switzerland, 179-187, 2003.

Sidle, R., Sasaki, S., Otsuki, M., Noguchi, S., Rahim, A., Sediment pathways in a tropical forest: effects of logging roads and skid trails. *Hydrological Processes, 18*, 703-720, 2004.

Sorbino,G., Foresta,V., Unsaturated hydraulic characteristics of pyroclastic soils. *Proc. 3rd International Conference on Unsaturated Soils, Recife (Brasil), Balkema, 1*, 405-410, 2002.

Tsaparas, I., Rahardjo, H., Toll, D.G., Leong, E.C., Controlling parameters for rainfall-induced landslides, *Computers and Geotechnics, 1,* 1-27, 2002.

Tucker, G.E., Bras, R.L., *Hillslope processes, drainage density, and landscape morphology. Water Resources Research, 34(10),* 2751-2764, 1998.

Ziegler, A.D., Giambelluca, T.W., Sutherland, R., Nullet, M., Yarnasarn, S., Pinthong, J., Preechapanya, P., Jaiaree, S., Toward understanding the cumulative impacts of roads in upland agricultural watersheds of northern Thailand. *Agriculture, Ecosystems and Environment [104],* 145-158, 2004.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Additional torrent control strategies on debris flow alluvial fans with extremely high vulnerable settlements

B. Mazzorana[a] **C. Scherer**[b] **N. Marangoni**[a]
[a] *Department of Hydraulic Engineering, 39100 Bolzano, Cesare Battisti Str.23*
(bruno.mazzorana@provincia.bz.it; nicola.marangoni@provincia.bz.it)
[b] *Hydraulic consultant (christian.scherer@libero.it; info@obrist-partner.it))*

Abstract: Recent extreme debris flow and hyper-concentrated flow events in the Alpine Regions have shown on the one hand the enormous destructive impact power on buildings, bridges and roads, and on the other hand, the limits of "classical" torrent control measures. Due to reliability constrained efficiency of consolidation protection measures with unsatisfactory performance in case of extreme events the sediment load volume is higher than expected due to remobilization phenomena. In many cases check dams with sediment retention or sorting function don't function properly and the openings are immediately obstructed. The results of an extended backward and forward oriented investigation clearly suggest a change of paradigm for the torrent control strategy at least where the possibility for hazard reduction is extremely limited, if pursuit with the construction of classical torrent control measures. This work suggests enlarging the toolkit for torrent control by going beyond "object protection" for buildings on the debris flow phenomena affected alluvial fans. The number of endangered objects located in the red, blue and yellow hazard zone has to be reduced to a minimum. Here we propose, starting from an in-depth understanding of the involved processes and overland flow dynamics, the construction of protection structures, the specific reinforcement and where necessary a functional reshaping of building's boundaries, terrain breaklines, and roads in order to obtain a maximum damage reduction diverting the flow towards sectors where the value at risk is lower. The goal is to identify the set of construction elements that minimizes the expected damage. It's an iterative procedure that involves a series of numerical simulations corresponding to each particular protection configuration. A simplifying graph theoretic schematization of the overland flow dynamics permits to identify suitable locations for the construction of these protection elements and to reduce the computational efforts reducing the number of simulations otherwise required. The procedure has been applied to a case study.

***Keywords**:* debris flow, natural hazards, object protection, risk analysis, torrent control.

Introduction

An in-depth analysis in form of a backward oriented investigation of the debris flow and hyper concentrated flow events, that have been documented since 1998 and which data have been made available in a proper database, shows on the one hand a series of significant results concerning the relationship between reliability of “classical”, in stream located, torrent control protection measures and event’s magnitude and their functional efficiency and the damage exposure of the objects at risk on alluvial fans on the other hand. Significant performance deviations have been registered for each of the following characteristics of an ideal protection system (Mazzorana, 2007):

- long durability (high reliability), easy and cheap maintainability
- high functionality (efficiency) with substantial mitigation effects for high return period events and a proper mitigation effect for low return period events
- high sediment transport regulation capacity with progressive reduction of the remaining sediment yield potential

- Low uncertainties about protection system responses to extreme events, easy integration and more effective implementation of early warning systems
- Resilient response to extreme loadings (beyond design return period intensities)

Once the peak discharge overloads the capacity of the alluvial fan channel cross sections, overland flow is triggered by the debris flow surplus volume. The objects at risk in most of the cases are directly exposed to the debris flow impact forces. In recent years first steps toward the ideation and implementation of integral constructive risk reduction approaches have been made. Here we report just some examples. Egli (1996, 2005) gives a complete overview of effective object protection techniques for mass movements, water related hazards and avalanches. Willi *et al.* (2006) explains how differentiated flood protection has been implemented at the Engelberger Aa. In this case the risk reduction for the settlements is realized from a technical perspective by a controlled dike failure that diverts lateral overflow towards floodplain areas where the value at risk is lower and where the flood effects can be controlled. Dike reinforcements and channel geometry adaptations complete the set of technical measures that are integrated with early warning and rescue action plans. Scherer *et al.* (2005) proposes a flood impact reduction concept for the city of Vipiteno in South Tyrol, where the city, in case of inundation, is protected by so called primary and secondary defence lines made up of iterative and temporary flood protection elements. Here we propose a stepwise problem solving approach that, starting from system response analysis (backward and forward oriented investigation) and a graph theory based representation of the dynamic behaviour of debris flow processes on alluvial fans, provides the framework for the identification of the set of construction elements that minimizes the expected damage. First an overview of the problem solving approach is given, then the graph theoretic elements that allow the simplified representation of flow phenomena on alluvial fans are presented and applied to practical cases and finally it is shown for one case study how to find the set of construction elements that maximizes the risk reduction.

Overview of the problem solving approach

The following list gives an overview of the problem solving approach:

1) The first step of the procedure is the system response analysis through backward and forward oriented investigation of the debris flow behaviour on the alluvial fan. The goal is to provide a wide range of scenarios in order to "capture" at least to a certain extent the variability of deposition phenomena as a results of debris flows of different magnitude, different rheological behaviours, unforeseen cross-sectional obstructions (e.g. due to woody debris transport) or other momentary "configurations" that cause unexpected diversions of the flow. The backward oriented investigation can provide, depending on the quality of the available data, a first idea of possible flow patterns and preferential flow directions. If deposition heights have been recorded, the deposition volume can be estimated. It has to be checked, on the basis of a geomorphological investigation of the alluvial fan, if other flow patterns could have existed in the past and could potentially be reactivated in case of extreme events. The forward oriented investigation consists in performing a series of simulations with a 2D debris flow propagation model as follows:

a) Debris flow simulation on the alluvial fan for events of relevant return period (e.g. recommendations for the production of hazard zoning maps) under the hypothesis that the "channel geometry" on the alluvial fan is not reduced and therefore available for propagation.
b) Under the same loading conditions of the system, assume that, due to in stream depositional phenomena, the "channel geometry" is not available for flow propagation.
c) Under the same loading conditions assume that the "channel geometry" is available for flow propagation except at bridges, culverts and abrupt channel restrictions.
d) Take into consideration possible climate change scenarios. According to the research results in this field increase the loading conditions of the system (duration and maximum discharge peak of the debris flow).

As a synthesis of the backward and forward investigation, the maximum flow depths and the maximum velocity maps are produced. These maps are called maximum

intensity propagation maps. Also the simulation results for each output time step interval are taken into consideration.
2) The next step consists in an overlay of the maximum intensity propagation map and the vulnerability map in order to obtain the "specific risk" map.
3) On the basis of the maps produced at points 1 and 2, the simplified graph theoretic representation of the debris flow dynamics follows.
4) Through heuristic engineering judgement the protection strategy is defined by determining the location and the type of additional risk mitigating object protection measures that "control" the flow and/or modify the flow propagation patterns.
5) According to the "selected" protection strategy the digital terrain model is modified and a new series of simulations is performed in full analogy with the forward oriented investigation described at point 1. A new risk evaluation follows. While the new maximum propagation map gives investigation to further the protection strategies, the comparison between the original risk situation and the new one shows the obtained risk reduction.
6) Steps 3 to 5 have to be repeated until the maximum risk reduction is obtained. The whole procedure is shown in the following flow chart (see figure 1).

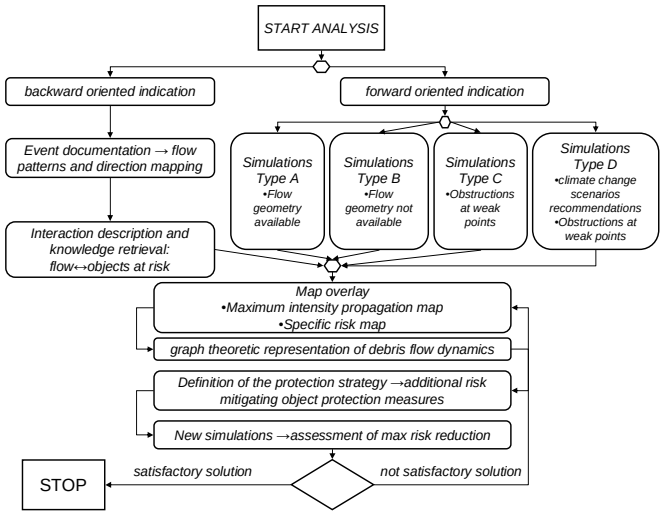


Figure 1. Flow chart of the problem solving approach

Representation of debris flow dynamics on alluvial fans: a graph theoretic approach

In this work graph theory is used to represent debris flow dynamics on alluvial fans (Ottmann *et al*, 2002). A digraph $G = \{V, E\}$ consists of a set V = {1, 2… |V|} of vertices and a set $E \subseteq V \times V$ of edges or arcs. A pair $(v, v') \in E$ is called arc from v (start vertex) to v' (end vertex); v and v' are adjacent and v (and also v') is incident with e. The vertices are represented by points and the edges by arrows with the arrow head toward the end vertex. We choose to store the graph objects in an adjacency list. The structure is shown in the next figure (left). For the scope of this work we identify a general graph structure (figure 2 on the right side) that combined correctly can reproduce, to an allowable degree of simplification, the "macroscopic" flow characteristics on the alluvial fan. This fundamental sub graph and its modifications consist of 1, 2… n input vertices and arcs, 2 intermediary vertices and arcs and 0, 1, 2... m output arcs and vertices. The second intermediary vertex, from which the output arcs depart, is the abstraction of any feature on the alluvial fan that influences significantly the flow. The arcs are labelled in relation to the specific risk. This fundamental sub graph can be used also to model in a simplified fashion the function of sediment dosing check dams.

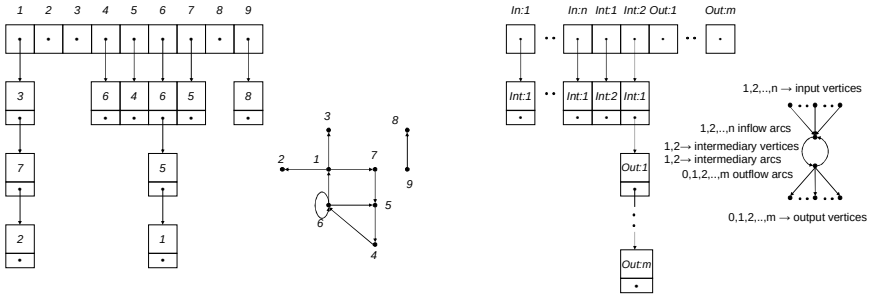


Figure 2. Graph adjacency list representation of the (left) and basic graph component for the description of the flow characteristics (right)

Similarly a channel tract with a partially obstructed bridge cross section, overflowing phenomena at the right and left side of the channel and backwater effects upstream of the bridge can be modelled (see figure 3).

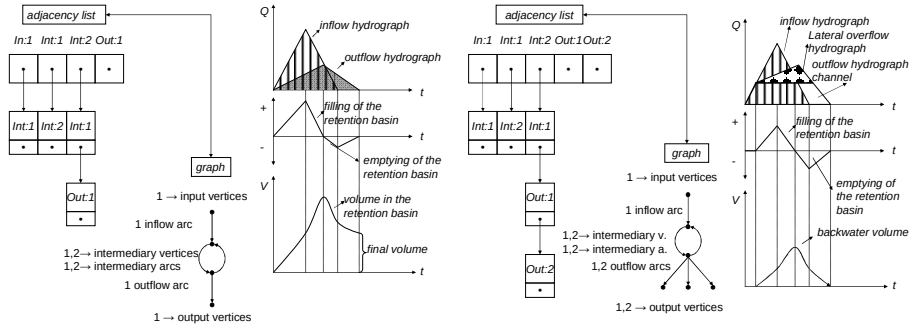


Figure 3. Graph representation of a sediment dosing check dam (left) and of a channel with obstruction and lateral overflow (right)

Given the maximum intensity propagation map –MIPM- and the specific risk map –SRM-, the relevant flow trajectories can be identified and the specific risk vertices –SRV-, as well as the flow changing vertices –FCV-, can be defined. Interpreting the flow patterns, taking into consideration both MIPM and SRM, the propagation graph –PG- is constructed and defined as follows: $PG = \{SRV \cup FCV, FA\}$, where FA is the set of flow arcs. The next steps consist in constructing the adjacency list of the digraph and in assigning the risk costs –RC- to the flow arcs –FA. The assigned risk costs derive from a risk analysis that takes into account the intensities of the process, the vulnerability of the damage exposed objects. The vertices are coloured according to the risk exposure, the arcs are labelled with the corresponding specific risk costs. The sources (roots) and the exit vertices are identified. A set of relevant risk objects to be protected are selected from SRV. The following paths are calculated: 1) maximum risk cost paths to exit vertices – MaxRCPtoEV- starting from root vertices, 2) minimum risk cost paths to exit vertices – MinRCPtoEV- starting from root vertices, 3) risk cost paths to the selected SRVs. The calculated paths at points 1, 2 and 3 are evidenced in the digraph.

A set of heuristics (principles) –P- can be used alone, or in combination, to modify the digraph in order to obtain a satisfactory risk reduction (Zobel, 2006). For some principle also the complementary principle –CP- can be applied in specific situations. Here a list of 7 principles is proposed:

1) P: Decomposition (diversion, deviation) ↔ CP: Composition (union)

E.g. P: deviation (collection) of MaxRCPtoEV towards MinRCPtoEV or towards low impact sectors

2) P: Separation (detachment) ↔ CP: Addition

E.g. P: Let $ROV \subseteq SRV$ be a set of relevant objects to be protected, then through a separation measure, part of the arcs – FA- that end in the vertices – ROV- are eliminated.
3) P: Combination
E.g. P: The implemented protection strategy is "inspired" by different protection principles. The problem is tackled from different perspectives
4) P: Multipurpose use ↔ CP: optimal single function performance
E.g. P: Protection measure obtained by landscape reshaping do not drastically interfere with agricultural practices when the system is not interested by natural hazard phenomena
5) P: Shortest path (minimum risk path)
E.g. P: MinRCPtoEV should be systematically favoured and the discharge should be conveyed toward it (adaptations of cross sectional geometry)
6) P: Smoothing in space and time (dosing, retention) ↔ CP: Controlled overloading
E.g. P: This is one of the most important impact reduction principles that should be considered throughout the whole protection strategy, e.g. favouring "natural" deposition phenomena in low risk sectors (Armanini, 2001)
E.g. CP: Controlled dike failure at a predefined location to unload the rest of the system
7) P: Substitution (Integration) of expensive permanent protection components with temporary protection components

Case study: Weissenbacher torrent

The Weissenbacher torrent basin (0.65 km²) is located in the Aurino basin in the eastern part of the Autonomous Province of Bolzano. Concerning the backward oriented investigation a relevant debris flow event happened in the Weissenbacher torrent at 13th August 2002. Concentrated rainfall was about 40 mm in 50 min coming from S-E direction. The debris flow event was triggered in the starting zone 5 smaller landslides (approx. 3500 m³). In the transit zone the debris flow caused deep and lateral erosion phenomena (approx. 7200 m³). The total debris flow volume that was approaching the alluvial fan apex was approx. 9500 m³, while approx. 1200 m³ were deposited along the incised channel. At the alluvial fan apex a consistent of the debris flow volume (coarser material) was deposited, while the remaining part propagated in five overland flow limbs and along the channel on the alluvial fan. Concerning in stream structures one culvert was completely destroyed, while another one was clogged. Several roads and approx. 4 ha agricultural land have been covered by the debris flow material and had to be cleared. Only four houses were damaged, but not severely. The next figure gives an overview of the event. Concerning the forward oriented investigation, first a hydrological analysis has been performed with the HEC-HMS model in order to obtain the hydrograph. The peak discharge of the debris flow was estimated using Takahashi's formula (Takahashi, 1991). Different scenarios have been simulated with the propagation model FLO-2D. Figure 4 shows the simulation results.

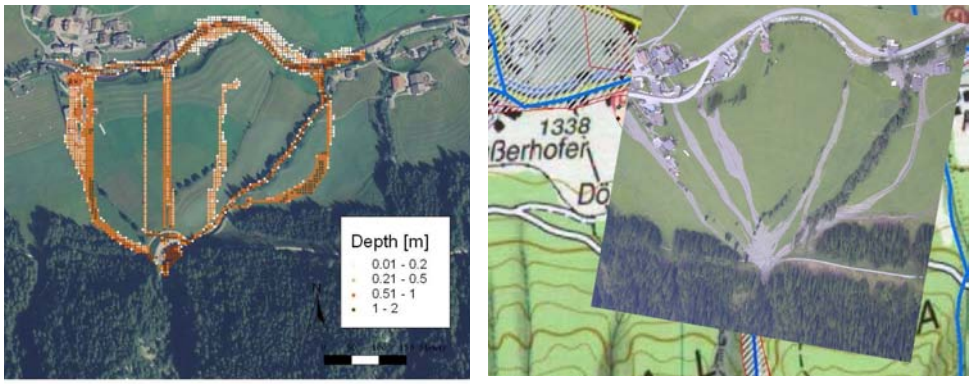


Figure 4. Left: simulation (depth - velocity) right: Event 13th August 2002

The results of the backward oriented investigation and forward oriented investigation (maximum system response) have been used to construct the propagation graph.

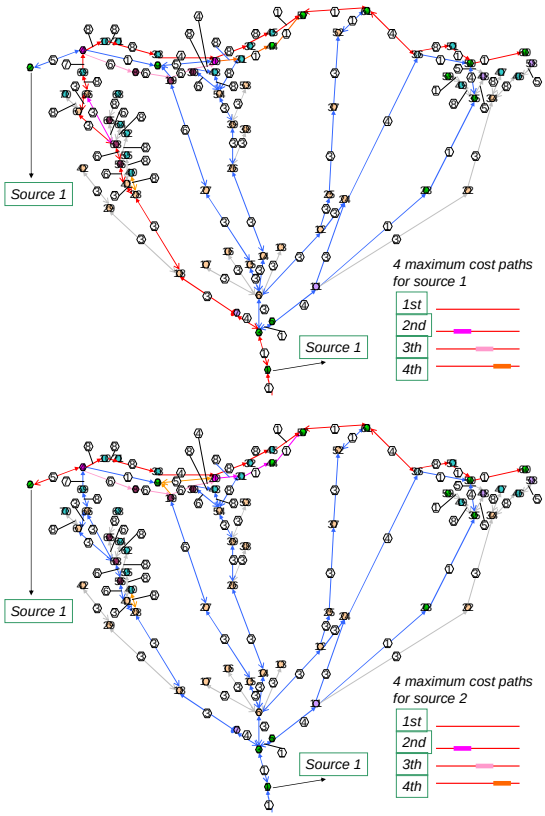


Figure 5. Maximum risk cost paths for source 1 and 2 respectively;

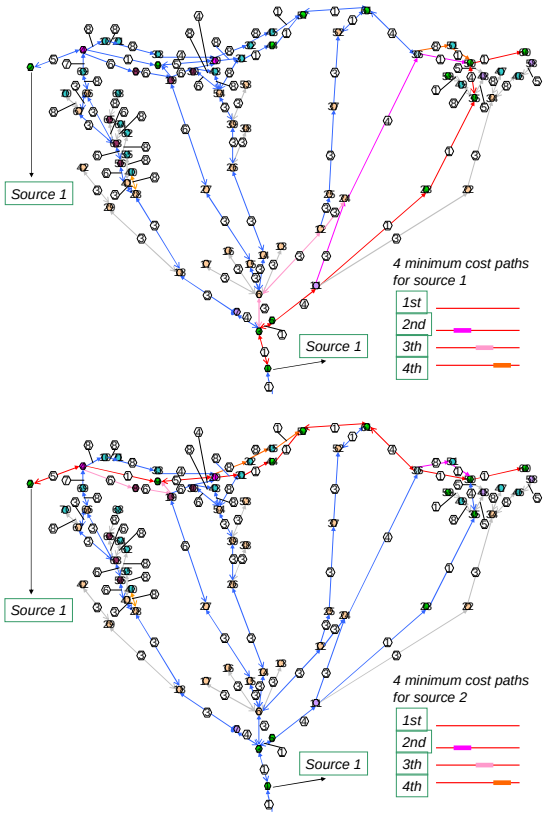


Figure 6. Minimum risk cost paths for source 1 and 2 respectively

Risk categories have been assigned to each vertex and the arcs have been labelled with the risk costs. The minimum and maximum risk cost paths for each root (source) of the graph have been determined. The results are shown in the figures 5 and 6.
The results of the methodology applied to the test case suggest the locations where it's efficient to intervene with technical measures and the heuristics indicate possible solution principles for risk mitigation problems. The definite choice and design of the technical measures that implement the solution strategies relies on expert knowledge. Consequently the heuristics have been applied and the protection strategy shown in the next figure has been developed. It's a combination of deviation and object protection measures in addition to measures that favour debris flow deposition phenomena in "low risk" alluvial fan sectors. The deviation protection measures consist in dam construction in combination of local reshaping and remodelling of the channel geometry. The protection strategy (see figure 7) includes also the removal of "weak points" along the torrents.

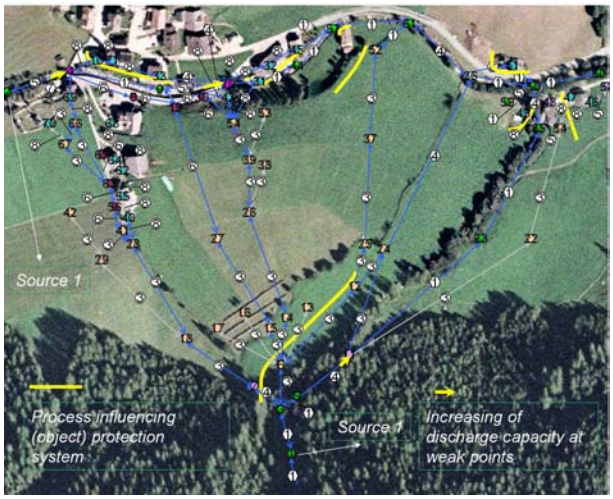


Figure 7. The proposed protection strategy

A series of additional simulations have been performed in order to check the risk reduction effect of the proposed protection strategy. The results are shown in the figure 8.

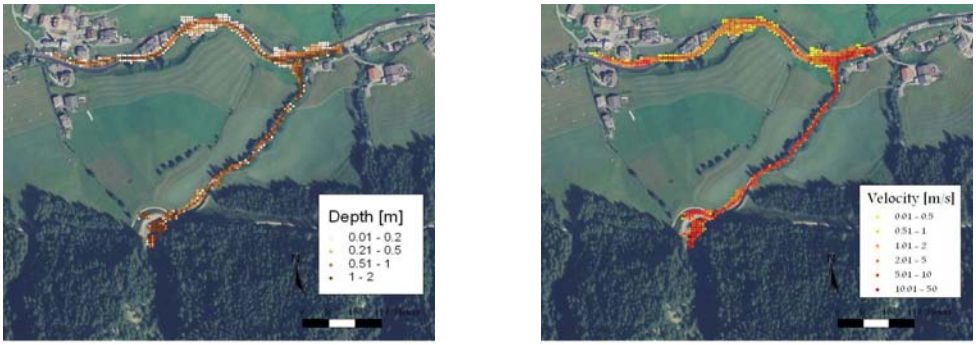


Figure 8. Simulation with protection strategy

Conclusions

In this work we proposed a procedure that structures the relevant information and knowledge from backward and forward investigation by means of the system response analysis resulting in a graph theoretic representation of flow patterns and object – flow interactions. On this basis a risk analysis can be methodically performed. In addition we proposed a set of heuristics that support the design of protection measures that efficiently reduce the overall risk exposure. The application of the case study -Weissenbach torrent- has shown that the procedure can successfully be applied to the investigation of "multi-source" debris flow related hazards in vulnerable regions. The calculation of maximum and minimum risk cost path as well as the risk cost path for the relevant objects is a crucial step

for the definition of the protection strategies that become comparable. The decision making process takes advantage from the availability of these indicators. Further research activities in this field will focus on the applicability of this procedure to flood related risk assessment. An automatic graph generation algorithm, that has to be added to complete the procedure, could facilitate the analysis of larger scale phenomena.
Hazard mitigation and increased risk reduction through strategically located control elements is an interdisciplinary task. The protection elements have to be hydraulically effective and suitable from a landscape-architectural view. Last but not least, the affected population has to be involved from the first planning steps with a consensus building approach. A transparent communication of hazard, risk and the inherent limits of "classical" protection measures is absolutely necessary in order to prepare the ground for a general acceptance of these additional protection strategies.

REFERENCES

Armanini, A., and M. Larcher, Rational Criterion for Designing Opening of Slit-Check Dam, *ASCE Journal of Hydraulic Engineering (February 2001)* Vol. 127, Nr.2, 2001.

Armanini, A. et al., Restoration of the basin Quebrada San Josè de Galipán, *International Workshop on the debris flow disaster of December 1999 in Venezuela*, 1999.

Egli, T., Hochwasserschutz und Raumplanung, *ORL Bericht 100/1996. vdf Hochschulverlag ETH Zürich,* 1996.

Egli, T., Objektschutz gegen gravitative Naturgefahren, *Kantonale Gebäudeversicherungen Vereinigung Kantonale Feuerversicherungen, VKF AEAI (Hrsg.), Bern*, 2005.

Mazzorana, B., Efficiency of protection systems in alpine torrents from a reliability, sustainability and mainainability perspective, *Taiwan, Slope land Disaster Mitigation Conference*, 2007.

Ottmann, T. and P. Widmayer, Algorithmen und Datenstrukturen, *Spektrum, Akad. Verl., Heidelberg; Berlin*, 2002.

Scherer, C., and B. Mazzorana, Bericht zur Hydraulik, *Interreg 3B Project „River Basin Agenda", Bolzano*, 2005.

Takahashi, T., Debris flow, *IAHR Monograph. Rotterdam, Balkema,* 1991.

Willi, H. P. and J. Eberli, Differenzierter Hochwasserschutz an der Engelberger Aa, *TEC21, Fachzeitschrift für Architektur, Ingenieurwesen und Umwelt,* 2006.

Zobel, D., Systematisches Erfinden, *expert-verlag, Renningen-Malmsheim,* 2006.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Modelling of the Risks of the Sidoarjo Mud Flow, Indonesia

S. Kivva and M.Zheleznyak

Ukrainian Center of Environmental and Water Projects, Prospect Glushkova,42, Kiev, 03187, Ukraine (slk/mark@ucewp.kiev.ua)

Abstract: The Sidoarjo mud flow flooded the part of the Porong area, Indonesia, surrounded by the set of the quickly constructed dams. The objective of the study was quantification of the risks imposed by the mud volcano towards the ECCO Indonesia factory site due to dam breaks and dam overtopping. The distance between the northern dike surrounding the volcano and the PT ECCO is at 4.3 km. This territory is crossed by many small rivers. The two- dimensional numerical model COASTOX_MUD has been developed for modeling of the variable density mud flow. The model is based on the system of four nonlinear partial differential equations using two components description of mud flow as non-Newtonian fluid. The model also has a simplified version which considers mud flow as one-component fluid with an averaged density , i.e. it is based on the system of three shallow water equations used by Laigle and Coussot [1997] and in other previous studies. The two component version of the model has supplementary benefits compared to single component models as it can simulate mud dilution induced by the rivers as they are traversed by the mud flow. The two component version of the COASTOX_MUD was used in the study by incorporating the specific weight of the water–sediment mixture and mud viscosity received from on-site sampling and laboratory measurements. The simulation for the dam breaks scenarios and the mud flow conditions of the summer 2007 shows insignificant risk of the factory site flooding by the mud flow due to the dam breaks.

***Keywords**:* Sidoarjo Mud Volcano; Mud Flow, Modelling, Hazards Mapping.

1. INTRODUCTION

The Sidoarjo mud flow, also named "Lusi", is an ongoing eruption of gas and mud from the earth in the sub district of Porong, Sidoarjo in East Java, Indonesia [Cyranoski D., 2007; Wikipedia, 2008]. The mud started to flow on May 28, 2006 and so far the mud has inundated parts of the Porong area. It is considered to be a mud volcano. So far, all efforts to stop the flow have failed. Some 50,000 m³ of hot mud were erupting every day as of August; in September 2006, the amount increased to some 125,000 cubic meters daily. By early September about 25 square km had been flooded by mud. A network of dams and barriers has been erected to contain the mud. The volcano site is located at the distance at 6 km, from the territory of the PT ECCO Indonesia factory site. The continuation of mud release was assessed as potentially long term process, - temporal scale - years [UN OCHA, 2006; Wikipedia, 2008] therefore the relevant risks of the mudflow for ECCO Indonesia factory site located in the vicinity of mud volcano was studied in 2007 by the team of Singaporean, Indonesian and Ukrainian scientists. The paper presents the methodology and the results of the risk assessments based of the modelling of the set of the scenarios of the mud flow propagation towards the PT ECCO Indonesia. The objective of the study was quantification of the risks imposed by the mud volcano towards the ECCO Indonesia factory site due to dam breaks and dam overtopping.

The previous simulations of the mud flow propagation for different scenarios including dam breaks and overtopping were provided on the basis of numerical one dimensional and two dimensional (depth integrated) models [O'Brien et al, 1993; Laigle and Coussot, 1997;

Ming and Fread, 1999; Pudasaini and Hutter, 2007]. The structural difference between the 2D Mud Flow Equations (MFE) and Nonlinear Shallow Water Equation (NSWE) is the introduction of a specific "resistance term" into the MFE which parameterizes mud rheology.

For the simulation of the Sidoarjo mud flow the numerical finite volume model COASTOX_MUD has been developed as the extension of the model COASTOX based on the NSWE. COASTOX model has been tested and implemented for a wide range of the water/sediment/radionuclide simulations including dam break problems for the water bodies of the vicinity of the Chernobyl Nuclear Power Plant within the and post-accidental studies described by Kivva and Zheleznyak [2001], Onishi at al.[2007].

COASTOX_MUD is new two component model based on the system of the four nonlinear partial differential equations that simulates mud flow using different fluid mud and water densities. The model also has a simplified version which considers mud flow as one-component liquid with an averaged density , i.e. it is based on the system of three 2D MFE used by Laigle and Coussot [1997] and in the other previous studies. The two component version of the model has supplementary benefits compared to single component models as it can simulate mud dilution induced by the rivers as they are traversed by the mud flow.

The two component version of the COASTOX_MUD was used in the study by incorporating the specific weight of the water–sediment mixture and mud viscosity received from on-site sampling and laboratory measurements.

2. MODEL DESCRIPTION

2.1 Mud Component Formulation

We consider the mud flow as a mixture of two components: a mud with mass fraction Ym and pure water with mass fraction Yf =1-Ym. The similar approach has been used by Herbert et al. [1988] and Oldenburg and Pruess [1995] for the modeling of variable-density groundwater flow. Assume that the volumes of fresh water and mud are additive:

$$V = V_m + V_f \tag{1}$$

where V, V_m, V_f are the volumes of mixture, mud and fresh water, respectively. The mass of mixture is given by

$$M = M_m + M_f \tag{2}$$

and the mass fractions of mud and water are given by

$$Y_m = M_m / M \tag{3}$$

$$Y_f = 1 - Y_m \tag{4}$$

Dividing (1) by (2) and substituting (3) and (4), we obtain for mixture density

$$\frac{1}{\rho} = \frac{V}{M} = \frac{1 - Y_m}{\rho_f} + \frac{Y_m}{\rho_m}$$

or

$$Y_m = \frac{\rho_m}{\rho} \frac{\rho - \rho_f}{\rho_m - \rho_f}$$

where the densities are given by $\rho_f = M_f / V_f$ and $\rho_m = M_m / V_m$.

1.2 Overland Flow

Considering the aqueous phase as a mixture of a mud and pure water, we describe the free-surface mud flow by the vertically averaged, shallow water equations for two-component mixture described in 1.1 above. These equations include continuity equations of mass fraction and two momentum equations presented as follows.

$$\frac{\partial}{\partial t}\left(h\rho Y_k\right)+\frac{\partial}{\partial x_i}\left(v_i h\rho Y_k\right)=I\rho_m Y_k(\rho_m) \tag{5}$$

$$\frac{\partial}{\partial t}(v_i h\rho)+\frac{\partial}{\partial x_j}\left(v_j\, v_i h\rho\right)+gh\rho\frac{\partial\xi}{\partial x_i}+\tau_i^{\eta}=0 \tag{6}$$

where t is the time variable (s); x_i is the spatial Cartesian coordinates (m); $h=\xi-\eta$ is the flow depth (m); v_i is the overland flow velocity in the x_i-direction (m s^{-1}); $\xi(x,t)$ is the free surface elevation (m); $\eta(x,t)$ is the bed surface elevation (m); g is the acceleration of the gravity (m s^{-2}); τ^{η} is the shear stress at the bed surface (kg m^{-1} s^{-2}); I is the flow rate of mud eruption (m s^{-1}) ; ρ is the fluid density; k=1,2 is the index of mixture component. The mass fractions Y_k are computed as a function of ρ, according to

$$Y_1(\rho)=\frac{\rho_f}{\rho}\frac{\rho_m-\rho}{\rho_m-\rho_f},\qquad Y_2(\rho)=\frac{\rho_m}{\rho}\frac{\rho-\rho_f}{\rho_m-\rho_f}$$

According to O'Brien and Julien (1993), the shear stress at the bed surface can be defined as:

$$\tau^{\eta}=\tau_y+\tau_v+\tau_t$$

where τ_y is the yield shear stress of the mud-water mixture; τ_v is the viscous shear stress; and τ_t is the turbulent-dispersive shear stress.

For turbulent flow Reynolds' stress dominates and viscous stress may be negligible. The boundary shear stress can then be approximated either by the Manning equation as

$$\tau_{t,i}=\rho g\frac{n^2}{h^{1/3}}v_i|v|$$

or by the Chezy equation as

$$\tau_{t,i}=\rho g\frac{1}{c^2}v_i|v|$$

where n is the Manning roughness coefficient (s m$^{-1/3}$) and c is the Chezy friction coefficient (m$^{1/2}$ s^{-1}).

The viscous shear stress is calculated by using the following relation

$$\tau_v=\frac{K\mu(\rho)v_i}{8h}$$

where K is non-dimensional friction parameter of the laminar flow that has the value of K=24 for a smooth bottom surface but it increases with higher roughness and irregularity of the riverbed; $\mu(\rho)$ is the dynamic viscosity of the mixture.

The yield stress τ_y and viscosity $\mu(\rho)$ increase with the sediment concentration C_V. Unless a rheological analysis of the mudflow site material is available, the following empirical formulas can be used for their calculations

$$\tau_y = \tau_y^0 C_V^{3.7}$$ [Wang, 2002]

and

$$\mu(\rho) = \mu_f (1 - C_V)^a$$ [Lee, 1969]

where $\tau_y^0 = 12000 Pa$, μ_f is the viscosity of clear water, and

$$a = -(2.5 + 1.9 C_V + 7.7 C_V^2)$$

The volumetric sediment concentration C_V is given by $C_V = (\rho - \rho_f) / (\rho_m - \rho_f)$.

3. NUMERICAL SOLUTION

The system of 2-D partial differential equations (5,) (6) is solved by the integrated finite-difference (finite volume) method [Kivva and Zheleznyak, 2001]. Transformation of the partial differential equations into algebraic form requires that the physical domain be spatially discretized into a computational domain composed of a number of non-overlapping control volumes. Each control volume surrounds a single grid point which defines the position of intrinsic property variables. The conservation equations are integrated over control volume by assuming a piece-wise profile that express the variation in the primary variable between grid points. The transient equations are solved with a fully implicit time-stepping scheme. The resulting system of algebraic equations is non-linear. The integral finite-difference method conserves quantities such as mass, momentum, and species over each control volume and the entire computational domain.

The non-linear algebraic forms of the conservation equations are converted to a linear form using a residual-based, global Newton's iteration technique. The technique generally yields quadratic convergence of the residuals with iterations, given initial estimates of the unknowns that are sufficiently close to the solution. A conjugate gradient scheme is used for the solution of the linearized algebraic form of the conservation equations.

4. TESTING OF THE NUMERICAL MODEL

The numerical model has been validated against a number of benchmark test cases widely documented in the publications. COASTOX was used for the simulation of a partial dam-break or rapid opening of a sluice gate with a nonsymmetrical breach for frictionless fluid. This test has been used by Mingham et al.[1998] and Zoppou et al., [1999] for the testing of the numerical models. The computational domain is defined by a channel 200 m long, 200 m wide with horizontal bottom. The non-symmetric breach is 75 m wide and centered at 125 m. The thickness of dam is infinitesimal. Initially, the upstream water depth is 10 m, downstream water depth is 5 m (wet bed) and 0 m (dry bed). The dam wall is then breached, and instantaneously water discharges from the higher to the lower level as a downstream-directed bore while a depression wave propagates upstream. Herein, for comparison with prior works, the computational domain is discretized by a uniform 5×5 m square mesh that exactly fits the domain. The results displaying 3D views of the water surface elevation after dam failure at time t=7.2 s for wet and dry beds (Figure 1). When there is a finite water depth downstream of the dam, there is always a sharp front moving downstream. For the dry bed case, however, there is no shock. The problem consists of a rarefaction fan, and there is a smooth transition in the water depth from the breach to the dry bed downstream of the dam.

The results of COASTOX numerical simulation are in close agreement with corresponding results presented in the above-mentioned publications.

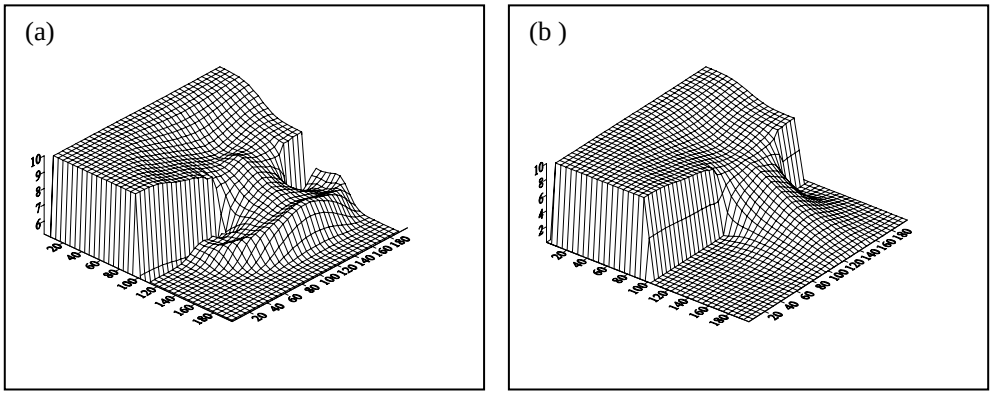


Figure 1. Solution of 2D Dam-Break Problem at 7.2 s: (a) wet bed case, (b) dry bed case.

5. NUMERICAL SIMULATIONS OF THE MUD FLOW PROPAGATION

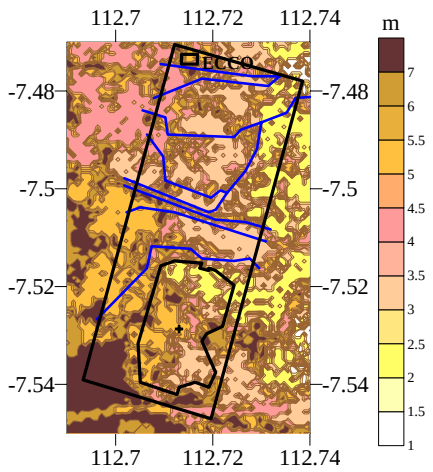


Figure 2. Topography of computational area with the depicted dams (broken black line), rivers (blue lines), center of volcano (cross), Ecco Site and the borders of the simulated area (black rectangular).

Basic information about the Digital Elevation Model in the vicinity of the mud volcano and Ecco facility was received from [USGS, 2007] as the data set with a 3 arc seconds horizontal grid dimension (Figure 2.) The resolution for the model was set at grid area of 90*90 meters. The Ecco site is located at 6 km northward from the volcano. The territory is a plain with the topographical elevation from 1 to 7 m above sea level that has a slight slope in North East direction (average slope is at 0.001). The territory is crossed by 8 small rivers flowing mainly from West to East with the discharges from 0.05 to 7 cub.m/sec. The distance between the northern dam surrounding the volcano and the PT ECCO is at 4.3 km. The black rectangular is the border of the simulated domain that will be presented on the following figures (from Figure 3 below) turned by such way that left and right borders of the simulated plot will be vertical. The black broken line inside the borders of the simulated area presents the bund around the volcano

The computational domain for the COASTOX_MUD model is defined by an area of 7620 m long by 3000 m wide between the ECCO facility and the mud impoundment. The computational domain was simulated using a non-uniform rectangle mesh, the nodes of which were spaced at 8 m – 60 m intervals in the length, and from 8 to 30 m in the width.

To estimate the risk to the ECCO facility from the mudflow the following scenarios have been considered in the numerical experiments.

Scenario 1: It is assumed that mud has the same consistency, density and flow properties as pure water and that no rivers or any other structures that could mitigate, impede and/or slow down the mud flowing in the direction of the Ecco facility boundary existed. The consequences of the formation of 100 m wide dam breach are modeled. This scenario represents a ‘worse case’ as pure water will have faster and less impeded flow characteristics than actual mud toward the Ecco facility boundary. The initial mud elevation in the mud basin is 20.0 m.

Scenario 2 is similar to the Scenario 1 however the real mud parameters are used. According to laboratory testing the mud has a density of 1270 kg/m^3 and a viscosity at room temperature is 12.84 Pas. The flow rate of mud eruption from volcano was 175000 m^3/day. The yield shear stress of the mud was equal to 20 Pa and the friction parameter *K* was equal to 2285.

Scenario 3 is used the basic assumptions of the Scenario 2, but accounting for actual river channel flows that exist between the mud impoundment and the Ecco facility boundary. The Manning roughness coefficient for each river was assigned as 0.025 s m-1/3 and 0.11 s m-1/3 on the remaining part of territory.

Four terrain points (numbers 1-4 at the crosses in Figures 3 and 4) were used to highlight the temporal dynamics of the water level near to the ECCO site boundary. The point 1 is located at 1200 m from PT ECCO in the direction of the potential approach of the incoming mud flow, point 2 is located at 200 m (south-west) from PT ECCO at the southern bank of Blingo River, the Point 3 is at 800 m South East from PT ECCO and the Point 4 is located 1500 m eastward from PT ECCO.

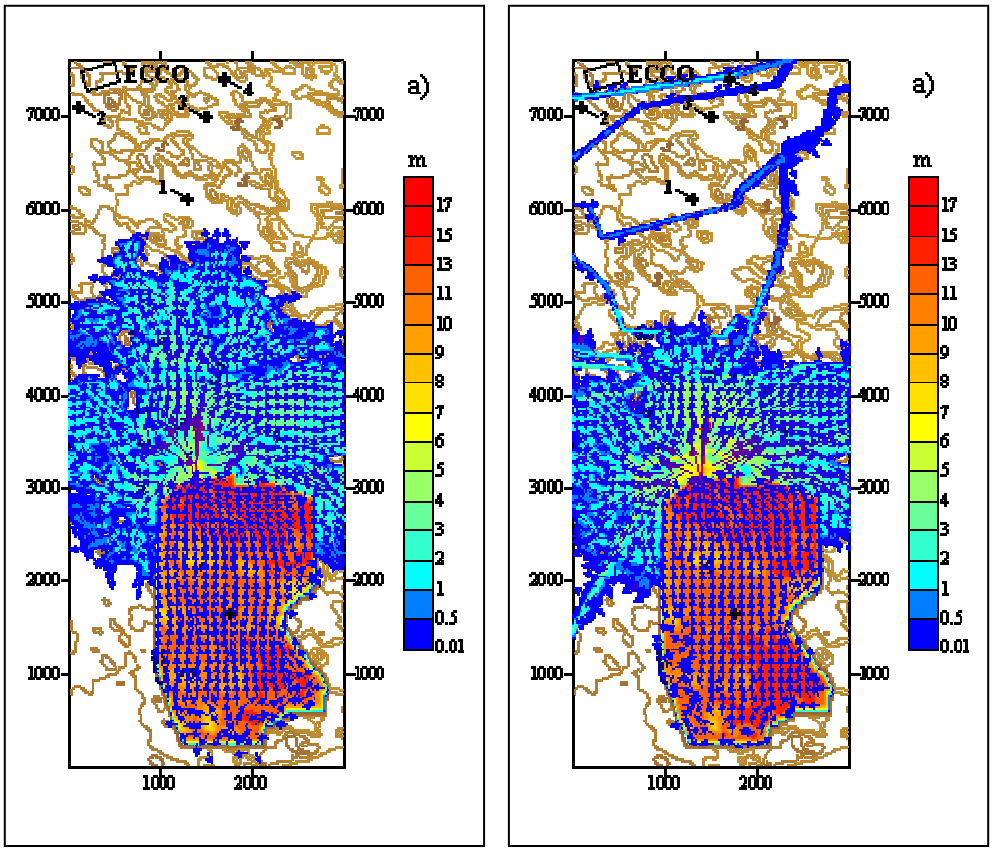


Figure 3. Flooded territory and water depth at 30 minutes s after the dam break for the Scenario 2 (left) and the scenario 3 (right)

The rivers delay and redirect the propagation of the mood flow after dam break (Figure 3). From the moment of 100 m breach formation the water reaches Point No 1 at 30 min in Scenario 1, at 1hour 20 min in Scenario 2 and at 5 hours after the dam break for the scenario 3 (Figure 4). . The maximum mud elevations at this point for three scenarios are 2.4 m, 2.3 m and 1.7 m respectively.

After passing the Point No 1 flow splits and propagates as two streams eastward and westward from PT ECCO (Figure 4). The territory upfront PT Ecco (south bank of Blingo River) practically is not inundated. In the closest to PT ECCO inundated point No 2 the water level increased 3 hours after breach formation till maximum value equals only 9 cm in the Scenario 1 and till 7.5 cm after 4 h in the Scenario 2. .

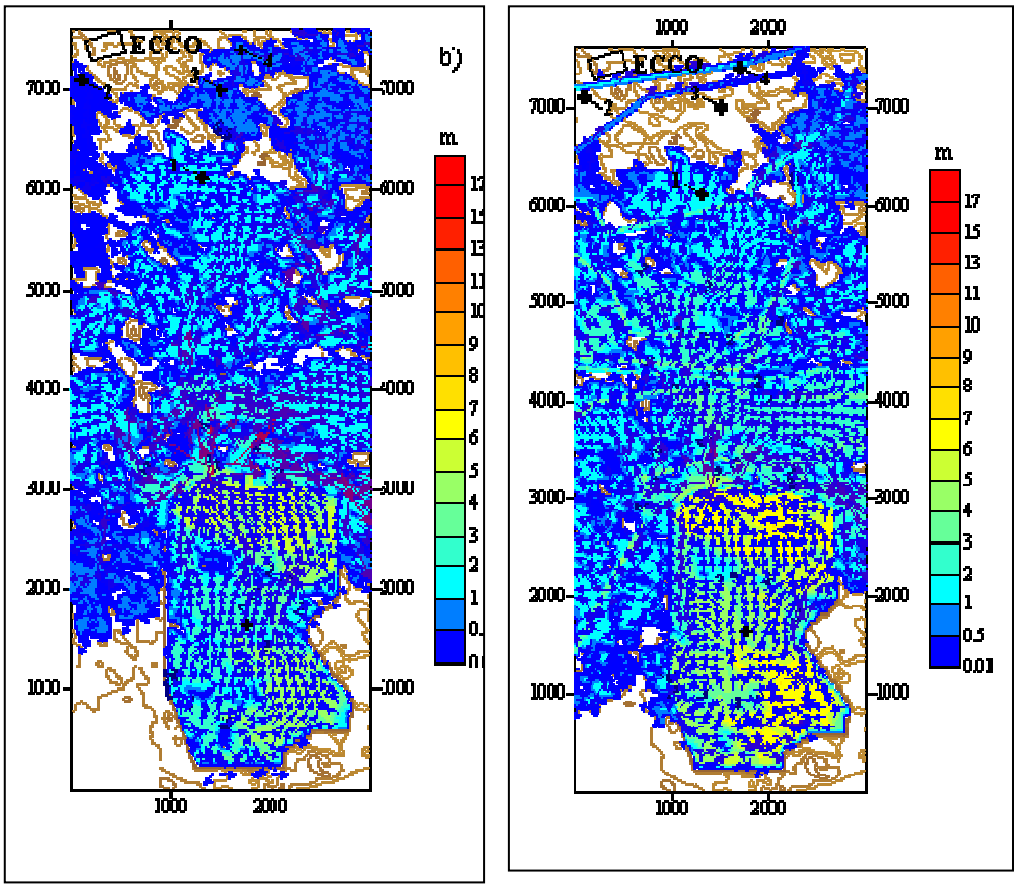


Figure 4. Flooded territory and water depth at 4 hours after the dam break. for the Scenario 2 (left) and the scenario 3 (right)

In the Scenario 3 under the influence of the river flows the mud flow is diluted and more significantly redirected to the North East than in the scenarios 1 and 2. The points No 2 and No 4 as also all right (South) bank of the Blingo river at PT ECCO will not be inundated in this Scenario driven by the mixing of the river flow and mud flow The simulated results of the Scenario 3 confirms the results of the more conservative scenarios 1 and 2 –the risk of the inundation for PT Ecco in a case of mud flow dam break is insignificant.

6. CONCLUSIONS

The two-dimensional finite volume model COASTOX_MUD has been developed for the simulation of variable density mud flow. The model was customized for the assessment of the consequences of the potential dam breaks at the Sidoarjo mud volcano, Indonesia on the basis of the topographical data, GIS data and mud parameters measured in the Brawijaya University, Malang, Indonesia and National University of Singapore. The two-components five equations model can simulate flow that is created by the mixing of the river channels flows and mud flow in the area between the mud volcano and the ECCO Indonesia factory. In the case of the homogenous mud density the derived system can be reduced to the 2-D three equations model, implemented in the previous studies by O'Brien et al, [1993]; Laigle and Coussot [1997]; Ming and Fread [1999]; Pudasaini and Hutter, [2007].

The conservative "worse case" risk assessment for the factory site inundation was provided in assumption that the mud flow has the same parameters as water flow and that all rivers between the volcano dams and PT ECCO site can be omitted. Within more

realistic scenario the measured mud density and viscosity were used in the model of permanent density. The simulations of the mixed river water – "mud flow were provided by the two components five equations version of the COASTAX _MUD model

The numerical results demonstrate that even in the "worse case " scenario the inundation at the ECCO factory after dam break will be less then 10 cm. For the scenario of the real density and viscosity the peak of the inundation is coming later and it will be lower. The simulation of the river flow impacts on the mud flow has demonstrated significant change of the flow direction (to North East) in this case that fully prevents the inundation of the PT ECCO site

However, in long term perspective if volcano would continue its activities, the next modeling studies should be provided after the mud sedimentation will cause the change of the depth in the mud basin.
The application of the variable density model in this case study shows how river flows can redirect mud flow and diminish significantly its elevation. Such results emphasize the needs for the implementation of variable density models for the realistic assessments of the mud flow propagation cross the rivernets.

ACKNOWLEDGEMENTS

The study was provided within the contracts between PT ECCO Indonesia and National University of Singapore (NUS) with the involvement of the experts from Ukraine and Indonesia.. The authors want to acknowledge to the Project Leader Dr. Jeff Obbard (NUS), Dr. Pavel Tkalich (NUS) and all other project team members for the fruitful cooperation.

REFERENCES

Cyranoski, D., Indonesian eruption: Muddy waters, *Nature*, *445*, 812-815, 2007.

Herbert, A.W., C.P. Jackson, and D.A. Lever, Coupled groundwater flow and solute transport with fluid density strongly dependent on concentration, *Water Resources Research*, *24*(11), 1781-1795, 1988.

Kivva, S., and M. Zheleznyak, Numerical modelling of 2-D open flow with moving boundaries: simulation of watershed runoff and long wave runup on a beach, *Computational Technologies*, *6*(2), 343-355, 2001. (in Russian)

Laigle, D. and P. Coussot, Numerical Modeling of Mudflows, *Journal of Hydraulic Engineering*, *123*(7), 617-623, 1997.

Lee, D.I., The viscosity of concentrated suspensions, *Transactions of the Society of Rheology*, *13*(2), 273-288, 1969.

Ming, J., and D. L. Fread, Modeling of mud/debris unsteady flows, *Journal of Hydraulic Engineering*, *125*(8), 827–834, 1999.

O'Brien, J. S., P. Y. Julien, and W. T. Fullerton, Two dimensional water flood and mudflow simulation, *Journal of Hydraulic Engineering, 119*(2), 244–261, 1993.

Oldenburg, C.M., and K. Pruess, Dispersive transport dynamics in a strongly coupled groundwater-brine flow system, *Water Resources Research, 31*(2), 289-302, 1995.

Onishi, Y., S. Kivva, M. Zheleznyak, and O. Voitsekhovich, Aquatic assessment of the Chernobyl nuclear accident and its remediation, *Journal of Environmental Engineering*, *133*(11), 1015-1023, 2007.

Pudasaini, S.P. and K. Hutter, Avalanche Dynamics: Dynamicsof Rapid Flows of Dense Granular Avalanches, Springer-Verlag, Berlin, 2007.

UNEP OCCHA, Environmental Assessment Hot Mud Flow East Java, Indonesia Final Technical Report: United Nations Disaster Assessment and Coordination mission in June & July 2006 and Follow-up mission in July 2006, Joint UNEP/OCHA Environment Unit Switzerland, 2006.

USGS. EROS –Earth Resource Observation and Science, http://eros.usgs.gov/products/elevation.html, 2007

Wang, Z.-Y., Free surface instability of non-Newtonian laminar flows, *Journal of Hydraulic Research, 40*(4), 449-460, 2002.

Wikipedia, Sidoraya mud flow, http://en.wikipedia.org/wiki/Sidoarjo_mud_flow, 2008.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Rheological Behaviour of Volcanic Granular Flows

K. Kelfoun

1. IRD - Instituto Geofisico, Escuela Politecnica Nacional, Quito, Ecuador

2. Laboratoire Magmas et Volcans - Université Blaise Pascal, CNRS, IRD – OPGC, 5 rue Kessler, 63038 Clermont-Ferrand, France (k.kelfoun@opgc.univ-bpclermont.fr)

Abstract

Numerical simulations of volcanic granular flows are increasingly being used for hazard assessment on volcanoes and appear to be essential for future hazard mitigation. A potential problem of such an approach, however, is that the rheological behaviour of such flows is very complex and currently impossible to describe fully from a physical point of view. As we cannot, at present, simulate all the complexity of the interactions at a microscopic level, we have to use simplified rheological laws. The frictional behaviour is often considered as a model that well describes the first order behaviour of such flows. The aim of this paper is to compare numerical simulations to well constrained field examples to check the validity of this law. Results show that the frictional model appears to be too simple to be used for pyroclastic flow and long runout avalanche simulations. A simple empirical law, using a constant retarding stress, represents a better alternative.

Keywords: Volcanic granular flow; Debris avalanche; Pyroclastic flow; Rheology.

1. INTRODUCTION

Long runout avalanches and dense pyroclastic flows are common in volcanic environments. In the following, they are called volcanic granular flows as they are formed of blocks and ash that interact during transport. Complex interactions also exist between particles and magmatic and atmospheric gases.

Volcanic long runout avalanches form when the flank of a volcano collapses. This destabilization may result from volcanic activity (e.g. magma injection), seismic activity or weakening of the rocks through time (e.g. by hydrothermal alteration). Huge masses of rocks (from a few km^3 to tens of km^3) are then transported away (up to tens of km) at inferred velocities of about 100 $m.s^{-1}$. They form widespread deposits, extending over several tens of kilometres squared and tens of metres thick. Their granulometry commonly ranges from micrometric ashes to decametre or kilometre-sized blocks.

Dense pyroclastic flows are fast-moving bodies of hot gas, ash and lava fragments. They are generated by the gravitational collapse of lava domes or eruptive columns, or by explosions from the volcanic conduit. Dense pyroclastic flows are generally a few metres thick, have a typical volume of 10^6-10^8 m^3 and their granulometry commonly ranges from micrometric ashes to centimetric-decimetric lapillis and blocks. Their velocity is of the order of 30-50 $m\ s^{-1}$.

Both phenomena are composed of broken rocks, present a very long runout compared to their drop height, and are thin relative to their area. Existing models of dense granular flows generally start from the assumption that the frictional (i.e. Mohr-Coulomb) behaviour is dominant in their rheology and plays the major role in deposit formation (see for example Pudasaini and Hutter, 2006, and references therein). This behaviour is chosen because it is exhibited by volcanic granular deposits and it also approximates the behaviour of sand flows in the laboratory. However, before using any model in hazard assessment, we should ensure that it correctly captures the first order features of the natural phenomenon. Thus, the following key questions need to be addressed: 1) is it realistic to consider the complex rheology of volcanic granular flows as mainly frictional? and 2) is this behaviour compatible with field observations of the geometry of their deposits, their relatively low velocity and their high mobility? To answer these questions, numerical simulations are compared to well-constrained field examples.

2. VOLCFLOW

Simulations have been done with the code *VolcFlow* developed at the *Laboratoire Magmas et Volcans*. It is based on the depth-average approximation. Using a topography-linked co-ordinate system, with *x* and *y* parallel to the local ground surface and *h* vertical, the general depth-averaged equations of mass (eq. 1) and momentum (eqs. 2, 3) conservation are:

$$\frac{\partial h}{dt}+\frac{\partial}{\partial x}(hu)+\frac{\partial}{\partial y}(hv)=0 \tag{1}$$

$$\frac{\partial}{\partial t}(hu)+\frac{\partial}{\partial x}(hu^2)+\frac{\partial}{\partial y}(huv)=gh\sin\alpha_x-\frac{1}{2}k_{actpass}\frac{\partial}{\partial x}\left(gh^2\cos\alpha\right)+\frac{T_x}{\rho} \tag{2}$$

$$\frac{\partial}{\partial t}(hv)+\frac{\partial}{\partial x}(hvu)+\frac{\partial}{\partial y}\left(hv^2\right)=gh\sin\alpha_y-\frac{1}{2}k_{actpass}\frac{\partial}{\partial y}\left(gh^2\cos\alpha\right)+\frac{T_y}{\rho} \tag{3}$$

where h is flow thickness, $\mathbf{u}=(u,v)$ is flow velocity, α is ground slope, $\mathbf{T}$ is retarding stress, ρ is the bulk density of the pyroclastic flow, k_{actpass} is the earth pressure coefficient [Savage and Hutter, 1991], and the subscripts denote components in the *x* and *y* directions. The model allows simulations of various rheologies, including frictional (with a basal and an internal friction angle), Bingham, viscous and Voelmy. More complex rheological laws can also be defined by the user.

The equations were solved using a shock-capturing numerical method based on a double upwind Eulerian scheme [Kelfoun and Druitt, 2005]. The scheme can handle shocks, rarefaction waves and granular jumps, and is stable even on complex topography and on both numerically 'wet' and 'dry' surfaces.

The value of k_{actpass} (eqs. 2 and 3) is a function of the basal (between the avalanche and the ground surface) and the internal friction angle φ_{bed} and φ_{int} respectively [Iverson and Denlinger, 2001].

For a dry frictional (i.e. Mohr-Coulomb) material, the retarding stress is of the form:

$$T_x=-\rho h\left(g\cos\alpha+\frac{\mathbf{u}^2}{r}\right)\tan\varphi_{bed}\frac{u}{\|\mathbf{u}\|} \tag{4}$$

The term $\mathbf{u}^2/r$ takes account of "overweight" due to the centrifugal acceleration on the topography curvature [Savage and Hutter, 1991]. The term $-u/\|\mathrm{u}\|$ allows the *x*-component of the retarding stress in the direction opposed to the displacement to be calculated.

Details and tests of the scheme are developed in more detail in Kelfoun and Druitt [2005]. Some additional details on the code, other examples and downloadable files are shown on: http://wwwobs.univ-bpclermont.fr/lmv/pperm/kelfoun_k/VolcFlow.html

3. EXAMPLE OF SOCOMPA LONG RUNOUT AVALANCHE

Socompa avalanche in northern Chile has been described in papers by Francis et al. [1985], Wadge et al. [1995] and Van wyk de Vries et al. [2001]. It formed by gravitational collapse of the northwestern flank of the 6000-m-high stratovolcano, leaving an amphitheatre 12 km wide at its mouth and with cliffs 300-400 m high. The avalanche flowed across a broad topographic basin northwest of the volcano up to a maximum distance of 40 km, and covered 500 km^2. The volume of rock transported is estimated to be about 25 km^3, with another 11 km^3 preserved as intact ('Toreva') blocks up to 400 m high at the foot of the collapse scarp. The morphology of the avalanche deposit is perfectly preserved in the hyper-arid climate of the Atacama Desert.

3.1 Frictional behaviour

Models were run assuming a frictional avalanche rheology (eq. 4) with three different combinations of basal and internal angles of dynamic friction: (1) $\varphi_{bed} << \varphi_{int} = 30°$, the static angle of friction for dry granular debris; (2) $\varphi_{bed} \neq 0°$ but $\varphi_{int} = 0°$; (3) $\varphi_{bed} = \varphi_{int} \neq 0°$ [Kelfoun and Druitt, 2005].

All three of these frictional models reproduce only very crudely the shape of the real avalanche deposit. A major failing is that, owing to the very low basal friction, the model avalanches flow off any gradients greater than 1 to 2.5° (depending on the case). After reaching their maximum limits, the avalanches drain back into the centre of the Monturaqui Basin. Consequently, the model deposits have negligible thickness along their limits of maximum extent, whereas thicknesses of up to 60 m are observed along the margins of the actual avalanche [Wadge et al., 1995]. The effect of topographic draining is to cause excess concentration of debris on the floor of the Monturaqui Basin.

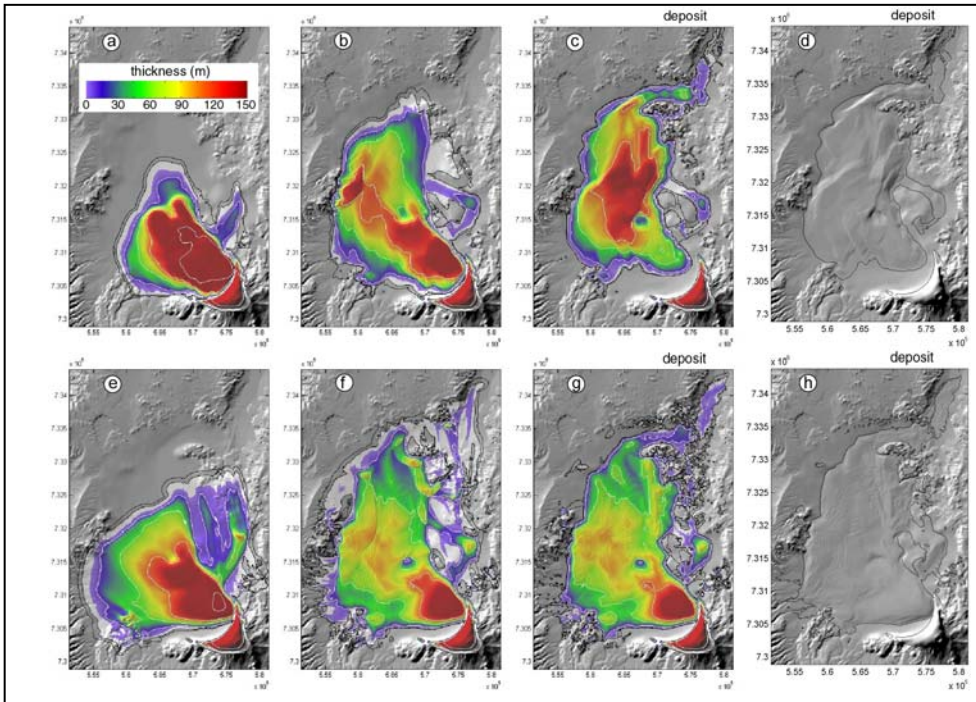


Figure 1. Snapshots of the emplacement of frictional avalanche models 1 and 2 at t = 200 s and t = 400 s, with the corresponding deposits (c and g) from Kelfoun and Druitt, 2005. (a-d) Model 1. Avalanche with $\varphi_{bed} = 1°$ and $\varphi_{int} = 30°$. (e-h) Model 2. Avalanche with $\varphi_{bed} = 2{,}5°$ and $\varphi_{int} = 0°$. The colour scale gives the thicknesses (m) of the avalanche. Figures d and h are shaded relief maps of the final deposits.

Simple frictional models are able to reproduce the approximate runout of Socompa avalanche only if very low values are used for the basal dynamic friction. However, they are unable to generate deposits with realistic thicknesses on slopes greater than about three degrees, nor realistic surface morphology.

3.2 Constant retarding stress

In view of the apparent inadequacy of the simple frictional models, Kelfoun and Druitt [2005] also ran models in which the retarding stress T in equations 2-3 was constant ($k_{actpass}$ was taken as unity). This very simple assumption was motivated by the study of Dade and Huppert [1998], who found that the field data for a large number of avalanches can be explained by an approximately constant retarding stress.

This empirical law simply states that the retarding stress is constant, independent of the velocity, thickness or any other parameter of the flow:

$$T_x = \text{Constant} \times \frac{u_x}{\|\mathbf{u}\|} \qquad \text{or} \qquad \mathbf{T} = \left\|\vec{T}x + \vec{T}y\right\| = \text{Constant} \qquad [6]$$

The models produce surprisingly good fits to the real avalanche, provided that T is about 50 kPa, depending on the initial slide geometry chosen. Unlike the frictional rheologies, this law produces a deposit with a well defined edge and leaves a deposit of realistic thickness on all slopes, irrespective of angle. Surface structures on the deposit from this simulation are remarkably similar to those of the real avalanche (Fig. 2).

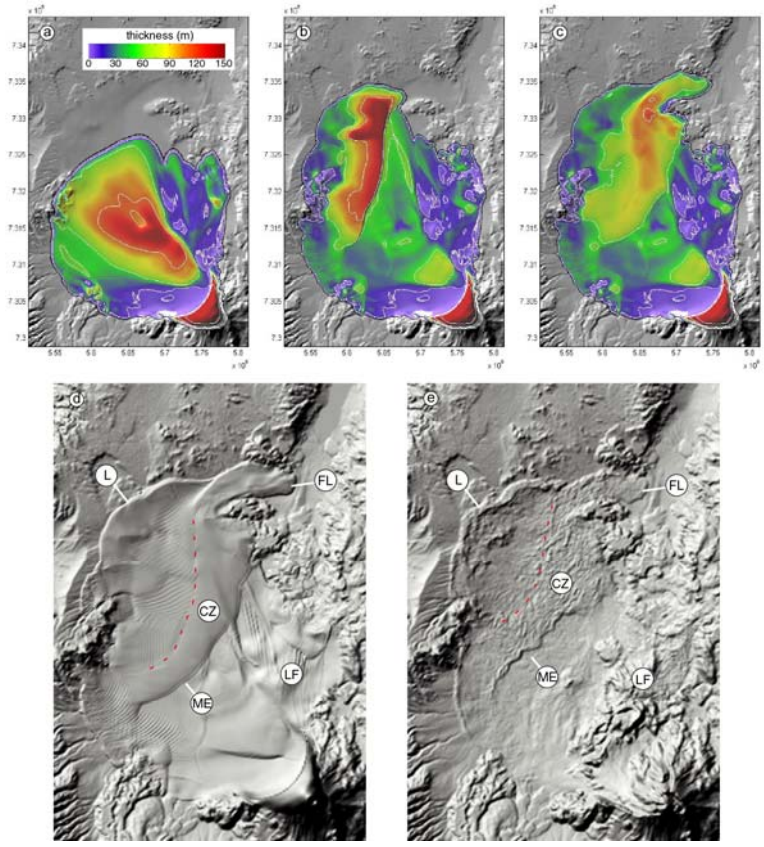


Figure 2. Avalanche evolution using a constant retarding stress T = 52 kPa [from Kelfoun and Druitt, 2005]. The colour scale denotes thickness. (a-c) Snapshots at 200 s, 400 s and 600 s. (d) Shaded relief map of the simulated deposit. (e) Shaded relief map of the real deposit.

A recent study of Kelfoun et al. [2008], based on field work and imagery analysis confirm the movement observed in the numerical model of Kelfoun and Druitt [2005]. It thus proves that the constant retarding stress is a good estimation of rheology of the avalanche.

4. EXAMPLE OF TUNGURAHUA PYROCLASTIC FLOWS

Tungurahua volcano in Ecuador erupted in July and August 2006, emitting pyroclastic flows that swept along the west flank of the volcano. The eruption was observed, described and monitored by the staff of the *Instituto de Geofisco* of the Escuela Politécnica Nacional (IG-EPN), Quito. Based on their data, it is possible to reconstruct the conditions of pyroclastic flow formation, transport and emplacement.

Results of this section are developed in Kelfoun et al. (submitted). The thickness of each unit is about 1 m, the total thickness of the deposit, composed of several units, being generally less than 10 metres. The feeding duration of the paroxysmal phase was 40 minutes and the total volume emitted between 5 and 10×10^6 m^3. Pyroclastic flows were formed in a quasi continuous manner close to the crater, with no, or very low, initial velocity. The maximal elevation at which pyroclastic flows are generated is 5000 m, the eastern rim is unbreached by pyroclastic flows. Field observations indicate that a greater volume of pyroclastic flows formed at the low points of the rim than at higher, and it is assumed that the mass rate at the feeding cells, expressed by the variation of the thickness *dh* with time *t*, is a linear function of the difference between the maximal elevation z_{gen} = 5000 m where pyroclastic flows were generated and $z_{rim,i}$, the elevation of the crater rim.

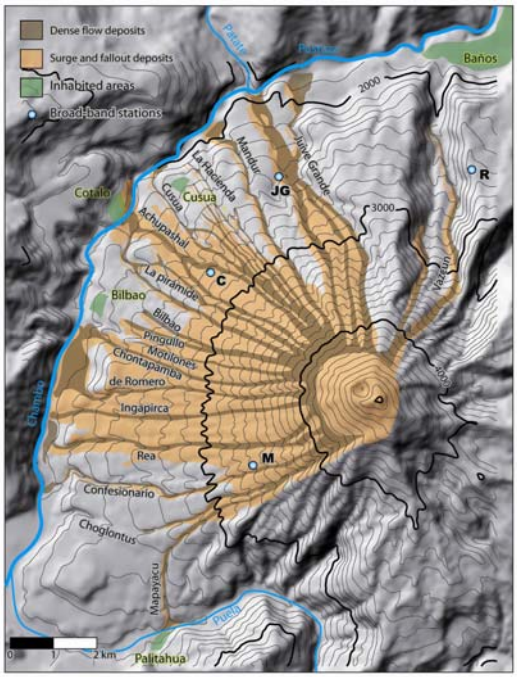


Figure 3. Map of deposits, Tungurahua, august 2006. Dense pyroclastic flows are in brown (from Kelfoun et al., submitted).

The new mass *dh* is assumed to be poured without velocity, the velocity of the cells at the feeding points being calculated by the conservation of momentum. In the simulation, the source is assumed as being at the cells located just outside the crater rim, at a distance of 350 m from the crater centre.

Frictional behavior

The frictional model that best reproduces the extension of the August 2006 eruptions of Tungurahua assumes φ_{bed} = 15° for φ_{int} equaling 0. The value of φ_{int} modifies runout and velocity of the flows but does not significantly change the conclusions of the paper (as for Socompa) and is not detailed here. At the onset of the eruption, the mass poured on the crater rim immediately began to flow and rapidly covered all the SW, W and NW flanks with a very thin thickness of <0.1mm. The main drainage channels were are affected, but also interfluves. The mass accelerated along the steep (>30°) slope of the terminal cone, reaching a velocity of more than 200m.s^{-1}. The thickness of pyroclastic flows on the interfluves does not significantly increase with time, but it increased in the valleys into which the mass drained. The mass began to decelerate and accumulate where the slope is inferior to the friction angle. After the alimentation stopped, at *t* = 2400 s, all the mass deposited on the flank progressively drained down and accumulated at the foot of the volcano on slopes slightly lower than the friction angle, where it formed sand piles with surface angle of about 15° (Fig.4b).

The incompatibility between numerical results and field data is not due to a problem in the numerical code or DEM resolution: put a grain of frictional material (sand for example) at the surface of a curved slope and the grain will immediately slide along the slope and will stop a little bit after the slope becomes less than the friction angle of the material. Put several grains and they will all stop at approximately the same place. Pour the sand over some tens of seconds and the sand will flow and accumulate to form a sand pile. Our frictional results reproduce exactly this frictional behaviour.

These results could point out a potential problem of using the frictional model for hazard assessment. One could argue that the shape of deposits is not important for hazard assessment and that a good map of inundation can be obtained with a simple model that does not capture all the complexity of the physics of the phenomenon. But results show that the frictional model cannot reproduce a correct map of inundation at Tungurahua. Firstly, all the interfluves are affected by dense flows and this largely overestimates the areas which are threatened (interfluves may have been affected by surges but the frictional model is not adapted for their simulation). Secondly, the mass accumulates in thick piles (up to 50m) and cannot spread. The model thus underestimates the hazard in flat areas. Thirdly, a velocity 5 to 10 times higher than in reality allows pyroclastic flows to cross ridges that they would not be able to cross in reality. The portion of the flows that crosses the ridge is no longer available to follow the drainage and to affect downstream areas. In this case of drainage bending, the frictional model can greatly underestimate areas affected. In the case of the Tungurahua, this occurs to the south: the deposits formed by the model at $y = 9.833 \times 10^6$ (Fig. 4) crossed a 100 m-high ridge. No deposits were found here in the field; instead the mass followed the main river to affect the populated area of Palitahua (Fig. 3).

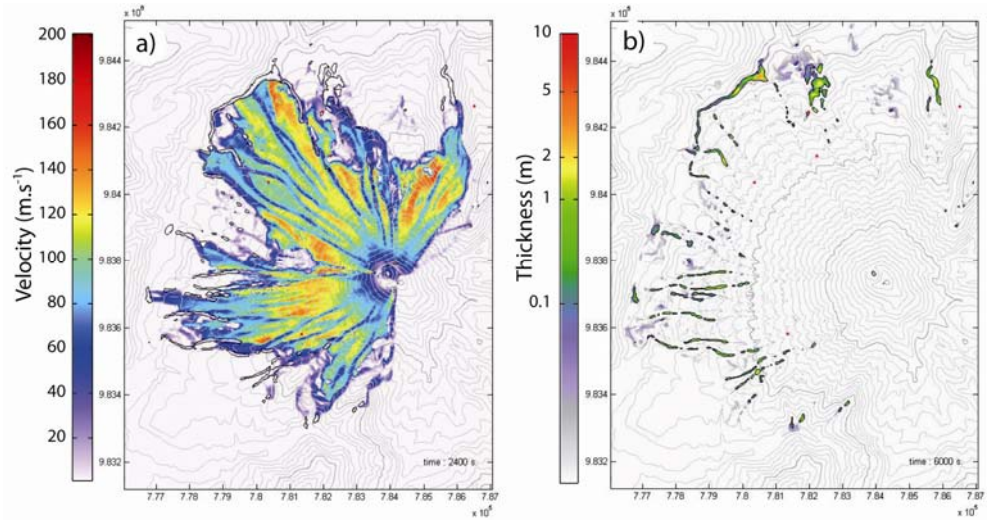


Figure 4. Velocity of pyroclastic flows at t=2400s (a) and thickness of deposits (b), simulated using the frictional model (15°). The colour scale of the thickness (b) is logarithmic. Black contours indicate 0.1 m deposits (no thinner pyroclastic flows were observed on the field). From Kelfoun et al., submitted.

It should also be noted that all the characteristics described here which show that frictional behaviour differs from pyroclastic flows are largely independent of the total volume, feeding conditions, details of the topography, value of the friction angle chosen, etc. The same behaviour occurs when the rate, volume, or the way the mass is injected are changed.

Constant retarding stress

The best fit results (extension, thickness, velocity) for the constant retarding stress are obtained with a constant retarding stress of 7kPa (Fig.5).

The first difference with the frictional model is that the mass accumulates around the rim before reaching a certain thickness (about 1 m) that allows flow initiation. Figure 5a shows that emplacement is carried out by pulses (in red) even if the source rate is constant: the mass accumulates until it reaches a critical thickness (about 1 m) which enables flow to take place. The mass then continues to be accumulated until the critical thickness is reached again, and a new pulse forms. Pulses are thicker and faster than the neighbouring material. Pulses generally run at less than 30m.s^{-1}, which corresponds to field observations. The front of the flows is well defined during emplacement. The simulated deposits also show a well defined front and a progressive decrease in thickness on steeper slopes. Simulated dense pyroclastic flows present a realistic thickness (0.5 – 2 m thick) and are restricted to valleys (Fig 5).

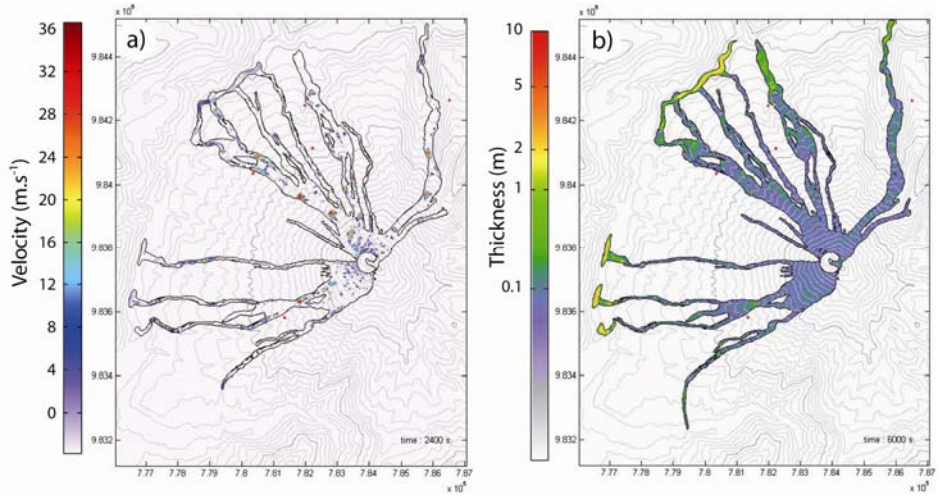


Figure 5: Velocity of pyroclastic flows at t=2400s (a) and thickness of deposits (b), simulated using the constant stress model (7 kPa). The colour scale of the thickness (b) is logarithmic. Black contours indicate 0.1 m deposits. From Kelfoun et al., submitted.

OTHER EXAMPLES

Socompa and Tungurahua are two examples among other field cases that have also been tested including pyroclastic flows on Lascar (Chile) and Merapi (Indonesia), and avalanches on Lastarria (Chile) and Llullaillaico (Argentina). Constant stress is always a better rheology than frictional stress. The value of the constant retarding stress is less than 10 kPa for pyroclastic and pumice flows, and about 50 kPa for long runout avalanches.

CONCLUSIONS

The frictional model appears not to be the best adapted to simulate pyroclastic flow and debris avalanche emplacement and deposits at Tungurahua and Socompa, as well as at other field example studied. Contact between particles at rest is surely frictional but other phenomena act when particles are moving to give a much more complex behaviour to the flow. The constant retarding model appears much better to capture this complex behaviour and to simulate pyroclastic flows.

The greater adequacy of the constant retarding stress proves that the ratio of driving / retarding stress in granular flows cannot be considered as constant – as for the frictional model – but decreases as the thickness increases. Above a certain thickness the flow is able to move. Less than this thickness, the retarding stress dominates the driving stress and the flow slows down. This produces flow deposits with a more or less constant thickness that

decreases on slopes, and this appears consistent with field observations. The behaviour is probably explained by an increase of the mechanical strength of the flows from their base or interior to the surface. In Socompa avalanche, a detailed study of the deformation of the surface lithology shows that the avalanche was formed by a fluidal interior covered by a rafted brittle crust (Kelfoun et al., 2008). This change in mechanical properties may be related to the variation of granulometry from bottom to top: the interior of the flow being constituted of finer particles than the surface, it may, for example, retain gas more easily thus prolonging its flowing capability. This increase of granulometry towards the surface, common in granular flow deposits, was observed at Socompa, and also appears very clearly in the Tungurahua and Lascar deposits.

The physics of pyroclastic flows is complex and many developments are necessary to fully understand it, and to obtain a robust physical model. Results of this paper will be useful for future evaluation of the quality of such a complex model that will have to explain why a constant stress model appears much more adapted than a frictional model to simulating pyroclastic flows. In the meantime, the frictional model appears to be too simple to be used for pyroclastic flow and long runout avalanche simulations. Constant retarding stress seems to represent an acceptable alternative.

ACKNOWLEDGEMENTS

This work was financed by two research programs of the French CNRS: "*Relief de la Terre*" and "*Aleas et Changements Globaux*" and by the *Institut pour la Recherche et le Développement* (IRD).

REFERENCES

Dade, W.B, and H.E. Huppert, Long-runout rockfalls, *Geology*, 26, 803-806, 1998.

Francis P.W., M. Gardeweg, C.F. Ramirez, and D.A. Rothery, Catastrophic debris avalanche deposit of Socompa volcano, northern Chile. *Geology*, 13, 600-603, 1985.

Iverson, R.M., and R.P. Denlinger, Flow of variably fluidized granular masses across three-dimensional terrain 1. Coulomb mixture theory, *Journal of Geophysical Research*, 106, 537-552, 2001.

Kelfoun K, and T.H. Druitt, Numerical modeling of the emplacement of Socompa rock avalanche, Chile. *Journal of Geophysical Research*, 110, B12202.1-12202.13 DOI:10.1029/2005JB003758, 2005.

Kelfoun, K., T.H. Druitt, B. van Wyk de Vries, and M.N. Guilbaud, Topographic reflection of Socompa debris avalanche, Chile, *Bulletin of Volcanology*, 10.1007/s00445-008-0201-6, 2008.

Kelfoun, K., Samaniego, P., Palacios, P., Barba D., Is frictional behaviour suitable for pyroclastic flow simulation: comparison with a well constrained eruption at Tungurahua volcano (Ecuador), submitted to the *Bulletin of Volcanology*.

Pudasaini, S.P., Hutter, K., Avalanche Dynamics: Dynamics of Rapid Flows of Dense Granular Avalanches, ISBN: 3-540-32686-3, Springer Verlag, 602 pp, 2006.

Savage, S.B., and K. Hutter, The dynamics of avalanches of granular materials from initiation to runout. Part I: Analysis, *Acta Mechanica*, 86, 201-223, 1991.

Van Wyk de Vries, B., S. Self, P.W. Francis, and L.A. Keszthelyi, A gravitational spreading origin for the Socompa debris avalanche, *Journal of Volcanology and Geothermal Research*, 105, 225-247, 2001.

Wadge G., P.W. Francis, and C.F. Ramirez, The Socompa collapse and avalanche event. *Journal of Volcanology and Geothermal Research.*, 66, 309-336, 1995.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Geomechanical modelling of triggering mechanisms for rainfall-induced triangular shallow landslides of the flow-type

L. Cascini[a], S. Cuomo[a], M. Pastor[b], J.A. Fernández-Merodo[b]

[a] University of Salerno, Department of Civil Engineering, Italy
(l.cascini@unisa.it, scuomo@unisa.it)
[b] Centro de Estudios y Experimentación de Obras Públicas (CEDEX), Madrid, Spain
(mpastor@cedex.es, jose.a.fernandez@cedex.es)

Abstract: Landslides of the flow-type are frequently triggered by rainfall as shallow landslides in different types of soils and geo-environmental contexts. When involving open slopes, these landslides often occur in triangular source areas where initial slides turn into avalanches through further failures and/or eventual soil entrainment. Similar phenomena are frequently observed in layered snow deposits and they are also recognised in ashy deposits on Mars. Significant examples are systematically recorded in Southern Italy where pyroclastic deposits overlie carbonate massifs. Particularly, in May 1998, rainfall triggered many destructive triangular shallow landslides of the flow-type along the slopes of Pizzo d'Alvano massif (Sarno-Quindici event). The available data-set allowed their failure and post-failure stages to be modelled. Numerical analyses were performed using limit equilibrium as well as hydro-mechanical coupled and uncoupled stress-strain approaches. The obtained results are discussed in the paper, providing a preliminary framework for this type of slope instability phenomena that is poorly addressed in the current literature.

Keywords: Rainfall, Shallow landslides, Flow-type, Triangular source areas, Avalanche.

1. INTRODUCTION

Landslides of the flow-type [Hungr et al., 2001, Pastor et al., 2008], when involving open slopes, frequently originate as small landslides (failure stage) that increase their initial volumes in triangular-shaped areas (post-failure stage) according to complex mechanisms not yet completely understood [Chen et al., 2006; Crosta et al., 2006]. Similar phenomena are often observed in layered snow deposits [Jamieson & Stethem, 2002] where different snow avalanches occur in relation to in-situ and climate conditions [Pielmeier et al., 2003]. Triangular avalanche phenomena were also recognised in some ashy deposits on Mars, where high resolution images [Gerstell et al., 2004] highlight the presence of impact craters at the head of the scars.

In shallow soil deposits, the failure stage of triangular-shaped landslides triggered by rainfall is associated with different triggering factors. Karst springs from bedrock are recognised as a possible cause [Cascini et al., 2005, 2008; Di Crescenzo & Santo, 2005] with significant examples for pyroclastic deposits of Southern Italy discussed by Budetta & de Riso [2004] and Guadagno et al. [2005]. Another triggering factor is the impact of failed soil masses on stable deposits [Hutchinson & Bhandari, 1971] with major examples available for torrent deposits in USA [Costa & Williams, 1984], for decomposed granitic soils in Japan [Wang et al., 2003] and for pyroclastic deposits in Southern Italy [Guadagno et al., 2005; Di Crescenzo & Santo, 2005]. The transition from initial slides to avalanches [Hungr et al., 2001] can be related to soil/material entrainment along the path of the failed masses [Jakob and Hungr, 2005; Crosta et al., 2006]. The increase of the initial failed volumes can be also associated with further failures occurring inside the landslides source

areas. Examples are available for events occurred in Southern Italy [Cascini, 2005, 2008, Cuomo, 2006] and for snow deposits [Pielmeier et al., 2003]. Moreover, the transition from slide to avalanche can be connected to liquefaction phenomena. In the literature, contributions based on theoretical evaluations and laboratory tests are available on the topic [Wang et al., 2003].

Due to the catastrophic consequences that these phenomena generally cause [Cascini et al., 2005], the understanding and modelling of both the failure and post-failure stages are a fundamental issue in mitigating the posed risk to life and facilities. The paper addresses this topic using different geomechanical approaches, with reference to a study area (Sarno-Quindici, Southern Italy) where disastrous events were recently recorded.

2. CASE STUDY

2.1 Landslides event

Shallow landslides of the flow-type have been systematically recorded for many centuries [Guadagno et al., 2005] in a populated area of Southern Italy where unsaturated pyroclastic soils [Cascini & Sorbino, 2002] overlie carbonate massifs. Particularly, on 4-5 May 1998, many destructive landslides of the flow-type were triggered by heavy rainfall [Cascini et al., 2000; Fiorillo & Wilson, 2004], causing 159 fatalities and relevant damage to property. Different and complex triggering mechanisms were recognised [Cascini et al., 2005, 2008]. One of the most frequent triggering mechanism (M2) was associated with triangular landslide source areas by Cascini et al. [2005, 2008]. The source areas are a few meters wide near the bedrock scarps and downwards enlarging (Fig. 1a). For these landslides, Cascini et al. [2008] recognised the karst springs from the bedrock (scheme M2a) and the impact of failed soil masses on stable deposits (scheme M2b) as the main triggering factors (Fig. 1b). However, also for the latter ones (scheme M2b), the presence of a karst spring was recognised as a necessary condition for the downwards enlargement of the impacted zone at the base of the scarps. Particularly, Cascini et al. [2008] observed that, in both cases, karst spring were only active for short periods of time (less than 24 hours) when they also had a high discharge (up to 10^{-4} m/s). On the other hand, the impact phenomena were related to small volumes of failed soil masses (10 - 100 m^3) falling from bedrock scarps (2 – 20 m tall). Figure 1c shows the areal distribution of the triangular-shaped phenomena that occurred in the study area in May 1998 [Cascini et al., 2008].

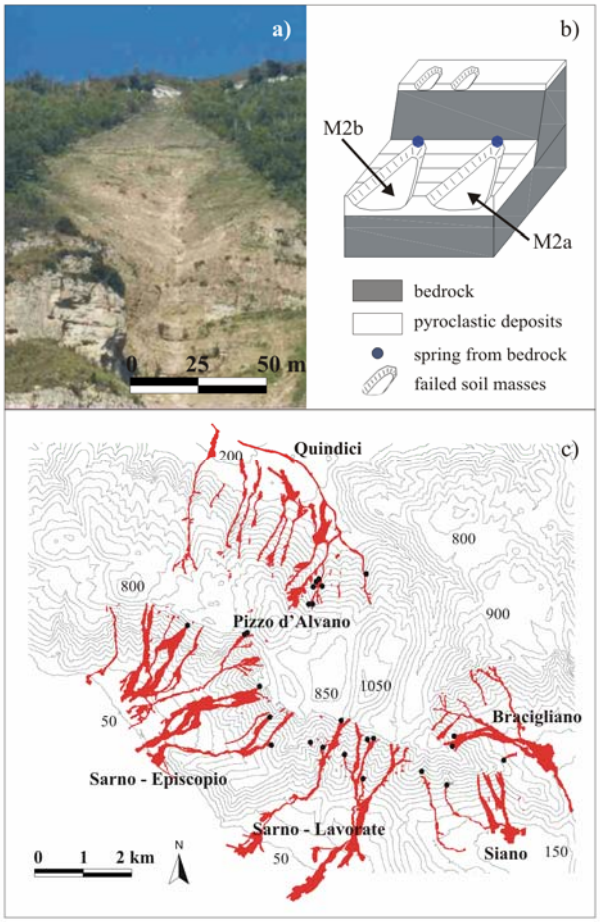


Figure 1. Triangular shallow landslides of flow-type occurred in the study area on May 1998.

2.2 Literature review

Despite their catastrophic consequences, these landslides are not exhaustively discussed in the literature as far as their failure and post-failure stages are concerned. Among the available contributions, morphometrical analyses suggest that landslides crown zones are characterised by apical angles ranging between 15° and 30° and steepness from 35° to 45° [Di Crescenzo & Santo, 2005]. However, similar analyses show that the geometrical features of the landslides source areas (apex angle of the source areas, height of natural and anthropogenic scarps, slope length, slope angle, initial volume) can be poorly correlated with each other [Guadagno et al., 2005].

These landslides have not been exhaustively modelled. Among the available contributions, Calcaterra et al. (2004) show that karst springs from bedrock affect the groundwater regime in large portions of the pyroclastic deposits (2-12 m in length) after 6-24 hours. However, the consequential slope failures are not analysed or modelled. On the other hand, Guadagno et al. [2003] highlight the role of bedrock scarps for upslope failures while slope instabilities occurring downslope the scarps are not addressed. As a contribution to the topic, the paper proposes a rational framework for geomechanical modelling aimed at investigating the failure and post-failure stages of the phenomena occurred in May 1998.

2.3 Proposed approaches, in-situ conditions and soil properties

The failure stage associated with springs from bedrock (scheme M2a) can be profitably analysed through an hydro-mechanical uncoupled approach. This approach is firstly based on pore water pressures modelling in saturated-unsaturated conditions through the numerical integration of the well-known Richards' equation. The obtained results can be then used as input to perform limit equilibrium analyses that relate the failure onset to slope safety factors approaching to one. To this aim, a rigid-perfectly plastic model can be adopted with a Mohr-Coulomb criterion extended to unsaturated conditions [Fredlund et al., 1978]. Referring to the same pore water pressures, stress-strain analyses can also be performed to validate the previous findings. Particularly, an elastic-perfectly plastic soil model can be used and both equilibrium and compatibility equations can be taken into account for the landsliding masses.

Failures induced by impact phenomena on stable deposits (scheme M2b) can be investigated using the above mentioned uncoupled approach, under the assumption of drained conditions. Conversely, if undrained conditions are considered, the excess pore water pressures must be necessarily evaluated referring, for instance, to the total isotropic and deviatoric stress variations respectively through the pore pressure coefficients α and β [Henkel, 1960]. Alternatively, hydro-mechanical coupled approaches [Pastor et al., 2002] can be used to take into account the effects of soil deformations on pore water pressures.

For the post-failure stage of these landslides, geomechanical modelling can be rigorously performed through hydro-mechanical coupled approaches that represent powerful tools especially when advanced soil constitutive models are available. However, uncoupled stress-strain analyses can be also used to investigate, under simplified assumptions, some aspects of the post-failure stage.

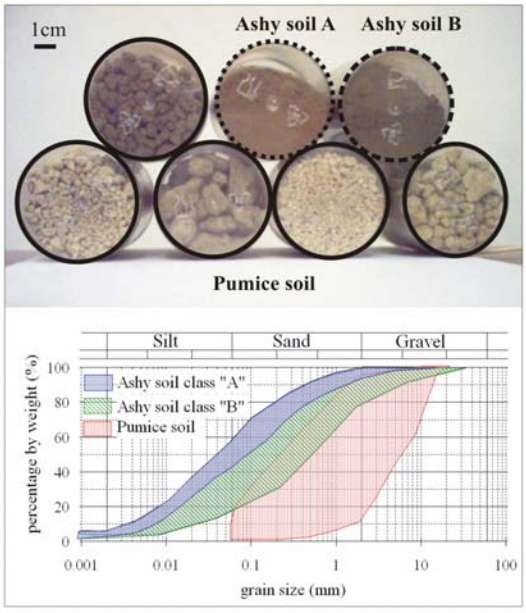


Figure 2. Pyroclastic soils from the study area.

The analyses discussed in the following sections were carried out on the available data-set that refers to both in-situ conditions and soil properties. As it concerns in-situ conditions, the typical stratigraphies consist of alternating layers (0.2-2m thick) that, as recognised by Cascini et al. [2005], belong to three main soil classes [Bilotta et al., 2005] i.e. pumice soils, coarser superficial ashy soils (class B) and finer deep ashy soil (class A) (Fig. 2).

As for the soil mechanical properties in saturated and unsaturated conditions, the literature provides soil water content and conductivity curves as well as the shear strength and stiffness [Bilotta et al., 2005] for typical suction values measured during the whole hydrological year [Cascini & Sorbino, 2002]. Typical values of the geotechnical properties are summarised in table 1.

Table 1. Soil mechanical properties.

	γ_d (kN/m^3)	γ_{sat} (kN/m^3)	n (-)	k_{sat} (m/s)	c' (kPa)	φ' (°)	φ^b (°)	ν (-)	E (MPa)	ψ (°)
Class B ashy soil	7.30	13.1	0.58	10^{-5}	0 ÷ 5	36 ÷ 41	20	0.26	5 ÷ 7	10 ÷ 20
Pumice soil	6.20	13.1	0.69	10^{-4}	0	37	20	-	-	-
Class A ashy soil	9.10	15.7	0.66	10^{-6}	5 ÷ 15	32 ÷ 35	20	0.30	1 ÷ 3	10 ÷ 20

3. PRELIMINARY GEOMECHANICAL MODELLING

3.1 Failure stage: the role of karst springs from bedrock

Using the uncoupled approach discussed in sect. 2.3 and the field data on karst springs from the bedrock (sect. 2.1), pore water pressures were computed through the Seep/W Finite Element code [Geoslope, 2004] for time periods shorter than 48 hours referring to the typical stratigraphies of the pyroclastic deposits (schemes 1-3 in Fig. 3) and soil mechanical properties listed in table 1. For each scheme, the initial conditions were assessed through a preliminary transient seepage analysis over the period January 1, 1998 – May 3, 1998. Unsaturated soil conditions were simulated throughout the slope section. The hydraulic boundary conditions were selected as follows: i) measured rainfall values at the ground surface [Cascini et al., 2008], ii) impervious boundary at the contact between bedrock and pyroclastic deposits; iii) water flux or hydrostatic distribution at upper lateral boundary of the pyroclastic deposits due to the presence of a karst spring. The achieved results outline that karst spring induces transient pore water pressures with maximum simulated values equal to 30 kPa for scheme 1 containing only class B ashy soils and 40 kPa for scheme 2 including class A and B ashy soils (Fig. 3). Similar results were obtained assuming hydrostatic distributions of pore water pressures at the upper lateral boundary of the deposits (i.e. corresponding to fractures 1.8 – 4.5 m deep), so highlighting the stratigraphy as a key factor in the groundwater regime. Particularly, ashy A soils strongly increase the simulated pore water pressures (scheme 2 in Fig. 3) while pumice soil layers determine an enlargement of the slope portion affected by the presence of springs (scheme 3 in Fig. 3).

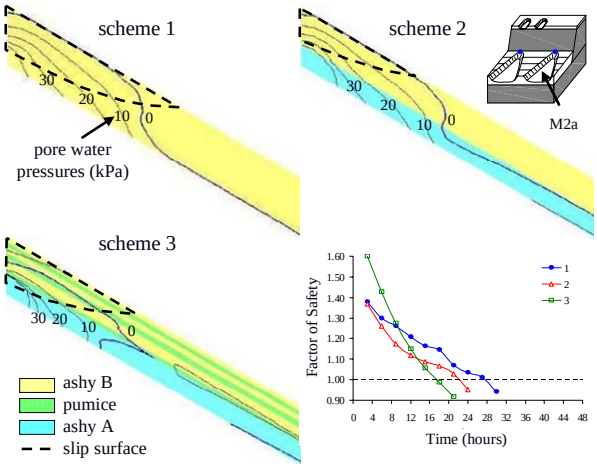


Figure 3. Failure stage induced by karst springs from

Once the pore water pressure regime was evaluated, the failure onset was investigated with limit equilibrium analyses, through the Slope/W code [Geoslope, 2004]. Planar

and curvilinear slip surfaces were considered. The curvilinear slip surfaces are associated to the minimum safety factors. Particularly, failure conditions were simulated assuming low discharges (3×10^{-5} m/s) over short time periods (> 20 h). Similar results were obtained by considering hydrostatic pore water pressure distributions at the upper lateral boundary of the pyroclastic deposit. In all the cases analysed, failed volumes depend on pore water pressures and soil shear strength properties, that are both strictly related to stratigraphy (Fig. 3). The latter is therefore outlined as a key factor for landsliding. The obtained results are confirmed by stress-strain analyses [Cuomo, 2006] performed using the Sigma/W Finite Element code [Geoslope, 2004], assuming an elastic-perfectly plastic constitutive model and soil properties in table 1. For instance, for the scheme 1, the simulated strains concentrate at the bedrock-cover contact well matching the slip surfaces assumed in the limit equilibrium analyses.

3.2 Failure stage: the role of impact of failed soil masses on stable deposits

This triggering mechanism was investigated, in static conditions, through a parametric analysis. As for the stratigraphy, the schemes of sect. 3.1 were considered. Impact loading pressures were estimated through the procedure proposed by Wang et al. [2003], referring to typical values for the heights of bedrock scarps and failing volumes (sect. 2.1). Impact loading pressures were so estimated as ranging between 5 kN/m and 30 kN/m. Moreover, distinct initial pore water pressures were considered, being related to the season and the eventual presence of karst springs from the bedrock.

Geomechanical modelling was firstly performed assuming drained conditions. As initial conditions, pore water pressures were assumed equal to those obtained with the previous transient seepage analysis over the period January 1, 1998 – May 3, 1998. Uniform distributions of pore water pressures were also considered with values ranging between 5 kPa and 60 kPa, that are the extreme values recorded during the in-situ investigations [Cascini & Sorbino, 2002]. Limit equilibrium analyses were performed through the Slope/W code, assuming uniform impact loading pressures. Curvilinear and planar slip surfaces were referred. The planar slip surfaces are associated to the minimum safety factors. The failure onset was not simulated, independently from the considered stratigraphy, pore water pressure and impact loading pressure. In fact, the impact loading pressures increase both the driving and resistance forces for the assumed slip surfaces essentially due to the slope geometry and loading conditions. As a consequence, slope safety factors change slightly with load application and achieve their initial value on load removal. As an example, figure 4 shows the results obtained for different stratigraphies and pore water pressures equal to those simulated at May 3, 1998.

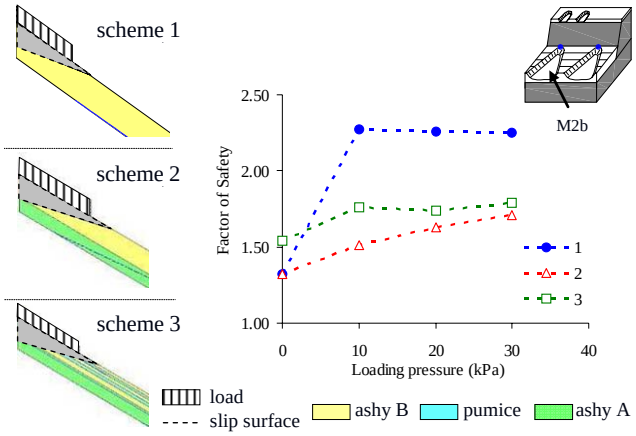


Figure 4. Effects of impact phenomena in drained conditions.

Based on the same pore water pressures, stress-strain analyses [Cuomo, 2006] outline the highest deformations in portions of the slope well matching the critical slip surfaces (Fmin) computed with the limit equilibrium analyses. Moreover, failure conditions are not predicted by the numerical analysis for any initial pore water pressure distribution.

Geomechanical modelling was then performed assuming that during the impact phenomena undrained conditions can exist. Referring to the uncoupled approach discussed in sect. 2.2, impact-induced variations of total stresses were firstly computed using the Sigma/W code.

Then, assuming nearly saturated soil conditions for suction values up to 5 kPa ÷ 10 kPa, for the ashy soils, the coefficient α was estimated (between 0.33 and 1) while the coefficient β was assumed equal to one. Finally, impact-induced pore water pressures variations were computed referring to the variations of total stresses and α, β coefficients. The performed analyses outline that deviatoric stresses strongly increase in the impact zone. Conversely, effective isotropic stresses significantly decrease due to the build up of the pore water pressures (Fig. 5). Indeed, the latter ones reach high values that are related to the soil stiffness, applied loading pressures and initial conditions. As a consequence, liquefaction phenomena may also occur. However, the accurate numerical modelling of these phenomena requires the use of advanced soil constitutive models whose calibration for these soils is not available in current literature. Consequently, a rigid-perfectly plastic constitutive model was consider first. Using this model, limit equilibrium analyses were performed assuming the stratigraphies of sect. 3.1 and the pore water pressures previously computed in undrained conditions. The example provided in figure 5a refers to scheme 1 of figure 3 (ashy B soils), having α and β respectively equal to 0.5 and 1, and initial pore water pressures equal to those computed on May 3, 1998. The obtained results outline that the slope safety factor, at load application, is still higher than unity (F=1.531), notwithstanding the increase of pore water pressures. Conversely, at loading removal, the same pore water pressures correspond to failure conditions (F=0.989). Similar slope instability scenarios can be drawn considering an initial suction value equal to 5kPa.

To validate the previous findings, the hydro-mechanical coupled approach proposed by Pastor et al. [2002] was used considering an elastic-perfectly plastic constitutive model. The obtained results match the previously simulated slope instability scenarios well as it concerns both pore water pressures and plastic strains. 3D stress-strain analyses also outline that the initial stress field is significantly modified by the assumed impact loading pressures. Particularly, the highest shear stresses (σ_{yz}) mainly concentrate along a direction forming an angle of 15-45° with the x-direction and plastic strains arise in a zone downslope enlarging (Fig. 5b).

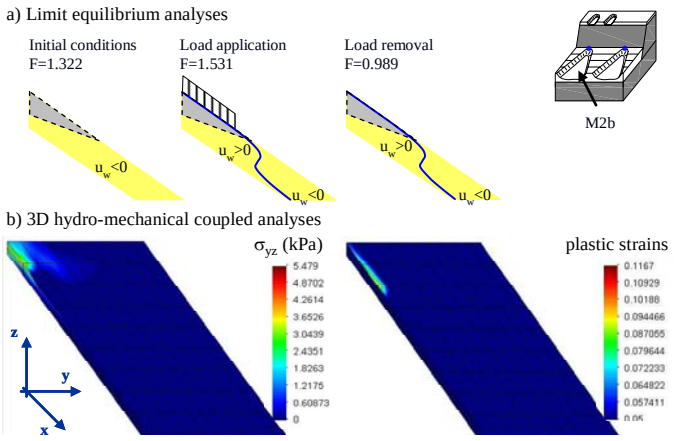


Figure 5. Effects of impact phenomena in undrained conditions.

3.3 An interpretation for the post-failure stage

The geomechanical modelling of the post-failure stage was performed using the simplified hydro-mechanical uncoupled approach discussed in sect. 2.2. To this aim, stress-strain analyses were performed referring, in static conditions, to planes parallel to the ground surface (s-y planes in Fig. 6). In these planes, at different depths, stresses along the longitudinal s-direction and lateral y-direction depend on slope angles and soil mechanical properties. At the upper boundary of these planes, loading pressures (σ_L) were assumed to be downwards applied along the longitudinal s-direction, to simulate the presence of a failed mass pushing the stable deposits downslope. The obtained results show that the applied loading pressures (σ_L) heavily modify the stress and strain fields in the s-y planes. Particularly, deviatoric stresses significantly increase and shear strains mainly concentrate along special directions that can be related to the stresses acting along the longitudinal s-direction and lateral y-direction. Yielding conditions were predicted by the numerical analysis over large portions of the slope, also for low pressures (< 20 kN/m) applied at the

upper boundary of the s-y planes, essentially due to the initial low lateral stresses (σ_y) along the y-direction. However, the lateral enlargement of the yielded zone is related to the stresses acting along both longitudinal and lateral directions. High longitudinal stresses (σ_s) predispose the arising of yielding conditions while initial high lateral stresses (σ_y) restrain the effects of the applied load.

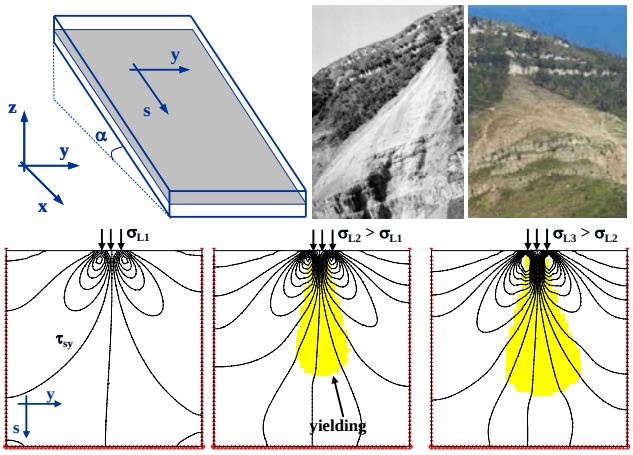


Figure 6. Stresses variations due to pressures applied at the upper boundary of a plane parallel to ground surface.

4. CONCLUSION

When involving open slopes, landslides of the flow-type often occur in triangular source areas where initial slides turn into avalanches through further failures and/or soil entrainment. Significant examples of this landslides typology were recorded in May 1998 in the pyroclastic deposits overlying the carbonate Pizzo d'Alvano massif (Southern Italy, Sarno-Quindici event). Referring to this event, the paper analyses two main triggering factors consisting in karst spring from the bedrock and impact of failed soil masses on stable deposits.

The failure stage is studied through both hydro-mechanical uncoupled and coupled approaches based on groundwater modelling, limit equilibrium and stress-strain analyses. Karst springs at the base of bedrock scarps are pointed out as a severe hydraulic boundary condition. In fact, the induced transient pore water pressures cause the failure onset mostly depending on the stratigraphy, even for short durations of the springs. Referring to the impact phenomena, the failure onset is not simulated assuming drained conditions. Conversely, in undrained conditions, the increase of pore water pressures determines the failure onset that is satisfactorily justified by limit equilibrium analyses. Analogous scenarios are obtained through stress-strain analyses that also point out the high decrease of mean effective stresses and the possibility of liquefaction phenomena.

The post-failure stage is analysed through the previous uncoupled approach based on simplified assumptions and stress-strain analyses. The achieved results outline that failure induced by karst springs from bedrock and/or impact loading can cause further failures in downslope stable deposits. Particularly, low pressures downwards acting along planes parallel to ground surface lead to yielding conditions in large portions of the hillslope whose extension and lateral enlargement depend on lateral stresses, depth and slope angle.

REFERENCES

Bilotta, E, Cascini, L., Foresta, V., Sorbino, G., Geotechnical characterization of pyroclastic soils involved in huge flowslides. *Geotechnical and Geological Engineering, 23*, 365-402, 2005.

Budetta, P., de Riso, R. The mobility of some debris flows in pyroclastic deposits of the northwestern Campanian region (Southern Italy). *Bull Eng Geol Environ, 63*, 293-302, 2004.

Calcaterra, D., de Riso, R., Evangelista, A., Nicotera, M.V., Santo, A., Scotto di Santolo, A. Slope instabilities in the pyroclastic deposits of the carbonate Appenine and the phlegrean district (Campania, Italy). *Proc. of the Int. Workshop "Flows 2003", Sorrento, Patron Ed.*, 61-75, 2004.

Cascini, L., Cuomo, S., Guida, D., Typical source areas of May 1998 flow-like mass

movements in the Campania region, Southern Italy, *Engineering Geology, 96*, 107-125, 2008.

Cascini, L., Cuomo, S., Sorbino, G., Flow-like mass movements in pyroclastic soils: remarks on the modelling of triggering mechanisms. *Italian Geotechnical Journal, 4*, 11-31, 2005.

Cascini, L., Guida, D., Nocera, N., Romanzi, G., Sorbino, G., A preliminary model for the landslides of May 1998 in Campania Region. *Proc 2nd Int. Symposium on Geotechnics of Hard Soil-Soft Rock – Napoli, Balkema, 3*, 1623-1649, 2000.

Cascini, L., Sorbino, G., Soil suction measurement over large areas: a case study. *Proc. 3rd Int. Conference on Unsaturated Soils, Recife (Brasil), Balkema 2*, 829-834, 2002.

Chen, H., Crosta, G.B., Lee, C.F., Erosional effects on runout of fast landslides, debris flows and avalanches: a numerical investigation. *Geotechnique, 56(5)*, 305-322, 2006.

Costa, J.E., Williams, G.P., Debris-flow dynamics (video tape): U.S. Geological Survey. Open-File Report 84-606., 1984.

Crosta, G., Imposimato, S., Roddeman, D.G., Continuum numerical modelling of flow-like landslides, *Landslides from Massive Rock Slope Failure, Evans et al. (eds.), Springer*, 211-232, 2006.

Cuomo, S., Geomechanical modelling of triggering mechanisms for flow-like mass movements in pyroclastic soils. *PhD dissertation at the University of Salerno, Italy*, pp. 274, 2006.

Di Crescenzo, G., Santo, A., Debris slides-rapid earth flows in the carbonate massifs of the Campania region (Southern Italy): morphological and morphometric data for evaluating triggering susceptibility. *Geomorphology, 66*, 255-276, 2005.

Fiorillo, F., Wilson, R.C., Rainfall induced debris flows in pyroclastic deposits, Campania (Southern Italy). *Engineering Geology, 75*, 263- 289, 2004.

Fredlund, D.G., Morgenstern, N.R., Widger, R.A., The shear strength of unsaturated soils. *Canadian Geotechnical Journal, 15*, 313-321, 1978.

Geoslope, User's guide. GeoStudio 2004, Version 6.13. Geo-Slope Int. Ltd. Calgary, Canada, 2004.

Gerstell, M.F., Aharonson, O., Schorghofer N., A distinct class of avalanche scars on Mars, *Icarus 168*, 122-130, 2004.

Guadagno, F.M., Forte, R., Revellino, P., Fiorillo, F., Focareta, M., Some aspects of the initiation of debris avalanches in the Campania Region: the role of morphological slope discontinuities and the development of failure. *Geomorphology, 66*, 237-254, 2005.

Guadagno, F.M., Martino, S., Scarascia Mugnozza, G., Influence of man-made cuts on the stability of pyroclastic covers (Campania-Southern Italy): a numerical modelling approach. *Environmental Geology, 43*, 371-384, 2003.

Henkel, D.J., The shear strength of saturated remoulded Clay. *Proc. of the ASCE Research Conference on Shear Strength of Cohesive Soil, Boulder, Colorado,USA*, 533-544, 1960.

Hungr, O., Evans, S.G., Bovis, M.J., Hutchinson, J.N., A review of the classification of landslides of the flow type. *Environ. and Eng. Geosci., VII (3)*, 221-238, 2001.

Hutchinson, J.N., Bhandari, R.K., Undrained loading, a fundamental mechanism of mudflow and other mass movements. *Geotechnique, 21[4]*, 353-358, 1971.

Jakob, M., Hungr, O., Debris-flow Hazard and Related Phenomena, Springer, pp. 739, 2005.

Jamiesion, B., Stethem C., Snow Avalanche Hazards and Management in Canada: Challenges and Progress, Natural Hazards 26, 35-53, 2002.

Pastor, M., Quecedo, M., Fernandez Merodo, J.A., Herreros, M.I., Gonzalez, E., Mira, P., Modelling tailing dams and mine waste dumps failures. Geotechnique, 52(8), 579-591, 2002

Pastor, M., Haddad, B., Sorbino, G., Cuomo, S., Drempetic, V., A depth integrated coupled SPH model for flow-like landslides and related phenomena. Int. Journal Numerical and Analytical Methods in Geomechanics, DOI: 10.1002/nag.705, 2008.

Pielmeier, C., Schneebeli, M., Developments in the stratigraphy of snow. *Surveys in geophysics, 24*, 389-416, 2003.

Wang, F.W., Sassa, K., Fukuoka, H., Downslope volume enlargement of a debris slide-debris flow in the 1999 Hiroshima, Japan, rainstorm. Engineering Geology, 69, 309-330, 2003.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A Semantic WEB Services Platform to support Disaster and Emergency Management

Maurizio Gloria [a],Valentina Lersi [a], Giovanni Minei [a], Diana Pasquariello [a],

Cristiano Monti [b], Antonio Saitto [b]

[a] MARS Center S.r.l., Via E. Gianturco, 31, 80146 Napoli - Italy
({gloria/lersi/minei/pasquariello}@marscenter.it)
[b] Telespazio S.p.A, Via Tiburtina, 965, 00156 Roma – Italy
({guest297.monti/antonio.saitto}@telespazio.com)

***Key words*:** Agent Platform, Jade, Ontology, Sesame, Semantic Web, Distributed Systems, Emergency Management

Abstract: The importance of telecommunications services for the management and control of critical situations dedicated to security of citizens has always been recognized: in particular, in this work we describe the solution designed and developed in an Italian research project, called CI6 (Centro Integrato per Servizi di Emergenza Innovativi). Focus of this project was to develop an integrated centre able to provide important informations concerning the state and possibile evolution of the crisis for decision makers. The knowledge base is managed by means of homogeneous schema to represent information (ontology of domain) and by means of semantic rules defined by the analysis of the domain. The communication layer has been provided by the CHIMERA platform we have developed, which is able to connect analogical / digital local radio network and satellite system for remote communications as well as to provide fast availability of communication and broadband using legacy equipment.

Keywords: Agent Platform; Jade; Ontology; Sesame; Semantic Web; Distributed Systems; Emergency Management.

1. INTRODUCTION

During crisis events, operators and decision makers work in difficult and complex environments that could require, in order to take decision, rapid exchange of informations as video, images, weather information, resources already engaged for emergency and resource yet available, and so on. The correlation of data can play a key role in preventing or immediately recognizing a critical event and managing the critical situation in the most effective way. Indeed, the importance of telecommunications services for the management and control of critical situations for the security of citizens has always been recognized [Mrozinski, 1975].

This paper presents a multi-layered platform (CI6) that supports data gathering and aggregation coming from various sources to produce value added information (say CI6 information). The goal of this project consists in developing a decision support system to coordinate and to simplify the management of critical situation. CI6's information domain is very rich; for example: mobility information, meteorology, current resources availability, cartography, textual communication with operators, information about rescue team and their equipment, stratification of different data, monitoring of emergency evolution, unicity of support system.

A hypothetical scenery could be a sea emergency, for example: a broken-down ship gives an alarm by satellite terminal interfaced by CI6. CI6 will be able to visualize alarm and its properties (typology, localization, vehicle's identification and so on..). A CI6's user can

visualize the alarm and can decide to manage it. An occurring alarm can be managed both through a predefined scenery, and creating a new one. When user creates a new scenery, the CI6' s decision support system suggests the possibility that two or more alarm could be about the same critical situation. Both creation of a new scenery and the selection of an existing one allow to use all services and functionalities before described.

2. MULTI-AGENT SW MIDDLEWARE ARCHITECTURE

2. 1 Introduction

The proposed architecture is a 3-tiers one: "3-tiers" is a client-server architecture where the user interface, functional process logic, data storage and data access are developed and maintained as independent modules.

This is a layered architecture, in which layers close to the user are built upon lower layers closer to the machine. In GUI and data-intensive applications like CI6 system, the separation (and distribution on different computing nodes) of the data management layer from the presentation and application logic layers provides concrete advantages allowing the replacement of one tier without affecting the others. Moreover, it makes easier to implement load balancing. The most volatile parts of the system (in case of changes in requirements) will be the user interface and the system's business rules. By separating user services from other services, the system's user interface can be changed without impacting the rest of the application. Similarly, by separating business services from other services, it's easier to change the business rules of the system with minimal impact to the rest of the system.

2.2 Middleware architecture

In CI6 project data layer is formed by all existing information sources; the business services layer implement the system's business rules; the user interfaces layer formats the information and presents it to the user (Figure 1).

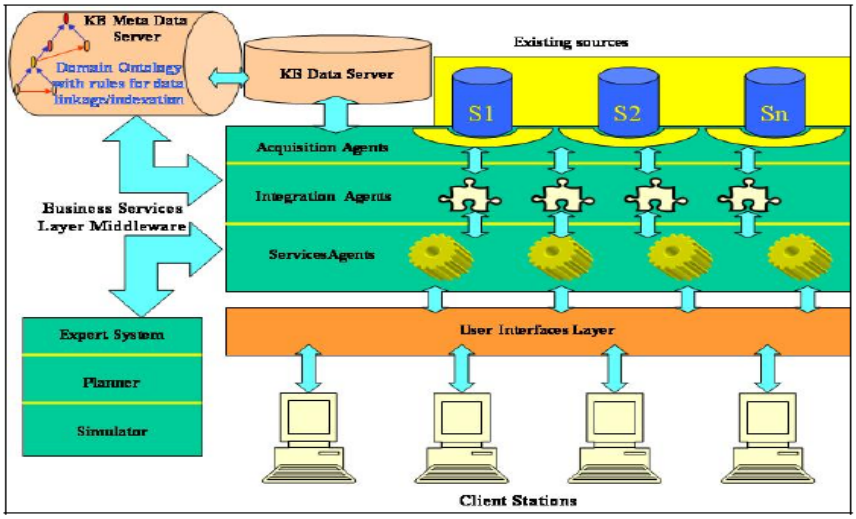


Figure 1. 3-tiers Multi-Agent SW Middleware Architecture

The data layer is composed by the available sources. The proposed architecture foresees the definition of a Knowledge Base (KB) that contains the whole knowledge related to the domain. The suggested approach to define the KB is based upon an ontological model and takes into account the most recent methodologies of knowledge representation and fruition.

A KB defined and implemented after this model offers more flexibility of use and an enhanced capability to find the useful information. Each domain object is represented at a first level in a taxonomy, a hierarchical structure of object's categories; the taxonomy, enriched with a set of functional relationships, constitutes the domain ontology.

The KB of the proposed architecture consists of two primary modules: **KB Data Server**: is the actual place were the data are stored; **KB Meta Data Sever:** uses the ontology for the classification of data and rules for the creation of a semantic relations among pieces of information.

The business services layer is defined as an intelligent middleware based on a MAS. In Multi-Agent Systems (MASs) the agents are considered to be autonomous entities, capable of modifying the way in which they achieve their objectives. Main characteristics of a MAS can be summarized as it follows: (1) each agent has incomplete information or capabilities for solving the problem (thus, has a limited viewpoint); (2) there is no system global control; (3) data/knowledge are distributed; and (4) computation is asynchronous. Such features makes a system fault tolerant (e.g.: a system is still able to work when a part of it is out of action), increased in flexibility and in adaptability. As a MAS, we used the JADE (Java Agent Development framework) environment [Di Napoli, 1998].

Implemented Jade agents categories comprehend "acquisition agents", "negotiators" or "service agent", that allow a user to access to suitable services during a management of emergencies [Di Napoli, 1998], [Bellifemmine, 2004], [Minei, 1997].

Starting from emergencies management domain, on the basis of the identified user needs an analysis activity has been focused on: 1) Identify and analyse the existing resources identifying existing data, associated services and available interfaces. 2) Identify new services and additional resources which will be useful to implement the cooperative approach. 3) Identify the service flow and the fruition of the established body of knowledge with an user oriented approach.

In our approach the service-oriented middleware platform is based on a 3-tiers architecture, composed by: 1) Acquisition Layer 2)Integration Layer 3)Service Manager. Each layer has been defined through a multi-agent system based on JADE paradigm. In particular:

- The acquisition layer has the task to interface the system with the available resources. The service-oriented architectural approach ensures the maximum amount of interoperability, as the information and services integration are platform independent. This approach proves very effective in the case that an additional resource should be integrated, so a real modularity is guaranteed to the system.
- The integration layer has the objective to organize the information associated with the resources. Its task is accomplished through an explicit representation schema, making use of a domain specific ontology and semantic integration rules. The ontology is the reference schema for the development of a knowledge base
- The service manager is based on a s/w agents' technology and guarantee the access to all the services of the system. The services are defined at this level and the s/w agents' community negotiate at this level the resource availability on the basis of users requests.

Another important layer of the proposed architecture is the User Interfaces. To make the CI6 working in optimal condition, the interface shall contain some important features: **adaptability** to data / user; **adaptivity** to different scenarios.

The adaptability and adaptivity are interface properties that offer a different typology of interaction (functions and modality of access) to different data types. Different data profiles are internally defined. The adaptability to the data is necessary because different types of data are defined; also a user adaptivity based on different user typology can be proposed. The suitability of the CI6 user interfaces is necessary due to the different scenarios in which the potential users could operate. The components are: **The Observation Module** – acquires the data from the user observing his behaviour (if the system is in adaptivity mode), from the user himself or acquiring them from context in which the user operates.

The Presentation Module – selects contents and modalities to display the information driven by the data acquired from the Observation Module. **The User Modelling Engine** – selects the user profile following the observation of the user and of the operative context. The engine can have additional components like inferential rules in case the acquired data is not perfectly adapted to the stored situation profile.

3. SEMANTIC ASPECTS: SESAME AND JADE INTEGRATION

The proposed architecture is based on the integration of JADE (for agent developed) and SESAME (for the management of relation among elements of domain). The middleware has the advantage to use both a software agents architecture and structured domain by means of RDF and RDF Schema (semantic net).

The domain of CI6 is composed by a set of resources, each associated to a set of properties. Two of these are strictly required, for they identify the resource and the related services. Other properties link the resource with other resources of the system. The two strictly required properties are: **HasService**: points to the description of a functionality belonged to resource and available to users; **HasSource**: points to the physical location of resource (for example: database, tool S\W, Web Services, etc.).

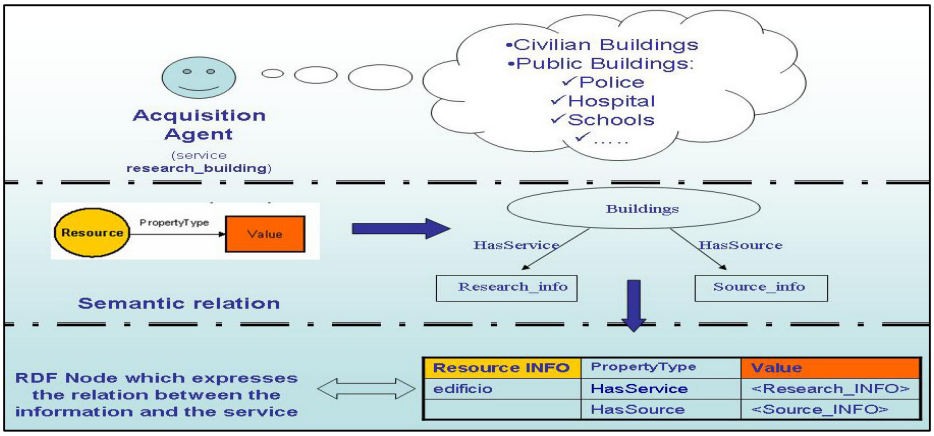


Figure 2. Schema of "collaboration" between Jade agent and a semantic net.

Let's introduce an example. The resource "Map Server" offers functionalities such as: selection of maps, zoom on maps, size of distance, geo-references information, etc. A physical address is associated to the resource. A form of "collaboration" between JADE agent and semantic net managed by SESAME is introduced in (Figure 2): 1) A jade agent has knowledge (ontology) about the particular area of domain it has to interact with. 2) The semantic relations associated to each node belong to (area subset of) the net of relations that describes the CI6 domain. 3) The structure of RDF node that belongs to semantic net managed by SESAME.

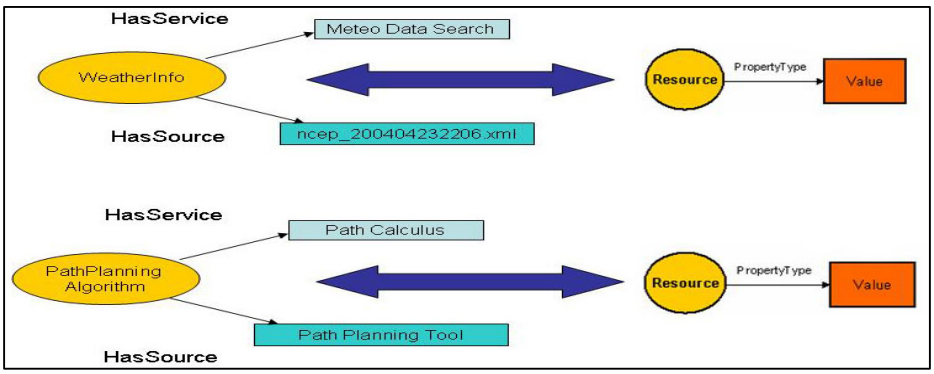


Figure 3. Example of resources "Weather" and "Path Planning Algorithm"

In (Figure 3) we can see the resources: “weather information” and “path planning algorithm”.

To explain the phases of interaction between JADE agent and semantic net, we use a simple example: the request, by a JADE agent, of a specific service.

CI6's architecture involves five entities (Figure 4): a Service agent; a DF agent; an acquisition agent; Sesame; one or plus database.

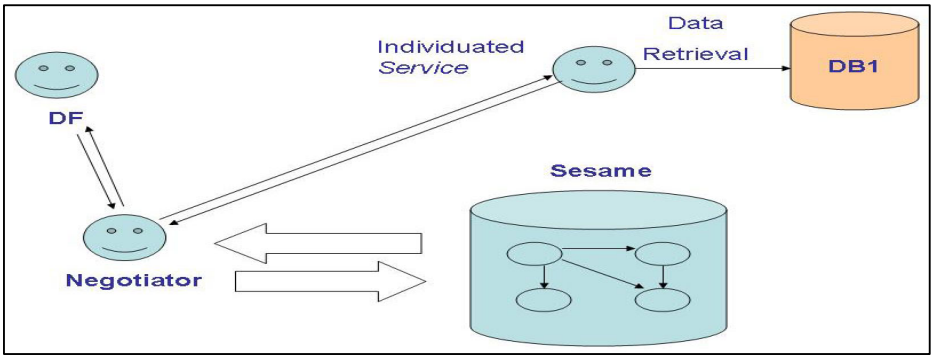


Figure 4. Entities involved in interaction between Jade agent and semantic net

The process of querying the semantic net by agents pass through the following steps (Figure 4): ***Interrogation***: Service agent forwards the request to SESAME to know the necessary information. ***Result***: by the interrogation, the service agent knows the value of relations “HasService” and “HasSource” (and of other eventual relations). ***Found Acquisition agent*** *can access to service*: Service Agent forwards to DF Agent the request to know the address of the application agent that can access to service; ***Access to information***: the Service Agent requests the service to Acquisition Agent and forwards it the source's address. The Acquisition Agent accesses to database and sends the information to service agent.

As an example of application of CI6's architecture, we briefly describe the service we called ***stratification service.*** Information about cartography, street map, weather, building maps, paths calculated by the integrated *path planning service*, are dynamically integrated on request. It’s possible to extract information according to the desiderated level of stratification.

In this way the user (the Emergency Room operator) can have an integrated view of the Emergency scenery, in order to choose the best solutions to face the emergency.

The service agent associated to the stratification service (the stratification agent) queries the Sesame repository according to the illustrated schema: ***Interrogation of semantic net***: Stratification Agent forwards to SESAME the request about the information that user (operator) intends to integrate. In fact relevant information vary according to the different Emergency scenarios. ***Result of query***: for example, the Stratification Agent knows, among other information, the URL of the Web Service to obtain the cartography and the Acquisition Agent’s name associated to the cartography resource. **Found Acquisition Agent can access to service**: Stratification Agent forwards to DF Agent the request to know the address of, for example, the Acquisition Agent for Cartography (the Cartography Service Agent), the Acquisition Agent for Weather (the Weather Service Agent), and so on. **Access to information**: the Stratification Agent requests cartography information to the Cartography Service Agent, the weather information to the Weather Service Agent and so on. When all acquisition agents respond to the Stratification Agent, it collects information and shows them to the user (the emergency operator).

For example, in the case of the occurrence of an emergency in via Gianturco, Naples, by means of the stratification service the operator can invoke a path planning algorithm to

calculate the best evacuation plan in function of the integrated information about traffic conditions, weather conditions and GPS signals of emergency operators (Figure 5).

The important thing to notice here is that the best response about evacuation plan is possible only on the basis of an integrated information, such as the information given by the stratification service. In fact, analysing only a partial information, only non-optimal evacuation plans could be calculated.

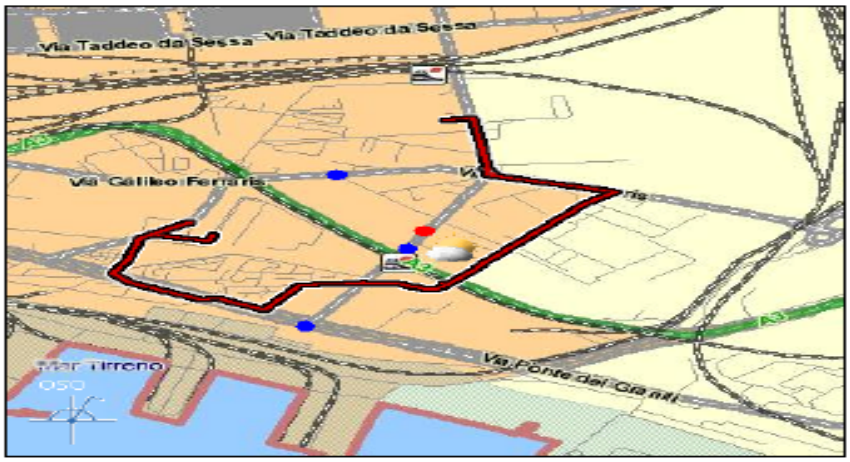


Figure 5. Example of Stratification Service

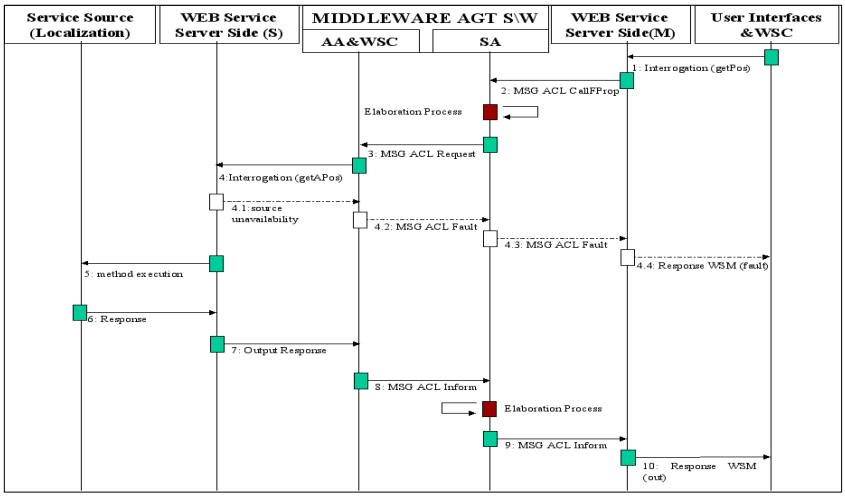


Figure 6. Flow of interaction of module of CI6

In Figure 6 is showed the flow of interaction of module of CI6 to obtain localization service.

4. COMMUNICATION MIDDLEWARE: THE CHIMERA PLATFORM

CHIMERA [Luglio, 2007] is an IP based platform implemented to support and achieve voice and data interoperability among the different operative teams on-field, who usually use heterogeneous radio technologies (i.e. HF, VHF, UHF, TETRA, etc.) and the emergency management infrastructure located in the Control Centre. Figure 7 shows the CHIMERA concept that is based on tree main levels: Terrestrial segment; Satellite segment; IP Gateway.

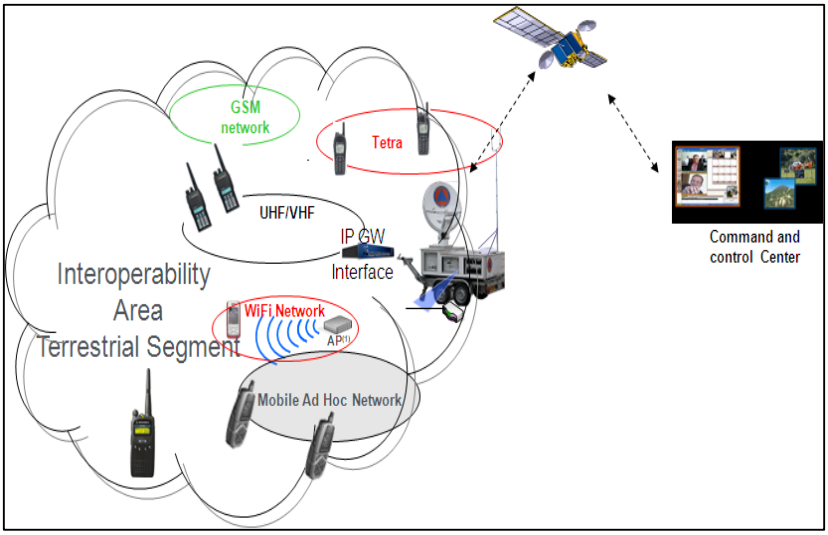


Figure 7 CHIMERA communication system

Terrestrial networks as WiMax, WiFi, Sensor Networks, MANET, TETRA and/or analog radio can be used. Moreover, this approach allows an easy way to set up interface with satellite systems. In particular, two types of satellite communications can be interfaced. A narrowband mobile terminals (GSM-like) to allow voice and low bit rate data communications from team members to any other location located even very far away, and a broadband terminals (ensuring IP interface) very small and light, easy to transport and to install when an even small size van is available and located nearby the rescue team.

The IP gateway is used to interface the terrestrial radio network and the satellite system. In particular, the Gateway is in charge to adapt the voice signal to IP protocol and is interfaced with the heterogeneous system. It is able to convert over IP format the data and voice coming from the terrestrial network. In this way, through the satellite system, it is possible to connect the network to another remote network, to a remote LAN or to interconnect users belonging to different networks ensuring the interoperability among different teams operating on the field. The IP Gateways are physically linked to radio devices and convert the voice and data radio communications to VoIP communications according to the standard protocol (H323 and SIP). The system may be deployed on a network already in use by other IP services or on a newly created network. The IP network may be logically divided into functional subnetworks dedicated to specific services thus dividing Radio Over IP services from other IP services. Otherwise the system may be deployed over a dedicated IP network. The physically or logically dedicated IP network has to be configured in order to support all the needed simultaneous voice communication links. The IP Gateways also allows interoperability between different radio networks direct linked to the same or to different Gateways. They can extend the radio communication capabilities through all the IP network. Through the Gateways half duplex and full duplex voice communications are available. They make possible data transmission and remote radio control depending on radio devices capabilities. Due to the integration between IP networks and radio devices, CHIMERA enables the communication among the Decision Support System, Decision Maker and teams on field using IP phones, handheld devices, PDA or also radio terminals.

5. CONCLUSIONS

The domain of emergency management is a complex area because it requires dynamism and flexibility in taking decisions. The decision makers need a support system to have current and complete information to support their decision and to know the available resources for managing the emergency.

These requirements are satisfied by our middleware based on software agent which, using a semantic layer, is able to give complete, flexible and significant answers to user's request.

In fact, although CI6 is still a prototype, it satisfied the tests.

We can mention test with port authorities of Palermo and civil security of Naples among more important tests.

In first case, we simulated a sea emergency as that described in introduction while, in the second case, we simulated the management of evacuation of a city area because of flood.

The result of simulation was positive in both cases; in fact although emergencies share common aspects in their being faced, the emergencies problem, 'cause of its being complex, intrinsically involves the presence of features specifically related to each situation. For example, one of the feature of a sea emergency is the difficulty of communication, that increases increasing the distance from the coast, and that arises problems both for the signal of alarm and for communication during emergency management. In this case, the use of a satellite terminal reveals as a good choice for this type of emergency.

In the second case, it has revealed important to have a 'path planning' service as a strategic tool for the decision support system.

The CI6's architecture responds in a satisfying manner to the requirements arisen by the problem of facing emergencies; in fact the possibility to have a semantic layer (semantic net) and a layer able to elaborate simple data to produce complex information (middleware based on agent), has revealed as an appropriate instrument to manage the complexity and the variety of the analysed domain. Moreover, being modular and extensible, the CI6's architecture offers the capability to integrate other services and functionality.

Future work may be prosecuted towards the automation of the emergency type identification and, as a consequence, in the automatization of the decision about the degree of integration required for facing emergencies.

REFERENCES

Mrozinski, R., The Application of Telecommunications to City Services, *IEEE Transactions on Communications*, 23(10), 1080-1084, October 1975.

JADE (Java Agent DEvelopment Framework), http://jade.tilab.com/

Bellifemmine, F., Caire, G., Rimassa, G., The JADE multi-agent systems platform: the first six years, *Intelligenza Artificiale*, 4(1), Dicembre 2004

Di Napoli, C., Mango Furnari, M., Mele, F., Minei, G., (1998), Explicit time management for agent internal and social activities, Proceedings of the Workshop “Intelligent Agents in Information and Process Management”, Bremen, Germany, September 1998

Minei, G., Palladino, M., Magopl: a language for multiagent oriented programming, Proceedings of the Workshop “Advances in User Modelling”, Sixth International Conference on User Modelling, Chia Laguna Sardinia 2-5 June 1997

Luglio, M., Monti, C., Roseti, C., Saitto, A., Segal, M., Interworking between MANET and satellite systems for emergency applications, *International Journal of Satellite Communications and Networking*, 25(5), 551-558, Settembre/Ottobre 2007.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Drought Mitigation using Operative Indicators in Complex Water Systems

G.M. Sechi[a] **and A. Sulis**[b]

[a] *Hydraulic Sector, Dept. of Land Engineering, University of Cagliari, Italy (sechi@unica.it)*

[b] *Hydraulic Sector, Dept. of Land Engineering, University of Cagliari, Italy (asulis@unica.it)*

Abstract: The definition of an effective link between drought indicators and drought mitigation measures in complex water systems is a tricky problem. In our view this gap between research and practice still limits the applicability of measures more than institutional or technological aspects. In this paper a methodology is presented with the aim of supporting water Authorities in the decision making process to face droughts in complex water systems. The methodology is a full integration of optimization and simulation tools. The exploratory power of the optimization allows to fast estimate drought risk indicators under future hydrological scenarios. The optimization model uses these indicators as triggers of measures in a pro-active approach. Then, a more robust and detailed simulation is used to test and evaluate these set of measures. Particularly in the case of too optimistic hydrological forecast, the pro-active approach does not completely eliminate the risk of drought and further measures in a reactive approach have to be implemented in the simulation. They include expensive and big impact measures to be taken later, during the drought event. Thanks to the collaboration with the regional water Authorities in South Italy, the proposed methodology has been currently under test in the Agri-Sinni water system. Early applications to the Agri-Sinni water system are presented in the paper showing the usefulness of the proposed methodology in mitigating the impacts of drought and selecting an economic efficient combination of pro-active and reactive measures.

Keywords: Drought; Pro-active Measures; Complex Water Systems; Operating Rules; WARGI

1. INTRODUCTION

Particularly in water resource systems that frequently experience severe drought events, the definition of drought mitigation measures becomes a central aspect in the systems planning and management. Even if the usual approach is still to manage the water emergencies rather than to prevent them through a set of measures coherently developed in a pro-active approach, it seems that something is changing in legal and institutional framework. At this time, in our view a practical application of mitigation measures to complex water systems is limited by a nebulous link between these measures and drought indicators.

Different water system state indicators could be used as a trigger to start measures. For example, the reservoir storages or the total available water (reservoir storage plus inflow) measured in key reservoirs at the end of the wet season shall trigger the implementation of measures in systems where supplies come from superficial reservoirs. Nevertheless, the traditional approach requires to use iteratively a simulation model to define a monthly probability distribution of satisfying predefined demand scenarios (or level of alert) based on the current values of drought indicators in the system. At times the definition of this matrix of probabilities can be long and difficult, and a high number of potential

combinations have to be compared. Also, in this approach mitigation measures are statically implemented as each demand scenario contains itself a set of measures to satisfy the demand. When the matrix is defined, simulation model is then used to test the time and the level of alert to be declared.

To improve the definition of an effective link between drought indicators and drought mitigation measures in complex water systems, a full integration of the simulation model WARGI-SIM (Water Resources Graphical Interface – Simulation Tool) and the linear optimization model WARGI-OPT has been recently developed (Sechi and Sulis, 2007) with the aim of identify and evaluate mitigation measures in a proactive approach that anticipate drought impacts. The pro-active approach includes measures implemented prior to the drought event with the aim of reducing the system vulnerability. Yevjevich et al. (1978) classified drought mitigation measures in three main categories: supply-oriented measures, demand-oriented measures and impact-minimization measures. While the impact-minimization measures are basically related to water users and various factors, which can minimize the economic, environmental and social impacts, supply and oriented demand measures are intended to reduce the risk of water scarcity. The mixed optimization-simulation approach aims to implement these two categories of measures (supply-oriented measures and demand-oriented measures) in a proactive approach considering a predefined infrastructural configuration of the water system.

In Sechi and Sulis (2007) the mixed optimization-simulation approach has been presented and applied to the South Sardinia (Italy) water system. Thanks to the regional water Authorities collaboration within a national research project (PRIN, 2005), the proposed approach has been currently under test in the Agri-Sinni (Southern Italy) water system. Preliminary results of the application to this complex water system are presented in the paper. The Agri-Sinni is a multireservoir and multiuse system that has frequently experienced extreme drought events in last decades. Results prove the usefulness of this approach in supporting water Authorities in the decision making process to face droughts in complex water systems.

2. OPTIMIZATION AND SIMULATION APPROACH

Despite difficulties in expressing all the system operations and constraints in a mathematical form, WARGI-OPT views the entire system as a single problem in time and space. This exploratory power of the optimization allows to dynamically defining a set of mitigation measures under different future hydrological scenarios. Then, WARGI-SIM is used to test and evaluate these set of measures. Particularly in the case of too optimistic hydrological forecast, the pro-active approach does not completely eliminate the risk of drought and measures in a reactive approach have also to be implemented in the simulation. The reactive approach includes more expensive and strong impact measures to be taken later, during the drought event, without reducing the system's vulnerability to future drought events. The effectiveness of these temporary actions may vary with the measures already implemented in the pro-active approach.

2.1 Pro-active measures in the mixed optimization-simulation approach

In the analysis of a water system on a time horizon T with a time step t (Figure 1), at each synchronization period τ_i in the simulation process, WARGI-OPT forecasts system evolution on a time horizon Δ based on current water storages in the system's reservoirs and a user-selected future hydrological synthetic scenarios. When dealing with hydrological uncertainty, the deterministic optimization method in WARGI-OPT can be implemented implicit stochastic environment (Hiew et al., 1989) with equally likely future hydrological scenarios.

WARGI-SIM uses the information provided by WARGI-OPT at each synchronization period τ_i as triggers of mitigation measures in a pro-active approach ($[z_\tau]$). These pre-emptive measures can modify system management in the subsequent periods until τ_{i+1}. As well known, the simulation time horizon T should be extended several decades in order to obtain a correct estimation of system performance.

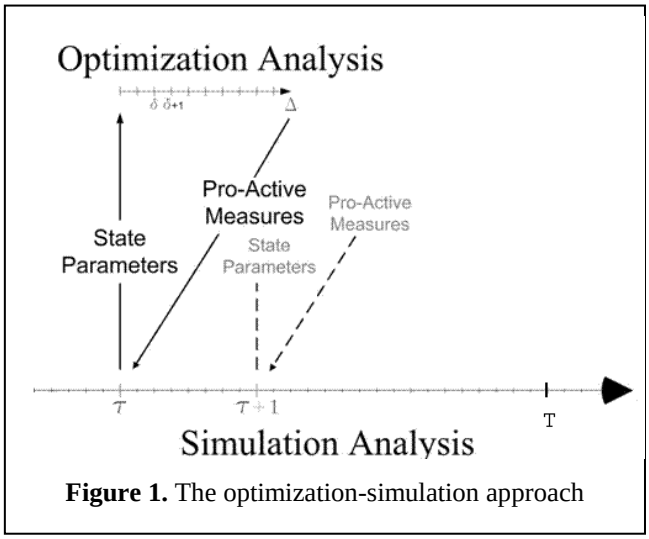


Figure 1. The optimization-simulation approach

The definition of the optimization time horizon Δ, as well as the hydrological scenario weight, and costs or penalties associated with pre-emptive measures are key aspects in the approach proposed (Sechi and Sulis, 2007).

2.2 Reactive measures in the simulation model

When pre-emptive measures $[z_\tau]$ do not make it possible to overcome the water scarcity in the simulation phase, WARGI-SIM introduces mitigation measures in a reactive approach ($[s_t]$). Unlike other simulation models designed to describe system behaviour using complex specific algorithm rules embedded in the code, WARGI-SIM defines reactive measures $[s_t]$ based on preferences and priorities (Sechi and Sulis, 2007). For each strategic reservoir, the user defines a reserved volume as a function of the period of the year. When storage volume is within the reserved zone, downstream flows are decreased to satisfy few demands only. Strategic reservoirs and priority demands are selected by the user. In such cases, based on a hierarchical list of combination of resources and demands supply flows could be activated to meet non priority demands from alternative or marginal resources, or temporary restrictions could limit some non essential uses in these demands.

2.3 Response function in the mixed approach

The goal of the proposed approach is to define the best combination of drought mitigation measures that minimize the response function R. In the mixed optimization and simulation approach, R can be written as the sum of costs associated to systems planning, OMR (operative, maintenance and replace costs) and mitigation measures:

$$R = C_\gamma Y + \sum_{\tau_i} C_{PD} z_\tau + \sum_{t=1}^{T} C_{NPD} s_t + \sum_{t=1}^{T} \min\left(C_{OMR} x_t\right) \qquad (1)$$

The set $[z]_\tau$ represents drought mitigation measures in a proactive approach obtained by the WARGI-OPT. Particularly in the case of water scarcities more severe than those forecasted by WARGI-OPT, the response function R includes measures $[s_t]$ in a reactive approach. The set of costs $[C_\gamma]$ is related in the simulation to project variables $[Y]$ as defined by WARGI-OPT. $[C_{OMR}]$ represents operative, maintenance and replace costs, associated with flow variables $[x_i]$. $[x_i]$ is the solution of a minimum cost flow problem between resource and demand nodes in WARGI-SIM. $[C_{PD}]$ and $[C_{NPD}]$ are costs respectively associated with drought mitigation measures in a proactive and reactive approach. $[C_{PD}]$ and $[C_{NPD}]$ include OMR costs for drought measures, agency income lost from reduced water sales, and reduced consumer surplus due to these measures.

The consumer surplus has been estimated by the willingness to pay method (Lund, 1995). When supply is less than users' maximum water demand, the consequent loss in total benefit value is given by the following integral formula (James and Lee, 1971):

$$D = \int_{Q_r}^{Q_s} P_s \cdot \left(\frac{Q_s}{Q} \right)^{\frac{1}{E}} dQ \qquad (2)$$

where Q_s indicates request for complete demand satisfaction, P_s the resource's marginal cost under conditions of completely fulfilled demand, Q_r reduced supply during the drought period, E the price elasticity of demand. Clearly, E is very low for demands in high priority (e.g. urban uses), where demand varies only slightly in relation to price; also, the value of elasticity in a short-term demand curve is lower than that in a long-term demand curve. Seasonal effects on water demand are not included in this approach.

3. APPLICATION TO THE AGRI-SINNI WATER SYSTEM

The Agri-Sinni water system (Figure 2) is located in Basilicata (South Italy) supplying demands also in Puglia and Calabria. The main reservoirs are Monte Cotugno (capacity of $556 \cdot 10^6$ m^3) and Persusillo (capacity of $159 \cdot 10^6$ m^3) along the Sinni and Agri River respectively. Marsico Nuovo and Cogliandrino are single purpose reservoirs (respectively for irrigation and hydroelectric use) with small regulation capacities. Four intake structures (Agri, Sarmento, Sauro and Gannano) were constructed on the main rivers for diversion of water.

Based on the observed monthly inflows at Monte Cotugno and Pertusillo over the period 1983-2005, inflows in the other sections of interest in the basin were generated. Inflows series reproduce the intense water scarcities in the Agri-Sinni that occurred in the period 1989-1990 and 2001-2002. Table 1 shows the main properties of the series.

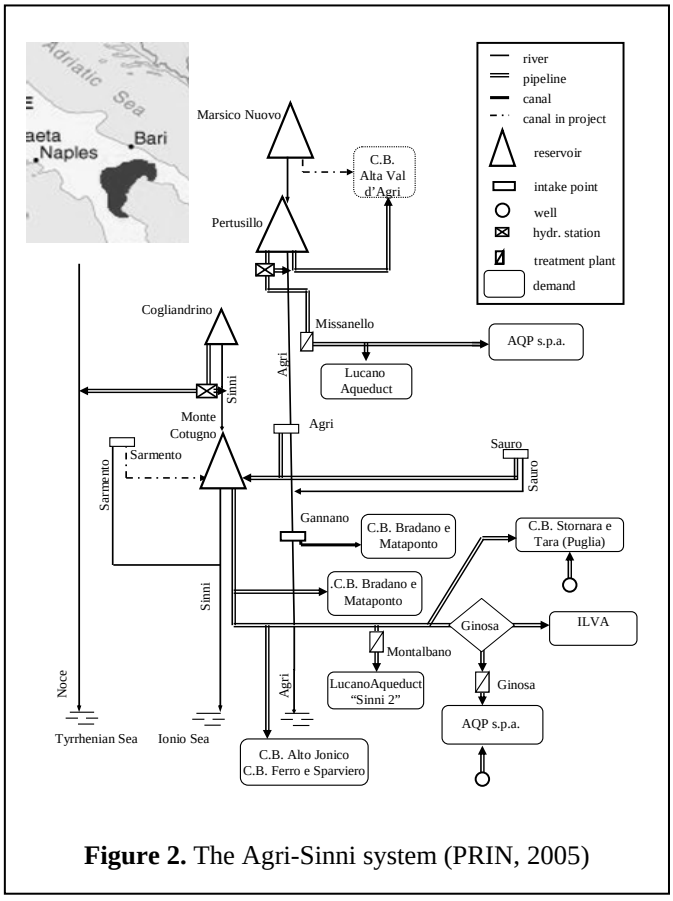


Figure 2. The Agri-Sinni system (PRIN, 2005)

Urban, industrial (ILVA in Figure 2) and irrigation demands (C.B. in Figure 2) are respectively $246.5 \cdot 10^6$ m^3/yr, $12.6 \cdot 10^6$ m^3/yr, and $240 \cdot 10^6$ m^3/yr.

The Agri-Sinni is a system often affected by intense water scarcities. Figure 3 shows the volume released from the Monte Cotugno reservoir for urban, industrial and irrigation uses over the period 1993-2003. During one of the most serious drought occurrence (2001-2002) water restrictions, mainly imposed on irrigation consumption, were equal to 50% of the average supplies over the period 1993-2000.

Numerous criteria could be used to summarize system performances (Hashimoto et al., 1982). Here, vulnerability (percentage value of maximum annual reduction on the total value of annual demand), time reliability (percentage value of the months in which reduction is less than predefined thresholds), volumetric reliability (percentage value of reduction on demand) have been used to quantify demand reductions. In this application to the Agri-Sinni system, mitigation measures have been finalized to preserve the complete satisfaction of priority demands for urban and industrial water requirements. Consequently, only the evaluation of system performances for irrigation use is reported when measures are implemented to face drought events.

Table 1. Statistical indexes of inflows in the period 1983-2005.

Stations	Mean ($m^3 \cdot 10^6$/year)	Stand. Dev. ($m^3 \cdot 10^6$/ year)	Max ($m^3 \cdot 10^6$/ year)	Min ($m^3 \cdot 10^6$/ year)
Pertusillo	212.15	57.72	328.54	118.25
Monte Cotugno	277.60	106.61	494.14	118.45
Cogliandrino	89.76	32.12	147.13	33.95
Marsico Nuovo	7.82	3.04	12.91	2.53
Gannano	105.54	88.56	389.03	11.72
Agri	115.54	64.43	241.55	17.92
Sauro	50.46	25.50	101.31	11.93
Sarmento	84.10	38.79	162.06	26.42

In a first phase, before applying the mixed simulation - optimization approach, a simulation-alone analysis using the WARGI-SIM module was carried out. System simulation considered the time horizon 1983-2005 and unit time period t equal to 1 month. Results provide an assessment of the system's capability to face water shortage situations when only a reactive approach is implemented in the Agri-Sinni system. Moreover this simulation analysis helps us assess the benefits of the pre-emptive measures defined by the mixed optimization-simulation approach. In WARGI-SIM a reserved volume was defined in Monte Cotugno and Pertusillo to assure a full-satisfaction of urban and industrial demands. The reserved volume is a monthly function defined as:

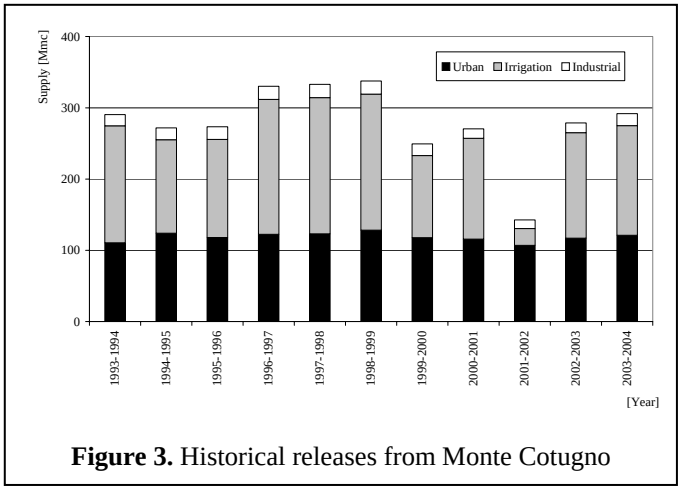


Figure 3. Historical releases from Monte Cotugno

$$V_{res,t} = \sum_{j=t}^{Oct} k\left(R_{urb} + R_{ind}\right)_j \qquad j = \{Mar, Apr, May, Jun, Jul, Aug, Sep\} \tag{3}$$

The reserved volumes considered in WARGI-SIM define a management policy in the Agri-Sinni that allocates resources for urban and industrial uses and reduced releases to irrigation use. The coefficient *k* must be carefully defined to avoid unnecessary restrictions. The allocation policy requires the frequent implementation of reactive measures.

Table 2 shows drought mitigation measures in the pro-active and reactive approach currently applied or planned by the water Authorities in the system. Additional measures have been also considered by the Authorities (e.g. construction of desalination plants or over-exploitation of aquifers) but not included in this study because of their uncertain fate.

Table 2. Drought mitigation measures in the Agri-Sinni.

Type of Measures	Pro-active approach (z_t)	Reactive approach (s_t)
Supply Increase	Reallocation available resources	
	Use of additional sources	
Demand Reduction	Pricing	Temporary restriction
	Use of agronomic techniques	Percentage rationing

Mainly, reactive measures in the Agri-Sinni system consist of reductions in irrigation water availability by mandatory percentage restrictions in supply and temporary restriction limiting the irrigation of some annual crops. Thresholds of measures implementation were statically identified as follows:

1. When deficit is between 0% and 50% a temporary restriction is introduced and irrigation for perennial crops is assured to avoid damages to the trees;
2. When deficit exceeds 50% ration allocations as a percentage of use during normal period is created.

Results obtained by WARGI-SIM alone (Table 3) highlight the lack of effective measures in a planning strategy to increase system's reliability in the case of intensive drought. A maximum annual reduction of 80.3% and 5 years where reductions exceed 50% determine unsustainable stress conditions in the irrigation sector.

Table 3. Irrigation performance index values in the simulation and simulation-optimization approach.

Index	SIM	OPT-SIM (A)	OPT-SIM (B)
Max annual reduction (% Demand)	80.3	67.8	52.4
Time reliability (% Years) Reduction =0%	43.5	34.8	65.2
Time reliability (% Years) Reduction ≤15%	65.2	56.5	95.6
Time reliability (% Years) Reduction ≤25%	69.6	65.2	95.6
Time reliability (% Years) Reduction ≤50%	78.3	87.0	95.6
Volumetric reliability (% Demand)	79.8	84.0	85.6

If the farmer has enough advance warning of drought before the irrigation season, he will have a number of drought management alternatives. Among these alternatives, he can choose between fully irrigating a lesser acreage, introducing water conservation practices and planting a more drought tolerant crop.

In the proposed approach, to mitigate the impact of these intensive drought in the predefined infrastructural configuration, water demand measures in a pro-active approach have been identified by WARGI-OPT based on synthetic series of monthly inflows (Δ=1 year). There is a trade-off between criticality of synthetic series in WARGI-OPT and the effectiveness of drought mitigation measures in WARGI-SIM. Highly pessimistic series suggest unnecessary pre-emptive measures whereas over-optimistic series provide no

useful information for the adoption of pre-emptive measures. According to the sensitivity analysis results, the 3rd worst annual series of monthly inflows has been used by WARGI-OPT to forecast system evolution 1 year ahead.

The pro-active approach applied to the Agri-Sinni includes measures that ration water demand for irrigation by pricing that influence the irrigation area and the use of less consumptive crops, and by using agronomic techniques that reduce consumption (Set A). Thresholds of pre-emptive measures implementation are dynamically defined by WARGI-OPT based on the current storage volume at the synchronization period τ_i = *April* (the beginning of the irrigation period in the Mediterranean regions) in each system's reservoir and the 3rd worst annual series of monthly inflows.

As shown in Figure 4, the Set A of pre-emptive measures is implemented in advance of the start of drought based on information provided by WARGI-OPT (e.g. demand reduction in a pro-active approach starts 2 years before the adoption of a reactive approach in 1999-2000 when a drought occurs). In the mixed optimization-simulation approach, pre-emptive measures do not need the adoption in the simulation phase of any ration allocation and significantly reduces the implementation of temporary restrictions (Figure 4). At the cost of distributing reduction over a larger period, (temporal reliabilities at low thresholds in Table 3), the mixed optimization-simulation approach also reduces the total amount of reductions both in the reactive and proactive approach equalling 67.8% instead of 80.3% (WARGI-SIM alone) during one of the most serious scarcity period (2001-2002). Also a significant reduction of the reserved volume for urban and industrial demands is presented with the coefficient *k* in (3), moving from 1 to 0.7.

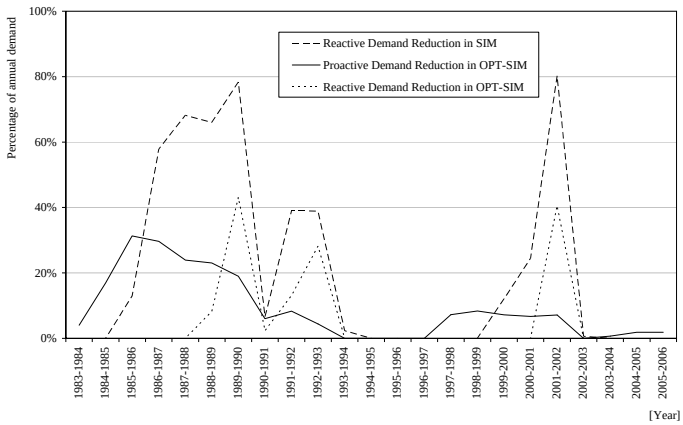


Figure 4. Temporal distribution of reduction measures in simulation (SIM) and in the mixed approach (OPT-SIM): the Set A case

A second set of drought mitigation measures (Set B) in the pro-active approach has been also implemented with the aim of reducing the risk of water scarcity by modifying supplies. These water-supply measures include:

1. Wastewater treatment plants with a total capacity of $60 \cdot 10^6$ m^3/yr serving irrigation demands;
2. Releases from the Cogliandrino reservoir serving different uses and a hydroelectric demand equalling $100 \cdot 10^6$ m^3/yr (20% in high priority).

In the pro-active approach, the volume exceeding the storage volume in the Cogliandrino reservoir to satisfy the hydroelectric demand, is transferred to the Monte Cotugno reservoir. In the event of forecasted scarcity events, WARGI-OPT increases water transfers during the spring season and saves water for future irrigation uses (Figure 5).

As wastewater use and releases from Cogliandrino are coherently implemented in a planning strategy defined by WARGI -OPT, the mixed approach needs to adopt demand reduction measures in a reactive approach (e.g. temporary restriction measure) only during 2 serious water scarcity events (Figure 6) with lower economic and environmental costs for the community than the simulation alone approach.

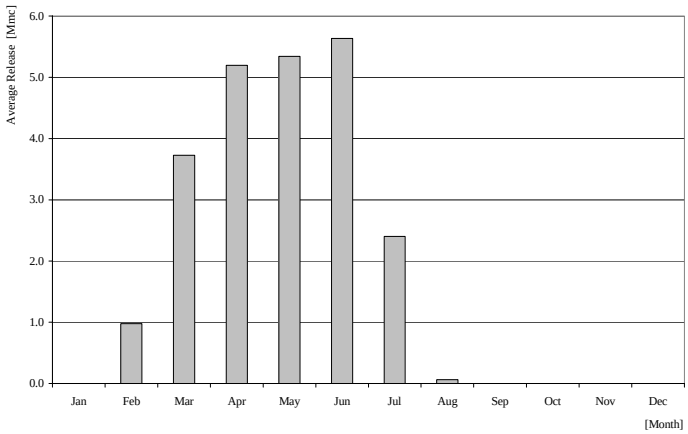


Figure 5. Mean monthly releases from the Cogliandrino reservoir as a water-supply measure in the pro-active approach

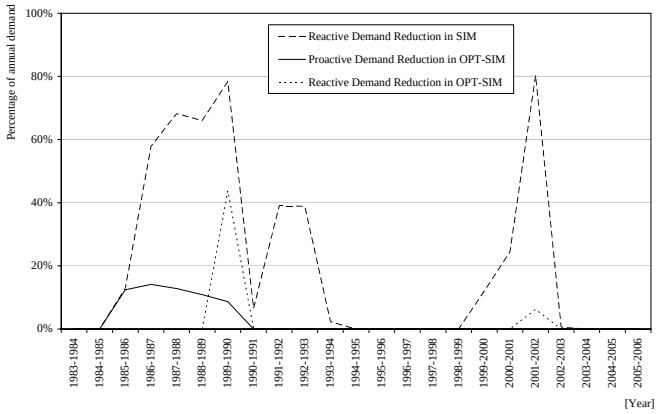


Figure 6. Temporal distribution of reduction measures in simulation (SIM) and in the mixed approach (OPT-SIM): the Set B case

According to Table 3 the comparison between maximum annual values shows a significant reduction (-35 %) of the total restrictions when water supply measures of Set B are implemented by the mixed optimization-simulation approach.

Also, it is to be noticed in the application results that demand-oriented measures in a pro-active approach define a new reduced demand in the irrigation sector as these measures give to the farmer time enough to make the water adjustments (Figure 7). The system evolution forecasting as defined by WARGI - OPT may provide the users with the flexibility needed to alter their water use from year to year. Choices of acreages and crops, and irrigation

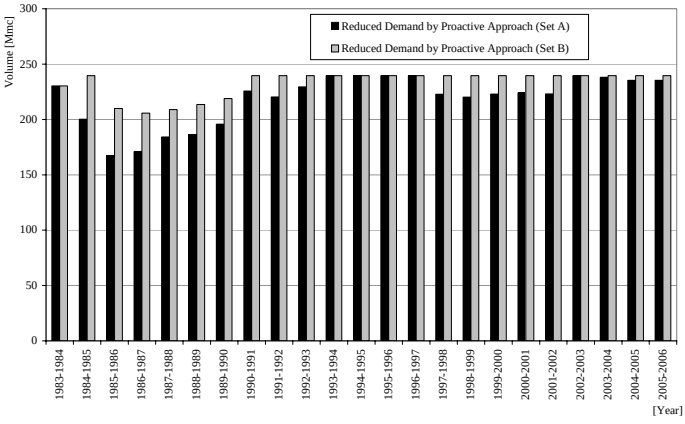


Figure 7. Reduced irrigation demand in the pro-active approach

technologies are assumed to adjust over time based on the information provided by WARGI-OPT. As demand-oriented measures are planned in advance and agreed by all users, the selected demand reductions are potentially effective at reasonable loss in total benefit value for the users. Consequently, the evaluation of system performances for irrigation use could be better defined by using time and volumetric reliability, and vulnerability related to the new dynamic demand configuration of irrigation sector. Table 4 shows high reliability values both in Set A and Set B, particularly regarding high thresholds (15% of the reduced demand or higher). Also, the reduced volumetric reliability is close to 100%.

Table 4. Irrigation performance index values based on reduced demand in the pro-active approach.

Index based on reduced demand	OPT-SIM (A)	OPT-SIM (B)
Max annual reduction (% Reduced Demand)	0.44	0.43
Reduced time reliability (% Years) Reduction =0%	0.87	0.74
Reduced time reliability (% Years) Reduction ≤15%	0.96	0.87
Reduced time reliability (% Years) Reduction ≤25%	0.96	0.87
Reduced time reliability (% Years) Reduction ≤50%	1.00	1.00
Volumetric reliability (% Reduced Demand)	0.94	0.93

4. CONCLUSIONS

Although the implementation of reduction measures is frequently the focus of arguments among people involved, it is sometimes necessary to adopt in advance these measures even when they seem unnecessary. In these discussions, the main disagreements are related to the amount of measures to be taken in advance and the timing. Bad system evolution forecasting tends to encourage early and unnecessary restrictions, whereas good forecasting could delay the implementation of insufficient measures. The high economic and environmental costs of the crisis management have provided incentive to define new approaches to improve the effectiveness of mitigation measures and to minimize the impact of droughts. The application of the proposed mixed optimization-simulation approach to the Agri-Sinni water system in South Italy shows its usefulness in defining mitigation measures in a pro-active approach. The major advantage of this approach compared with traditional simulation approaches is the ability to dynamically adopt measures based on different scenarios of system evolution given by the optimization model. Also some improvements over the approach used in this application are possible. Promising improvements include the use of artificial intelligence techniques such as artificial neural network in the evaluation of some of the key aspects of this approach as the optimization time horizon or the reserved volume to assure full satisfaction in high priority demands.

ACKNOWLEDGEMENTS

This research was developed with the financial support of the Italian Ministry of Education, University and Research, under the PRIN 2005, “Decision Support Model in complex system management under shortage conditions”, prot. 2005082390. The Authors wish to thank two anonymous reviewers for their useful comments on earlier draft.

REFERENCES

Hashimoto, T., D.P. Loucks and J.R. Stedinger, Reliability, Resiliency, and Vulnerability Criteria for Water Resource System Performance Evaluation, *Water Resources Research*, 18(1), 14–20, 1982.

Hiew, K., J. Labadie and J. Scott, Optimal operational analysis of the Colorado-Big Thompson project. In: *Computerized decision support systems for water managers*, Labadie et al. (eds), Reston, Va., 632-646, 1989.

James, L.D., and R.R. Lee, *Economics of water resources planning*, McGraw-Hill, 1971.

Lund, J.R., Derived estimation of willingness to pay to avoid probabilistic shortage, *Water Resources Research*, 31(5), 1367-1372, 1995.

PRIN, Decision Support Model in complex system management under shortage conditions, prot. 2005082390, Italian Ministry of Education, University and Research, 2005.

Sechi, G.M. and A. Sulis, Mixed simulation-optimization technique for complex water resources systems analysis under drought conditions, Chapter 11. In: *Methods and Tools for Drought Analysis and Management*, Rossi, G., T. Vega, and B. Bonaccorso (eds), Springer, 217-237, 2007.

Yevjevich, V., W.A. Hall and J.D. Salas, *Proc. Conference on Drought Research Needs*, Water Resources Publications, Fort Collins, 1978.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Living With Peak Discharge Uncertainty: The Self-learning Dike

J. L. de Kok[a] and A.Y. Hoekstra[a]

[a] *Department of Water Management (WM)*
Faculty of Engineering Technology, Twente University
PO Box 217, 7500 AE
Enschede, The Netherlands (j.l.dekok@utwente.nl)

Keywords: Flood Defence; Uncertainties; Climate Change; Adaptation; Rhine; Dike; Risk Analysis.

Abstract: Although river dikes still play a key role for flood protection in the Netherlands there is a growing interest for other measures to deal with larger peak discharges, such as lowering or widening the floodplains. Regardless of the strategy chosen the assessment of its effect on the flood risk depends on the peak discharge statistics. A problem here is that the statistical analysis of peak discharges relies on probability distributions based on the limited time series of extreme discharges. The extrapolation of these distributions are subject to considerable uncertainty, because there is a measuring record of only about 100 years and the natural variability can be expected to change as a result of climate change. This raises the question whether a more direct response to the effects of climate change is possible. The natural variability of the peak discharge changes, the changes in this variability due to e.g. climate change and the new statistical distribution can only be established after the actual change has happened. Even with regular updates of the statistical distributions it is inherent that the actions taken to reduce the flood risk are not anticipatory but delayed. As an alternative, this paper presents an adaptive or so-called self-learning approach to deal with the uncertainty in the peak discharge statistics. The difference with the probabilistic design of flood defense works, which depends on the analysis and prediction of uncertain peak discharges, is that the dike is adapted in direct response to peak water levels exceeding the dike height minus a certain safety margin. The results indicate that, on average, adaptive flood management based on observed peak water levels is at least as safe as a probabilistic approach, which necessarily relies on uncertain discharge statistics. Other advantages of the adaptive strategy are also obvious: the rule of response is simple and easy to communicate to the public, and peak water levels are less difficult to measure. In general the example demonstrates that flood management can be based on a direct response to the effects of climate change, without tedious statistical analysis of peak discharge records.

Keywords: Flood Defence; Uncertainties; Climate Change; Adaptation; Rhine; Dike; Risk Analysis

1. INTRODUCTION

After the 1953 flood catastrophe in the Southwest the Netherlands decided to follow a probabilistic approach for the design of coastal flood defence works in the Netherlands [Van Danzig, 1956; Vrijling, 2001]. From the 1980s onwards this approach was adopted for the design of the dikes in general [Vrijling, 2001]. The current approach is based on a design water level corresponding to the return period for overtopping of the dikes, which is increased with a freeboard of at least 0.5 m [MTPWWM, 2005a]. By law the design water

levels are re-evaluated every five years to deal with uncertainties with respect to the actual peak discharge variability and changes in the river characteristics. A problem encountered here is that the statistical distributions for the peak discharge are extrapolations derived from limited time series of discharge data. The knowledge about discharges and water levels with a return period of 1250 years is limited because historic data are available for a hundred-year period only. Furthermore, the volume and variability of peak discharges are subject to change due to factors such as climate change and changing upstream flow conditions. Due to climate change the peak flows in the lower Rhine corresponding to return times of 100-1000 years are expected to increase by 5 % or more over the next 50 years [Middelkoop et al., 2001]. As an extreme scenario the design discharge at Lobith could increase from 16,000 m^3s^{-1} to 20,000 m^3s^{-1} in the year 2100 [Middelkoop and Kwadijk, 2001]. In the probabilistic design a change in the discharge regime will only lead to a different design long after the change took place, even if the discharge statistics are regularly updated. It is therefore inherent to this strategy that the actions taken to reduce the flood risk are not anticipatory but following. Here we evaluate the usefulness of the probabilistic dike design policy against the background of the older design philosophy, which was to apply a safety margin on top of the highest water level ever recorded. We refer to this older design philosophy by the term 'self-learning dike', where 'self-learning' refers to the fact that dike adjustments immediately follow actual extreme flood levels. Metaphorically speaking, the dike learns through monitoring actual water levels and adapting according to a simple rule, without applying any statistics. The aim of the analysis is not to challenge the full risk-based approaches (Vrijling, 2001; FLORIS, 2006), but only to compare two different approaches to cope with discharge uncertainty. The two dike heightening strategies differ in the rule that governs the heightening of the dike. The probabilistic approach is based on the Design Flood Level (DFL), which is the water level corresponding to a flood return time of 1250 years, plus a freeboard of 0.5 m. The DFL is calculated on the basis of the historic record 1901-1998, and updated every five years by determining again the peak discharge probability distribution with the additional discharge data added to the existing time series. In the self-learning design strategy the dike height is compared with the highest water level that occurred so far. If this water level exceeds the existing dike height minus a safety margin *s*, the dike is heightened to the water level plus this safety margin. Obviously, a larger safety margin leads to a safer but more expensive dike design. For this analysis the safety margin has been taken equal to the dike height for the probabilistic dike design (including the freeboard) minus the largest historic water level. This ensures that the initial dike height is identical for both strategies.

2. METHODOLOGY

A comparison between the probabilistic flood prevention strategy and the self-learning strategy has been carried out for the case of the border gauge Lobith of the Rhine river. The endangered dike ring 48 "Rijn and Ijssel" has a total length of 57 km. In case of flooding the worst case estimate of the flood damage is 6.8 billion euro [Eijgenraam et al., 2005]. The two dike heightening strategies are compared for three different scenarios for the peak discharge statistics:

1. current peak discharge statistics without including uncertainty;
2. peak discharge statistics with uncertainty included;
3. gradual climate change trend with slowly increasing peak discharges.

In fact scenario 2 reflects the knowledge uncertainty related to limited data and scenario 3 reflects the intrinsic uncertainty due to a lack of knowledge due to climate change.

The results of the simulations are compared on the basis of the average number of dike overtopping instances, the average number of times the dike has to be heightened, and the total extra height added to the initial dike height during the 100-year simulation.

The comparison of the different dike heightening mechanisms is based on extension of the historical discharge data for the gauge station at Lobith. A time series of the yearly

maximum discharges for the period 1901-1998 [Parmet et al., 2001] was used. The complete time series has been homogenised so as to represent the river condition of 1977 [Lorenz and Kwadijk, 1999; Parmet et al., 2001]. The homogenised peak discharge data have been used to derive the parameters for the Gumbel Extreme Value distribution with the cumulative probability distribution:

$$F(Q) = \Pr ob\,(q_{max} \le Q) = \exp(-\exp(-b(Q-a)) \qquad (1)$$

where Q is the peak discharge in m^3s^{-1} and a and b are the moment estimators for the parameters of the Gumbel distribution. The stationary Gumbel distribution is justified here because statistical analysis using the Spearman ranking test did not reveal a significant trend in the historic discharge time series. In addition application of the probability plot correlation test statistic [Stedinger et al., 1993] to the discharge data proved that the historic discharge data can be described with a Gumbel distribution. For the simulated extension of the peak discharge time series in the first two scenarios there is no significant trend as well. For the scenario with a gradual climate change the probability distribution becomes non-stationary [Khaliq et al., 2006]. In this case the Gumbel parameters a and b are made time dependent by means of multiplication with a common factor, which ensures the chosen design discharge is reached after 100 years.

For the initial year of the simulation the values a = 5170 m^3s^{-1} and b = 6.584 × 10^{-4} $m^{-3}s^{-1}$ were determined from the rescaled discharge data to ensure that the 1250-year design discharge corresponded to the value of 16,000 m^3s^{-1} at Lobith, which was agreed upon in 2001 [MTPWWM, 2005b]. The return period for a peak discharge Q is then given by

$$T(Q) = \frac{1}{P(Q)} = \frac{1}{1-\exp(-\exp(-b(Q-a))} \qquad (2)$$

Once the representative Gumbel probability distribution was established the parameters were used to simulate an artificial 100-year time series following the historic period. For each year the new discharge is simulated by:

$$Q(t) = F^{-1}(x_t) \qquad (3)$$

where $Q(t)$ is the discharge in year t, $F^{-1}(x)$ the inverse of the cumulative distribution function of Eq. (1) that is determined from the extended discharge record, and x_t a random uniform number in the range [0,1]. To obtain a reliable estimate of the probability of overtopping during the 100-year period the simulations were repeated 100,000 times. This number of iterations ensures that the values of μ and σ for the simulated discharge approach the parameters of the Gumbel distribution within less than 1.5 %. For any year the water level corresponding to the discharge is determined from the known stage-discharge relationship (Figure 1) for the gauge station at Lobith, which is an extrapolation of the stage-discharge relationship based on [Schielen, 2007, Van den Brink et al., 2007]. This water level is first compared with the existing dike height to determine whether overtopping of the dike crest occurs. Depending on the dike management strategy chosen the water level is also used to adapt the dike when necessary. The results of the simulations are used to determine the average probability of overtopping during a 100-year period for each combination of the dike strategies and discharge scenarios.

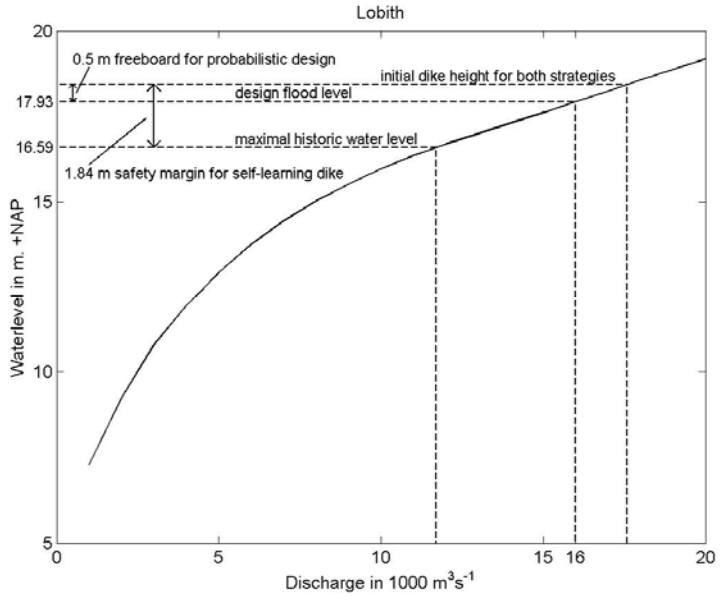


Figure 1. Stage-discharge relationship for the Rhine gauge station at Lobith based on [R.J.M. Schielen, 2007; N.G.M. Van den Brink et al., 2007].

For a fair comparison of both dike heightening strategies it is necessary to include both the costs of dike heightening and the prevented flood damage. A cost-benefit analysis has recently been carried out for dike ring 48 in the framework of the project "Room for the River" [Eijgenraam, 2005]. The investment costs of heightening the dikes along the complete 57 km length of the dike for dike ring 48 have been obtained by extrapolation of the investment costs given in Table 3.6 and Appendix C of [Eijgenraam, 2005]. For dike ring 48 the fixed investment costs are 0.5×10^6 euro, whereas the variable investment costs are 1.2×10^6 euro per km length of dike for an extra dike height of 50 cm, 1.8×10^6 euro per km length of dike for an extra dike height of 75 cm, and 3.0×10^6 euro per km length of dike for an extra dike height of 100 cm, respectively [Eijgenraam, 2005].

3. RESULTS

The performance of the two dike heightening strategies has been evaluated based on the three different discharge scenarios (Table 1). All results are based on the synthetic discharge record for the period 1998-2098. The difference between the standard and second scenario is that in the latter scenario the parameters a and b in Eq. 1 are not constant but drawn randomly in a range that has been deducted from the discharge data.

Table 1. Average performance of the probabilistic (A) and self-learning (B) dike heightening strategy for the three discharge scenarios over the 100-year simulation period.

	standard		with uncertainty		climate change	
	A	B	A	B	A	B
# dike overtoppings	0.024	0.021	0.034	0.030	0.047	0.037
# dike adaptations	2.13	1.00	2.40	1.09	3.62	1.40
extra height per adaptation	0.07	0.48	0.07	0.51	0.07	0.51
total extra height	0.14	0.48	0.17	0.55	0.24	0.71

The results indicate that, on average, the self-learning dike performs better in terms of safety and requires less adaptations of the dike height. The average extra height per adaptation and total extra height after 100 years, however, are larger for the self-learning dike. Adaptations to the dike height that are relatively small are less desirable because of

the relatively high fixed costs of dike heightening. To reduce the number of dike height adaptations one could add an extra margin of, for example, 0.5 m to the dike height when it is adapted.

Table 2 shows the discounted costs aspects for both dike strategies, based on the average of 100,000 simulations. The average flood damage is based on a 2 % yearly economic growth rate as used by Eijgenraam [2005]. The Net Present Value is based on an effective discount rate of 4 % [Eijgenraam, 2005] and determined from:

$$NPV = \sum_{t=1}^{100} \frac{C(t) + D(t)}{(1+r)^t} \qquad (4)$$

where $C(t)$ are the investment costs in year t, $D(t)$ is the flood damage in year t, and r is the discount rate. For the probabilistic design the fixed investment costs are larger than for the self-learning dike, whereas the variable costs of dike heightening are smaller. As expected the flood damage due to dike overtopping is smaller for the self-learning dike due to the higher safety level. The total expected costs (sum of investment costs of dike heightening and flood damage) are always a bit smaller for the self-learning dike. Compared with the damage incurred without dike heightening the total costs are only lower for the self-learning dike for the scenario with climate change. From a purely economic perspective this would mean that dike heightening is not necessary for most of the combinations shown in Table 2, but this changes if safety is taken into account.

Table 2. Comparison of the average costs over 100 years in million euro of dike heightening and average expected flood damage in million euro for both dike strategies for the three scenarios.

	standard		with uncertainty		climate change	
	A	B	A	B	A	B
total fixed costs	22.7	8.2	24.6	9.0	27.1	9.8
total variable costs	4.6	17.5	5.4	21.3	5.7	21.9
total flood damage	74.8	69.5	107	98.8	122	106
costs per adaptation	12.8	25.5	12.5	27.9	9.0	22.6
total expected costs	102	95.1	137	129	155	138
total expected costs without dike heightening	85		123		147	

Figure 2 below shows the year-by-year development of the average costs of dike heightening, number of dike overtoppings, flood damage, and total expected costs for both dike strategies for the climate change scenario with a gradual increase of the design discharge to 18,000 m^3s^{-1}. The gradual increase of the peak water levels (due to climate change) is directly reflected in a gradual increase of the dike height for the self-learning dike. Therefore, the average costs per year are constant for this strategy. For the probabilistic design the costs decrease first due to the increased safety level and then increase due to the effect of climate change on the discharge statistics.

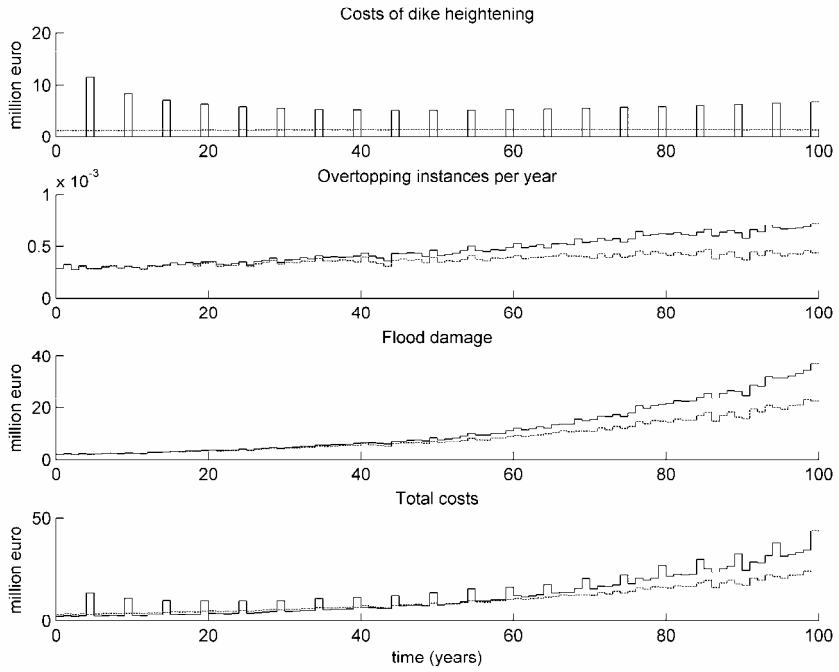


Figure 2. Average investment costs of dike heightening, number of overtopping instances, flood damage, and total expected costs for the probabilistic strategy (solid line) and the self-learning dike (thin line) over the 100-year simulation period. The discharge statistics correspond to a gradual increase of the design discharge from 16,000 to 18,000 m^3s^{-1}.

4. CONCLUSIONS

The effectiveness of the probabilistic dike design strategy has been compared with the self-learning dike strategy on the basis of the average number of overtopping instances over a 100-year period. The results of the simulations clearly indicate that the safety level and total expected costs of the self-learning dike are somewhat lower than for the probabilistic design, although the self-learning dike requires larger adaptations. Under conditions of uncertain or gradually changing discharge statistics, the average safety performance of the self-learning dike is also better than for the dike based on a probabilistic design. We acknowledge the relevance of other dike failure mechanisms and the importance of analysing how effects of flooding can be reduced in addition to examining methods to reduce flooding probabilities. Nevertheless, any flood security policy will need to include some policy towards dike heights. The approach of the self-learning dike can also be applied to other measures such as floodplain lowering or dike shifting with the same response criterion. If the water level exceeds the dike height minus the safety margin, for example, the dike is shifted backwards over such a distance that the water level for the corresponding discharge is lowered to the critical level. However, to estimate the appropriate distance it is necessary to have information on the stage-discharge relationship. Three important advantages of the self-learning dike are:

- the use of a simple rule for response which needs recording of peak water levels only and is easier to implement;
- no dependency on uncertainties in the extrapolation of discharge statistics, nor on the use of an uncertain discharge–water level relationship;
- in terms of safety communication towards the protected population the rule of the self-learning dike is more transparent.

Without challenging the full risk based or cost-benefit approaches (Vrijling 2001; Eijgenraam, 2006; FLORIS, 2006) the general conclusion is that for Lobith, on average, the self-learning dike is at least as safe as the probabilistic design. In comparison to the risk-based approach the merit lies in the more simple and direct response mechanism to uncertainties related to the discharge statistics. In a risk-based approach probability is accounted for, but the uncertainties in the statistics are usually not explicitly included.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Martijn J. Booij of the University of Twente for useful suggestions with regard to the hydrological aspects of this work and Dr. Ralph M J Schielen for providing the stage-discharge relationship used in this study.

REFERENCES

Eijgenraam, C.J.J., Veiligheid tegen overstromen - Kosten-batenanalyse voor Ruimte voor de Rivier, deel 1 (Flood safety - Cost-benefit analysis for Room for the River, part 1). CPB Document No. 82., Netherlands Bureau for Economic Policy Analysis (CPB), The Hague, The Netherlands, 2007. http://www.cpb.nl/nl

Eijgenraam, C.J.J., Optimal safety standards for dike-ring areas, CPB Discussion paper no. 62, Netherlands Bureau for Economic Policy Analysis, March 2006. http://www.cpb.nl

FLORIS, Flood risk and safety in the Netherlands (Floris), Floris study - full report, DWW report 2006-014; ISBN 90-369-5604-9, Road and Hydraulic Engineering Institute of the Netherlands Public, Works Department (Rijkswaterstaat), March 2006. http://www.helpdeskwater.nl/algemene_onderdelen/kennisdesk/?ActItmIdt=4286

Khaliq M.N., T.B.M.J. Ouarda, J.C. Ondo, P. Gachon, and B. Bobée, Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology 329*, 534-552, 2006.

Lorenz N.N., and J.C.J. Kwadijk, Onderzoek 1/1250 afvoer bij Lobith. Representativiteit meetreeks, verdelingsfuncties en statistische extrapolatie (Investigation of the 1250-year discharge at Lobith. Representativeness of the discharge record, distribution functions and statistical extrapolation). WL|Delft Hydraulics, Delft, The Netherlands. In Dutch, 1999.

Middelkoop H., K. Daamen, D. Gellens W. Grabs, J.C.J. Kwadijk, H. Lang, B.W.A.H. Parmet, B. Schädler, J. Schulla, and K. Wilke, Impact of climate change on hydrological regimes and water resources management in the Rhine basin. *Climatic Change 49*, 105-128, 2001.

Middelkoop H, and J.C.J. Kwadijk, Towards integrated assessment of the implications of global change for water management - the Rhine experience. *Physics and Chemistry of the Earth Part B 26(7-8)*: 553-560, 2001.

MTPWWM Veiligheid Nederland in Kaart - Tussenstand onderzoek overstromingsrisico's (Mapping the Safety of The Netherlands - Intermediate Status of the Examination of Flood Risks). Ministry of Transport, Public Works and Water Management, The Hague, The Netherlands. In Dutch, 2005a. http://www.projectvnk.nl

MTPWWM Planologische Kernbeslissing Ruimte voor de Rivier - Deel 3 - Hoofdstuk 4 - Strategische Beleidskeuzen (Key National Spatial Planning Decision Room for the River - Part 3 - Chapter 4 - Strategic Policy Decisions). Ministry of Transport, Public Works and Water Management, The Hague, The Netherlands, 2005b. In Dutch. http://www.ruimtevoorderivier.nl

Parmet, B.W.A.H., W. van de Langemheen, E.H. Chbab, J.C.J. Kwadijk, F.L.M. Diermanse, and D. Klopstra, Analyse van de maatgevende afvoer van de Rijn te Lobith (Analysis of the design discharge of the River Rhine at Lobith), RIZA report 2002.012. Ministry of Transport, Public Works and Water Management, Directorate Rijkswaterstaat, Institute for Inland Water

Management and Waste Water Treatment (RIZA), Arnhem, The Netherlands, 2001. http://www.rijkswaterstaat.nl/rws/riza/home/publicaties/rapporten/2002

Schielen, R.M.J., Personal communication Ministry of Transport, Public Works and Water Management, Directorate Rijkswaterstaat, Institute for Inland Water Management and Waste Water Treatment (RIZA), 2007.

Stedinger, J.R., R.M. Vogel, and E. Foufoula-Georgiou, Frequency analysis of extreme events. *Handbook of hydrology*, Chapter 18, pp 18.27-18.28, McGraw-Hill, USA, 1993.

Van Danzig, D., Economic decision problems for flood prevention. *Econometrica 24*: 276-287, 1956.

Van den Brink N.G.M., D. Beyer, M.J.M. Scholten, E.H. van Velzen, Onderbouwing Hydraulische Randvoorwaarden 2001 voor de Rijn en zijn takken (Foundation for the Hydraulic Boundary Conditions for the Rhine and its branches). RIZA report 2002.015. Ministry of Transport, Public Works and Water Management, Directorate Rijkswaterstaat, Institute for Inland Water Management and Waste Water Treatment (RIZA), Arnhem, The Netherlands. In Dutch, 2007.

Vrijling J.K., Probabilistic design of water defense systems in The Netherlands. *Reliability Engineering and Systems Safety 74*, 337-344, 2001.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

An integrated multimedia model and multicriteria analysis approach to managing sewage sludge application on agricultural soils: framework and methodology

A. C. Passuello[a], M. Schuhmacher[a,b], V. Kumar[a,c], M. Nadal [b], J. L. Domingo[b]

[a] *School of Chemical and Process Engineering, ''Rovira i Virgili'' University, Av. Paıses Catalanes 26, 43007 Tarragona, Spain*
[b] *Laboratory of Toxicology and Environmental Health, School of Medicine, ''Rovira i Virgili'' University, San Lorenzo 21, 43201 Reus, Spain*
[c] *Catchment Science Centre, University of Sheffield, S3 7HQ, Sheffield, UK*
(ana.passuello@urv.cat)

Abstract: Sludge application to agricultural soils is a managing practice of increasing use, because of its benefits to soil and crops. The Council Directive 86/278/CEE regulates concentrations of heavy metals for sludge application, but has no restrictions related to organic compounds contamination. The Decision Support System (DSS) presented here, integrates a multimedia risk model of persistent organic pollutants (POPs) to human and ecosystem (P), a model concern to areas' vulnerability to nitrogen (N), and also an economical model (E), to indicate the best decision management for sludge application on agricultural soils. The three dimensional representation of the model allows a simple visualization of the multicriteria analysis, permitting transparency on decision making.

Keywords: Decision Support System (DSS), Sludge Final Disposal, Diffusive Transport Model, Risk Assessment.

1. INTRODUCTION

The application of sewage sludge to agricultural soils is a widely extended practice, and that is progressively increasing with the years. The number and capacity of wastewater treatment plants has increased in the last few years in Spain, resulting in an enlargement of 47% on sludge production between the years 1997 and 2003 [Applus, 2007]. The sludge from wastewater treatment plants is a residue that must be properly disposed, in order to protect the environment. Its increasing use on agriculture soils without proper treatment makes it a topic of concern. However with careful management, the application of sludge to land is clearly beneficial to agricultural soils and crops. The main benefit for agricultural proposes is the increase of the nutrients as well as the organic carbon content in soil, which improves the crops production.

The Council Directive 86/278/CEE determines values for concentrations of heavy metals in soils to which sludge is applied, concentrations of heavy metals on sludge and the maximum annual quantities of such heavy metals which may be introduced into soil intended for agriculture. However, in Spain, there are no legally enforceable restrictions related to organic compounds which may be present in sludge.

An extensive and diverse range of organic compounds are known to exist in sewage sludge [Katsoyiannis and Samara, 2005], many of which become transferred to sludge-amended agricultural soils [Wild and Jones, 1992]. These polluting agents move through the different compartments, and they accumulate throughout the food chain. As is, a multicompartimental model is used as a tool for predicting pollutants distribution on the

environmental compartments (air, water, soil, crops), and evaluate its risk for the population as well as the ecosystem, through exposition and risk assessment models.

Another point of concern is that sludge application on soils may lead to groundwater nitrification, as a result of nitrogen movement through the lixiviate. The nutrients content of sludge is a variable related to application dose. Studies have been proved that sludge application improves soil fertility along the years [Fuentes et al., 2008]. On the other hand, care must be taken on areas that are vulnerable to nitrogen pollution, as recommended on EC Nitrate Directive 91/676.

Economically, sludge application improves soil fertility and exempt fertilizers use to correct the soil. Also, the crop production is increased because of the nutrients content of the sludges.

In order to determine the best scenarios for sludge application on agricultural soils for crops production in Spain, a decision support system (DSS) that integrates economical and environmental issues is presented. The objective of this DSS is to minimize the environmental risks (to human and ecosystems) and maximise the economic return, considering the characteristics of the different sites that sludge may be applied. The DSS is presented in Figure 1.

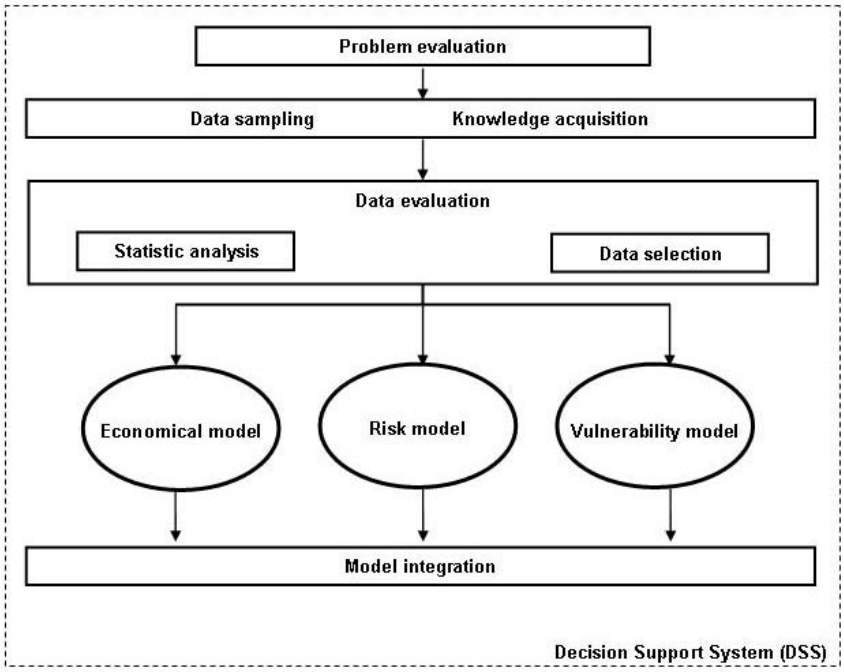


Figure 1 Phases of the decision support system (DSS)

The DSS presented here allows analysing different scenarios, according to economical and environmental criteria. For that, three models are used: one related to risk impacts of persistent organic pollutants (POPs) to human and ecosystem (P), another related to vulnerability to nitrogen (N), and also an economical model (E). The models will be applied to different scenarios, according to local site and sludge characteristics. The development of these models is presented in the following section.

2. MULTICRITERIA ANALYSIS

2.1 Risk model

Persistent organic pollutants (POPs) constitute a class of man-made chemicals with a pronounced persistence against chemical/biological degradation, environmental mobility, a

tendency for bioaccumulation in human and animal tissue, and significant impacts on human health and the environment, even at extremely low concentrations [Katsoyiannis and Samara, 2005]. In a previous study, several samples of sludge were collected and characterized. The sludge analysis defined the POPs to be evaluated (PAHs, PCBs, PBDEs, PCNs, Nonilfenol, PFOS and PFOAS).

The fate model considers the rates of POPs' release because of sludge application and its fate through soil, air, crops, soil invertebrates and soil-dwelling mammals. Vegetation and soil have been treated separately, as their characteristics of exchange with the atmosphere are different [Wania et al., 2006].

The model is based on Level IV fugacity models [Webster and Mackay, 2007], which is most useful for estimating recovery times of a contaminated system that is experiencing reduced emissions [Mackay, 2001]. That's the case of sludge amended soils, which receive an amount of sludge once a year.

Physical chemical properties data and concentrations in sludge are used to calculate the concentrations in the compartments, and uptake into vegetation, soil invertebrates and soil-dwelling mammals. The soil component of the model allows different layer of soil to be defined. The model is dynamic and allows seasonally variable environmental parameters, since the seasonal and multi-year time scale is of most interest for POPs [Wania et al., 2006].

The parameters used to develop the model are: sludge characteristics and application dose, soils characteristics (porosity, hydraulic conductivity, bulk density, field area, density of organic matter and mineral matter, bioavailability factor, layer depth, volume fraction of air and water), kind of crop, temperature, and pollutant properties (solubility, K_{OW}, vapour pressure, H, persistency, half life in soil and toxicity).

The parameters considered to estimate crop contamination are biosolids properties (K_{OW}, K_{OA}), contaminant concentration, and crop properties, for two types of cereals: corn and wheat. Also, the climatology is considered, as the crops are cultivated at separate seasons. For invertebrates and mammals, the volume, density, volume fractions of lipid and water are taken into account.

The key processes considered are those who could lead to loss of chemical during the transport, at the two layers of soils (Figure **2**).

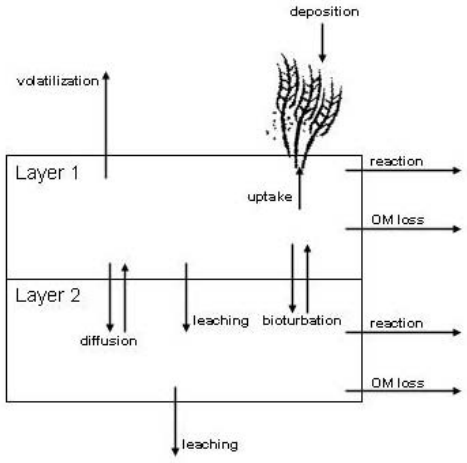


Figure 2 Transfer processes considered on the different environmental compartments. Adapted from Webster and Mackay [2007].

The environmental compartments are linked by inter-compartmental transfer processes. The first layer soil loses chemical by five processes: volatilization to the air above the soil compartment, leaching to the second layer, sorbed phase transport (bioturbation) to the second layer, diffusion in pore air and pore water to the second layer, and degrading reactions. The second layer loses chemical by sorbed phase transport and diffusion to the

first layer, by leaching, and by degrading reactions. Chemical is lost from the entire soil system by degradation in each layer, by volatilization from the top layer, and by leaching from the lower layer [Webster and Mackay, 2007].

A fugacity approach is used in order to estimate the mechanisms of volatilization, diffusion, leaching, reaction, bioturbation and loss of organic matter. A sensibility analysis is assessed with the aim of defining the most relevant parameters.

For model validation, the analysis of collected samples of sludge and soils where the sludge is applied will be used. Also, an experiment on field scale is being developed, in order to quantify toxicological impacts and pollutants movement. In this experiment, the different types of soils and climate will be assessed, considering two cereal types: corn and wheat.

Organic contaminants may impact humans and ecosystem through different pathways [Schowanek, et al., 2004]. The exposure and risk model considers ecosystem and human exposure to polluted soil, food and air, with the different routes of contamination: air pollution due to volatilization, vegetal pollution through soil and polluted air, dermal absorption by contact with soil, ingestion of contaminated soil and crops, inhalation of re-suspended soil particles.

The Risk Assessment Analysis is based on CEC [2003]. Risks for adverse human health effects are estimated assuming to be carcinogens or non-carcinogens [Schuhmacher et al., 2001]:

• Non-carcinogenic risk: To determine if the contaminant poses a risk to human health, daily intake is compared with the reference dose (RfD) for chronic exposure. The tolerable daily intake (TDI) established by the WHO was here used.

• Carcinogenic risk: It is calculated by multiplying the estimated dose by the carcinogenic potency factor. The predicted carcinogenic risk is an upper-bound estimate of the potential risk associated with the exposure.

The risks to the ecosystem are calculated using the PECs for all environmental compartments, as described in the EU Technical Guidance Document [CEC, 2003].

Chemical properties may be highly uncertain due to difficulty in measurement, while landscape data are highly variable at spatial scales [Bennett et al., 1999]. Therefore, it is very important to present clearly the propagation of parameter variances into model outputs that reflect environmental persistence of chemicals [Luo and Yang, 2007]. For that, the Fuzzy Latin Hypercube Sampling (FLHS), a technique that allows characterization of both uncertainty and variability in the input variables, will be applied as described by Kumar et al. [2008].

2.2 Vulnerability model

To determine the areas that are vulnerable to nitrogen, the lixiviate movement through groundwater is assessed; considering the geographic boundaries, the biosolid's contaminant concentration, the kind of agricultural application, soil and pollutant properties.

The transport of nutrients to groundwater will be based on an improved version of DRASTIC [Rupert, 1999], which calibrates the point rating scheme to measured nitrite plus nitrate as nitrogen ($NO_2 + NO_3$ - N).

DRASTIC index is a numerical ranking system developed by Aller et al. [1987] to assess groundwater pollution potential in various hydrogeologic settings. In the DRASTIC methodology, groundwater pollution potential is evaluated by seven factors: D – depth to water; R – net recharge; A – aquifer media; S – soil media; T – topography (slope); I – impact of the vadose zone media; and C – hydraulic conductivity of the aquifer. Each of the DRASTIC factors is assigned a relative weight ranging from 1 to 5. The most significant factor has a weight of 5 and the least significant has a weight of 1.

Once the DRASTIC Index has been computed, it is possible to identify areas, which are more likely to be susceptible to GW contamination relative to others. The higher the DRASTIC Index, the greater is the GW pollution potential.

2.3 Economical model

For the economic evaluation, the costs and benefits of the sludge application are calculated. The costs considered on this evaluation are transport, storage and sludge application on land. Thus, the main economical variables are the distance between the agricultural area and the treatment plant, the amount of sludge produced and the nitrogen content of sludge, which is related to the application dose.

The organic matter content of soil is improved with the sludge application, fact that is not observed on fertilisers application on soil [Fuentes et al., 2008]. For this reason, the agricultural productivity of the alternative sites (scenarios) where sludge is applied is estimated as described by Koo and O'Connell [2006]. The agronomic model is based on nitrogen response curves and the crops' price. The yield (in t/ha/year) is a function of nitrogen application (kg N/ha/year), and the price is estimated by the crops market value (€/t).

3. AGGREGATION

In order to define the normalization indexes, different scenarios will be applied for each of the models. The normalization values are defined through a linear function [Afgan and Carvalho, 2004].

Then, the models are integrated considering the normalized values aggregated as a three dimensional vector **(**Figure 3**).** For that, weighting values may be defined by the decision maker. The best option is, thus, the one who gets closer to an optimum value.

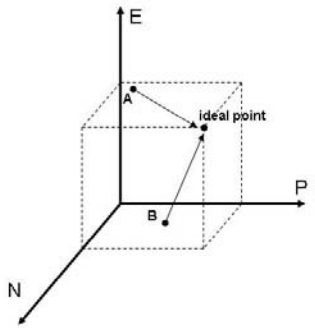


Figure 3 Representation of two scenarios as three dimensional vectors: risk impacts of persistent organic pollutants (POPs) to human and ecosystem (P), vulnerability to nitrogen (N), and economical model (E).

It may be noted in Figure 3, that the point A represents an alternative viable economically, but with low concern on environmental problems. On the other hand, point B is not interesting on economical point of view, but has a major concern on the environmental issue than point A. In order to calculate which point is closer to an "ideal scenario", the following equation is suggested:

$$d_{ij} = \sqrt{\lambda_P^2\left(u_j - u_i\right)^2 + \lambda_N^2\left(v_j - v_i\right)^2 + \lambda_E^2\left(w_j - w_i\right)^2} \quad (1)$$

Where d_{ij} is the distance between the selected scenario *i* and the optimal point *j*, λ_P, λ_N, and λ_E, are the weighting values for the scenarios P, N and E; u_i and u_j are the vector P values

for the scenario *i* and the optimal point *j*, v_i and v_j are the vector N values for the scenario *i* and the optimal point *j*, and w_i and w_j are the vector E values for the scenario *i* and the optimal point *j*.

4. CONCLUSIONS

At present, the Spain's legal regulation regarding sludge application on agricultural soils only concerns about metal levels. Several works have shown that the existence of vulnerable areas, as well as the impacts of the POPs present in sludge are issues of concern.

The DSS presented here consists in a simple way of aggregating different criteria. It permits using models based on different evaluation techniques, as well as on decision maker's experience. It represents a first screening of the main parameters to be considered on a DSS of sludge application on agricultural soils, and will be applied to a reduced number of scenarios. Then, a site specific evaluation must be done, with the aim of assessing the uncertainty of the selected variables.

The economic evaluation is significant, as it consists on main criteria for the decision maker. In this study, not only direct cost, but long range estimation was considered, in order to define the best scenario along a larger period. Besides, because of the tendency of water resources shortage, avoiding the groundwater nitrification leads to a prevention of future environmental, and consequently, economical problems.

Moreover, this article presents an innovative way to assess the risk of POP's exposure to humans and ecosystem, evaluating the fate of these contaminant, followed by a exposure and risk analysis.

The three dimensional representation allows viewing the contribution of each of the selected criteria (economical, vulnerability and risk) on the final result. The weighting factor applied by the decision maker may be clearly observed, as well as the individual contribution of the models, enhancing the transparency and reliability of the method.

ACKNOWLEDGEMENTS

The authors wish to thank Ministry of Science and Technology and Ministry of Industry, Tourism and Trade.

REFERENCES

Afgan, N. H. and M. G. Carvalho. Sustainability assessment of hydrogen energy systems. *International Journal of Hydrogen Energy* 29(13): 1327-1342, 2004

Aller, L., T. Bennett, J.H. Lehr, R.J. Petty and G. Hackett. *DRASTIC*: a standardized system for evaluating ground water pollution potential using hydrogeologic settings. United States Environmental Protection Agency, 622 pp., 1987.

Applus. *Proyecto SOSTAQUA*. Contaminantes orgánicos persistentes. Evaluación del riesgo medioambiental y sanitario asociado a la gestión de lodos de depuradora. Memoria parcial julio 2007. Applus Labaqua, 24pp., 2007.

Bennett, D. H., W. E. Kastenberg, and T.E. McKone. A multimedia, multiple pathway risk assessment of atrazine: the impact of age differentiated exposure including joint uncertainty and variability. Reliability *Engineering & System Safety* 63(2): 185-198, 1999.

CEC. *Technical Guidance Document on Risk Assessment of Chemical Substances following European Regulation and Directives*. April 2003, second ed. European Chemicals Bureau (ECB), JRCISPRA, Italy, 2003.

EEC. Council Directive 86/278/EEC of 12 June 1986 on the protection of the environment, and in particular of the soil, when sewage sludge is used in agriculture. *Official Journal* L 181. P. 0006 – 0012, 1986.

EEC. Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources.

Fuentes, J.S., F.C: Sagué, J.M.P. Grau, L.S: Genoher, C.N. Romeral. *Resultats de l'aplicació reiterada de fangs d'EDAR en els cultius del cereal*. IRTA Mas Badia, 13pp., 2008.

Katsoyiannis, A. and C. Samara. Persistent organic pollutants (POPs) in the conventional activated sludge treatment process: fate and mass balance. Environmental Research 97(3), 245-257, 2005

Koo, B. K. and P. E. O'Connell. An integrated modelling and multicriteria analysis approach to managing nitrate diffuse pollution: 2. A case study for a chalk catchment in England. *Science of The Total Environment* 359(1-3), 1-16, 2006.

Kumar, V., M. Mari, M. Schuhmacher, and J.L. Domingo. Partitioning total variance in risk assessment. Application to a case study: a municipal solid waste incinerator. *Environmental Software and Modelling*. In Press, 2008.

Luo, Y. and X. Yang. A multimedia environmental model of chemical distribution: Fate, transport, and uncertainty analysis. *Chemosphere* 66(8): 1396-1407, 2007.

Mackay, D. *Multimedia Environmental Models*: The Fugacity Approach, Second Edition. Trent University, Ontario, Canada, 261pp., 2001.

MAFF, (Ministry of Agriculture, Fisheries and Food), Agricultural land classification of England and Wales, MAFF, 60 pp., 1988.

Saaty TL. Decision-making with the AHP: why is the principal eigenvector necessary? *European Journal of Operational Research*, 145, 85–91, 2003.

Schuhmacher, M., M. Meneses, A. Xifro, and J.L. Domingo. The use of Monte-Carlo simulation techniques for risk assessment: study of a municipal waste incinerator. Chemosphere 43(4-7): 787-799, 2001.

Schowanek, D., R. Carr, H. David, P. Douben, J. Hall, H. Kirchmann, L. Patria, P. Sequi, S. Smith, S. Webb. A risk-based methodology for deriving quality standards for organic contaminants in sewage sludge for use in agriculture -Conceptual Framework. Regulatory Toxicology and Pharmacology 40(3), 227-251, 2004.

Vaidya, O. S. and S. Kumar. Analytic hierarchy process: An overview of applications. *European Journal of Operational Research* 169(1), 1-29, 2006.

Wania, F., K. Breivik, N.J. Persson, and M.S. McLachlan. CoZMo-POP 2 - A fugacity-based dynamic multi-compartmental mass balance model of the fate of persistent organic pollutants. *Environmental Modelling & Software* 21(6), 868-884, 2006.

Webster and Mackay. *Modelling the Fate of Contaminants Introduced into Agricultural Soils from Biosolids*. Canadian Environmental Modelling Centre. Trent University. Peterborough, Ontario, 54pp., 2007.

Wild, S. R. and K. C. Jones. Organic chemicals entering agricultural soils in sewage sludges: screening for their potential to transfer to crop plants and livestock. *The Science of The Total Environment* 119, 85-119, 1992.

Assessment and Management of Environmental Risks Resulting from Operation of Large-scale Irrigation Delivery Systems: a Case Study in Southern Italy

D. Zaccaria[a] and M. Scimone[b]

[a] *Land and Water Division, International Center for Advanced Mediterranean Agronomic Studies, Mediterranean Agronomic Institute of Bari, Italy (zaccaria@iamb.it)*

[b] *Department of Biology, University of Trieste, Italy (mauro.scimone@iol.it)*

Abstract: The cumulative effects of irrigation delivery schedule on aquifer and on agricultural soils were analysed in a study conducted on a large-scale irrigated area located in southern Italy. These effects likely result from the inadequate operation of the large-scale irrigation delivery system. Agriculture in the study area is market-oriented and tightly dependant on irrigation, due to hot and dry climatic conditions and to the poor distribution of seasonal rainfall. The study area is characterized by fairly high levels of groundwater salinity likely due to intensive exploitation of the aquifer by farmers especially during peak irrigation demand periods. The concentration of groundwater pumping during these periods results most likely in seawater intrusion. Rising of salinity levels are in fact noted in the groundwater from the start of irrigation season up to the months of July and August. The effects of alternative irrigation distribution schedules to groundwater salinity were assessed in a previous study by simulating water deliveries through a soil-water budget approach under different combinations among crops, soils and climatic conditions. The water distribution currently carried out by the local managing authority poses risks both to the aquifer and to agricultural soils. The present work aimed at conducting a generic Risk Assessment and Risk Management of the soils and aquifer degradation in the study area. Risk generating processes for the major environmental hazards were clearly identified and risks significance was also estimated. Some feasible management options were tentatively evaluated by using different criteria such as the social acceptability, the effectiveness in risk alleviation and the necessary costs for implementation. The Risk Assessment and Risk Management procedures were carried out in view of eventual new agricultural/territorial planning as well as of potential counter-measures for mitigating the existing environmental problems and for reducing the potential risks.

Keywords: Irrigation delivery systems, Environmental Risk Assessment, Seawater intrusion.

1. INTRODUCTION

In many Mediterranean irrigation schemes distribution of water to farms is conducted by means of large-scale delivery networks. In the past, these systems played a significant role in developing, stabilizing and diversifying agricultural production. At present, the major focus is on the effective distribution of limited water resources that otherwise would be accessible only to few users (Lamaddalena and Sagardoy, 2000). Under scarcity conditions, water allocation and distribution to farms usually follows a supply-driven approach

Despite their promise as engines of agricultural growth, in many areas of the world irrigation delivery systems perform far below their potential (see Small and Svendsen, 1992). As pointed out by Plusquellec et al. (1994) and by Lamaddalena et al. (1995), this

may be due to unrealistic designs, to inadequate water delivery schedules and to operational constraints. When crop production is strongly dependant on irrigation, the effective and responsible water delivery to farms is critical for maintaining economic and environmental sustainability of agricultural areas. If water delivery by the management authority is unreliable, inadequate in terms of the conditions of delivery, rigid or not timely matching crop water demand or growers' practices, farmers tend to rely on groundwater resources as main source of irrigation.

Many researches have proved that the over-exploitation of groundwater for agricultural purposes causes a drawdown of water levels and gives way to the seawater intrusion in coastal areas, consequently resulting in the deterioration of quality of the irrigation water pumped from wells (Paniconi et al., 2001). Seawater intrusion into the groundwater (Capaccionia et al., 2005) occurs during periods of intensive pumping, also leading to the subsequent salt build-up in the cropped soils.

Under the perspective of responsible use of natural resources, there is a specific need for providing land-use managers with reliable tools and techniques for assessing environmental man-induced risks. This can generate the opportunity of evaluating the reliability of mitigation and management options on a sound scientific basis.

This paper describes the work for conducting the generic Risk Assessment and Risk Management (RA&M) of the soils and aquifer degradation in a large-scale irrigation scheme located in southern Italy and managed by a local Water User Organization (WUO). Some management options, which represent alternatives to the "business-as-usual", were identified and evaluated for mitigating existing environmental problems in the area.

2. THE STUDY AREA

The "Sinistra Bradano" large-scale irrigation scheme is an agricultural area located in the western part of the province of Taranto, (Apulia Region, southern Italy) covering a total topographic area of 9651 ha. The whole scheme is subdivided into 10 operational districts that include an overall irrigated area of 8636 ha. The water source is the San Giuliano dam located on the Bradano River in the nearby region of Basilicata. The dam has a total capacity of 70 Mm^3, out of which 35 Mm^3 are usually available for irrigation of the Sinistra Bradano scheme. Water is conveyed from the San Giuliano dam to the study area by means of a main canal for about 30 Km. The water conveyance and distribution systems are managed and operated by a local WUO. The area is characterized by a hot and dry climate, which is typical for the Mediterranean region. The precipitation is scarce, averaging around 550 mm $year^{-1}$, and occurs mainly in the period between November and March. Summer droughts are frequent. Irrigation water is supplied to farms through a pressurized distribution system. The distribution network is operated by rotation delivery schedule with a fix delivery interval of 10 days. The rotation is fixed for the entire irrigation season (April through October) with a flow rate of 20 l s^{-1} and 5 hours of delivery of to each user.

The main crops are citrus, table grapes, olive trees and summer vegetables. Soils are mainly of alluvial type and the applied water gets drained in 2 or 3 days. Because of semi-arid climatic conditions, profitable farming in the area strongly depends on irrigation. Nevertheless, farmers report that usually the water distribution conducted by the WUO is not matching the actual crops' needs nor the farmers' requirements, both in terms of timing and of conditions of delivery. Delivery intervals, flow rates and pressure heads available at hydrants are found to be inadequate by farmers for the prevailing farming practices. Consequently, during the last 10 years many farmers have utilized groundwater pumping as main source of irrigation for large part of the irrigation season. A concentration of groundwater pumping occurs during the peak water demand periods (July and August). This progressively led to high pressure on groundwater resources and started originating aquifer contamination and soil degradation. In details, seawater intrusion in the groundwater and salt build-up in the agricultural soils are the major causes of environmental degradation in the study area.

Seasonal changes in the groundwater quality were assessed in a previous study by comparing the groundwater salinity in winter versus that in summer. Summer and winter salinity maps were in fact developed based on the spatial interpolation of point-measured values of the Total Dissolved Solids (TDS) and Electrical Conductivity (EC) of samples

collected from the groundwater in the Sinistra Bradano area during the irrigation season of 2006. The comparison between winter and summer maps showed a relevant increase in salinity from winter to summer, from EC of 1.5 - 1.8 g l^{-1} in winter to 3.1 g l^{-1} measured in summer. This salinity increase is most likely related to the intensive pumping from the aquifer during period of peak demand, specifically from May to August. This can be reported as the main resulting consequence of the inadequate delivery schedule enforced in the area. Several interviews with local operators and growers revealed that many water users rely exclusively on groundwater pumping for irrigating their crops in order to achieve the desired flexibility. The rigid rotation supply may in fact cause wasteful water use such as improper timing, over- irrigation and runoff, and may inhibit good farm management (e.g. Merriam et al. 2007).

The local climatic conditions as well as the intensive farming of agricultural areas and non-optimal distribution of water supplies make "business-as-usual" not sustainable in the area on the long run.

Given the above hazardous phenomena and the need for a strategic change with respect to the existing situation in the study area, a generic Risk Assessment and Management (RA&M) procedure was conducted through a new framework in view of possible counter-measures for mitigation of existing environmental problems and risks. The RA&M framework was conceived under an EU-funded project named "*Remotely Accessed Decision Support System for Transnational Environmental Risk Management* (Project STRiM), within the Interreg III B - CADSES Program, and is outlined hereafter.

3. THE STRiM RA AND RM FRAMEWORK

The STRiM RA&M framework (Figure 1) consists of 5 iterative steps and is predominantly based on the Environmental Risk Management guidelines issued by the Department of Environment Food and Rural Affairs (DEFRA, 2002) of United Kingdom, which focus on risk management and applicability to any type of environmental risk. The STRiM RA&M framework is linked to other key environmental protection decision-making procedures such as the Environmental Impact Assessment (EIA), the Strategic Environmental Assessment (SEA) and the framework conceived by the European Environmental Agency (EEA) on Driving Forces, Pressures, State, Impacts and Responses (DPSIR). Environmental Risk Assessment (RA) and Risk Management (RM) require datasets to support decision-making, often in the form of indicators. In order to harmonise environmental protection management, the STRiM framework links the DPSIR indicator and monitoring framework with RA and RM, something not attempted before. The framework embeds risk assessment into the risk management process and, as such, includes a number of key aspects emerging throughout the various steps of the process. Among these issues, the most relevant are: **a)** the importance of accurately defining the actual problem; **b)** the need to prioritize all relevant risks prior to quantify them and proceed with the data collection; **c)** the need to consider the risks taking into account feasible management solutions by using options appraisal from the initial stages; **d)** the iterative nature of the process.

4. POTENTIAL HAZARD IDENTIFICATION

The potential hazard occurring in the study area is the aquifer over-exploitation. Under the specific situation of the study area, the intention (defined as "*any course of action, intentional or otherwise, which by its nature may pose a risk to the environment - natural or built - and the life it sustains*") for which the RA&M is required is the management of business-as-usual, which corresponds to the intensive farming along with the adopted irrigation delivery schedule by the water management authority. The secondary hazards resulting from the aquifer over-exploitation are those reported in the Figure 2, whereas the sources, pathways, receptors and impacts are indicated in the Table 1 for each hazard.

The potential sources for aquifer over-exploitation are the intensive pumping (S1) by farmers during peak irrigation demand periods (July and August) and the inadequate water distribution through the irrigation networks (S2). This situation occurs due to the existing

market-oriented agriculture based on water-demanding crops and to the current operation of the irrigation distribution system, which is not matching crops and farmers' water requirements. The primary pathway (P1) is through groundwater pumping, which in some periods may occur beyond the safe yield of aquifer due to concentration of withdrawals. This has the effect of depressurizing the aquifer, giving way to seawater intrusion and to aquifer contamination by saline water. The receptor of salinization by seawater is thus the aquifer itself.

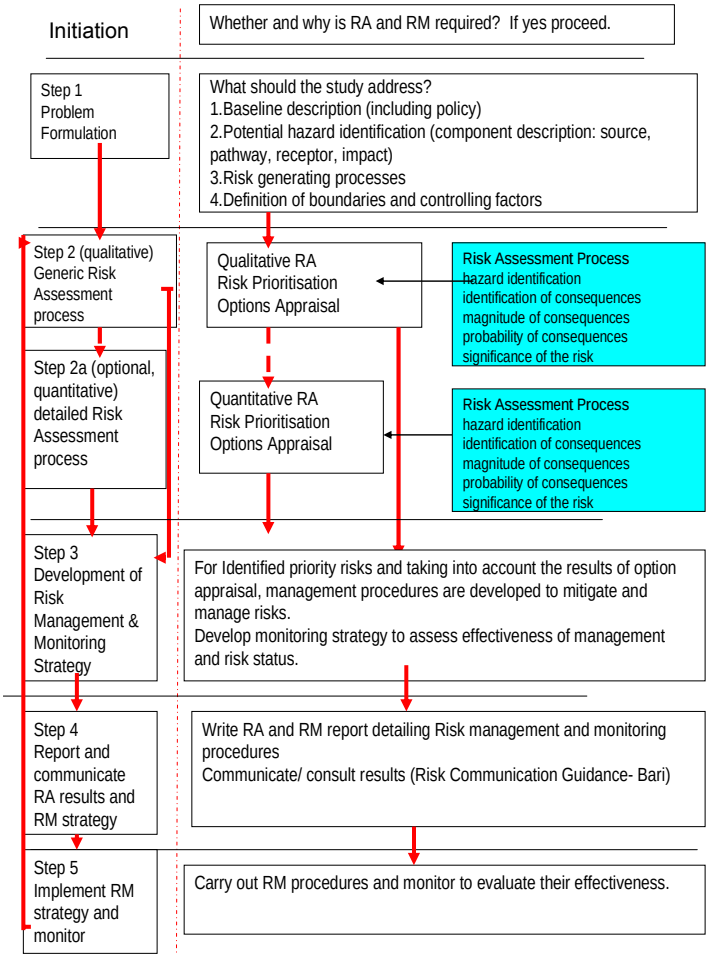


Figure 1. STRiM Risk Assessment and Management Framework (modified from DEFRA, 2002)

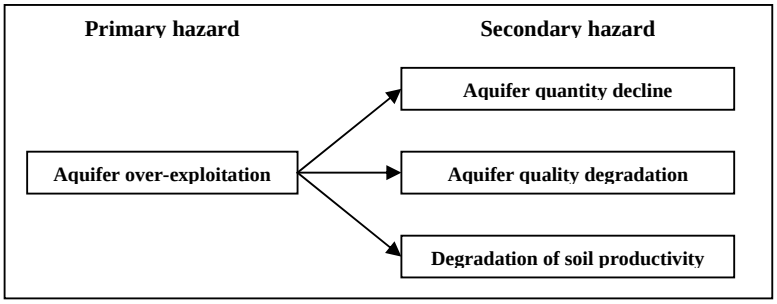


Figure 2. Primary and secondary hazards identified in the study area

Table 1. Hazards sources pathways receptors and impacts

Hazard	Source	Pathway	Receptor	Impact
H1 Aquifer over-exploitation	**S1** - Intensive pumping by farmers during peak demand periods	**P1**-Aquifer	**R1**-Aquifer	**I1.1** - Aquifer Depletion
				I1.2 - Salinization by seawater intrusion
	S2 - Poor water distribution through the irrigation networks	**P2**-Aquifer	**R2**– Soils	**I 2.1** - Salt build-up in the soils

The standardized map of salinity distribution within the aquifer is reported in Figure 3. Figure 4 represents the spatial distribution of environmental vulnerability as related to aquifer pumping, to salinity distribution within the aquifer and to aquifer recharge.

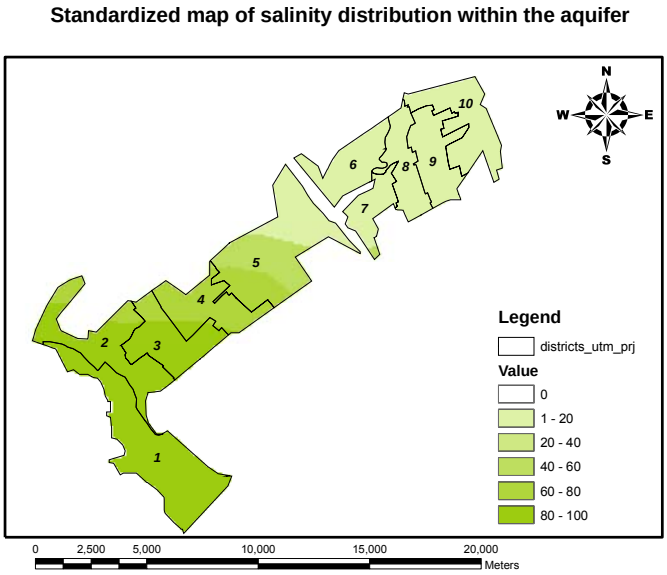


Figure 3. Standardized map of salinity distribution within the aquifer

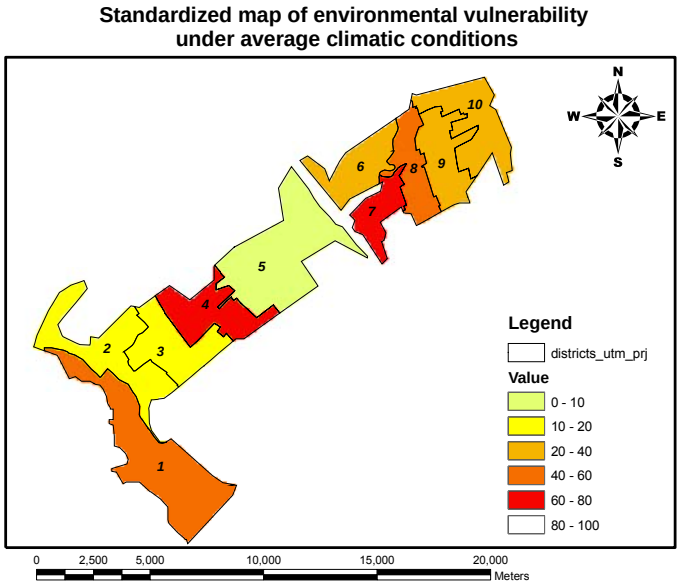


Figure 4. Standardized map of environmental vulnerability under average climatic conditions

The secondary pathway (P2) is again through groundwater pumping by farmers and through the distribution of saline water onto irrigated fields. The main potential impacts are: the quantitative depletion of aquifer (I1.1), which occurs beyond the safe yield and/or beyond the natural recharge capability; the aquifer salinization due to seawater intrusion (I1.2); the accumulation of salts in the irrigated soil (I2.1) resulting from the distribution of saline irrigation water and from the water evaporation and transpiration processes.

5. IDENTIFICATION OF RISK GENERATING PROCESSES

Risk generating processes for the over-exploitation of aquifer.

As pointed out in previous sections, many growers report that the rotation delivery schedule operated by the water management organization is too restrictive and not timely matching the actual crop water needs for the whole study area. Also the conditions of water delivery, namely the flow rate and pressure heads at hydrants, are not those required at farm level for proper operation of the on-farm irrigation systems. Thus, the current irrigation deliveries are not adequate for the existing farming conditions. Being the aquifer relatively shallow in the whole study area, a large number of farmers in the last ten years started drilling on-farm tube-wells and using them for irrigation purposes. This led to a very large number of irrigation wells, either licensed or unlicensed, which pump irrigation water from the aquifer during large part of the irrigation season, with the highest concentration of withdrawals occurring during the months of July and August. As in many other coastal areas, the over-exploitation of groundwater resources causes drawdown of water levels and gives a way to intrusion of seawater. Besides the aquifer depletion beyond the safe yield, this also results in the qualitative deterioration of groundwater due to salinization. This also results in an increasing process of salt build-up in the soils, as farmers utilize salinized groundwater to irrigate their cropped fields. Salts loads are thus progressively brought onto cropped soils through the irrigation water. If leaching is not properly conducted on a regular basis or in case salts are not flushed away from the root zones by means of winter rainfalls, soils are progressively subjected to salts build-up, which may negatively affect their productivity.

6. CONTROLLING FACTORS OF THE HAZARDOUS PHENOMENA AND MAGNITUDE OF IMPACTS

The aquifer over-exploitation depends upon the following factors:

- Crop water demand, which is tightly dependant on the evapo-transpiration demand, on water-holding capacity of soils, on the growth stage of crops, on the prevailing farming and irrigation practices, and on the effective rainfall (amount of rainfall stored in the depth of the root zone). In the study area a concentration of crop water demand is observed for the months of July and August. Also most of farms are not equipped with storage facilities (holding ponds) which may allow buffering crop demand with the actual water delivery.

- The adopted delivery schedule, which basically depends on the available flow rate, on the water demand, on the design and capacity of the existing distribution network, as well as on operational resources and skills provided by the technical staff of the WUO. In the study area irrigation water is distributed to farms by a fix 10-day rotation. The rotation is not agreed with farmers and is dictated by a supply-driven approach by the WUO. More flexible arranged deliveries would allow partially overcoming the rigidity of the water distribution.

- On-farm irrigation practices, which can range from full satisfaction of irrigation requirements to different levels of deficit irrigation, based on the crops grown, on the specific sensitivity of the different growth stages to water deficit, on the target yield, as well as on the farmers' ability in water management at field level. In the study area full irrigation is the most common irrigation practice. Crops are mainly irrigated by micro-irrigation methods, which allow maximizing crop yields even with the use of saline

water. Leaching of salts from soil top layers is usually not carried out by the majority of farmers, but salts flushing mainly occur due to fall and winter rains.

- Natural leaching and aquifer recharge, which depends mainly on rainfall (intensity and distribution), on vegetation cover, on soils' hydraulic features, and on slope. Natural leaching and partial aquifer recharge usually occur in the study area during fall and winter months but, as resulted from previous investigations, are not capable of avoiding the progressive salinity increase in the aquifer and salts build-up in the soil on the long run.

The overall magnitude of impacts was estimated based on three criteria, namely the spatial spread of impacts, the time-duration of impacts and the time necessary for the impacts to be onset. Within each of the above criteria, the impacts were assigned a partial score based on a scale ranging from 1 to 4. For instance, the scale related to the spatial spread of impacts assigns scores based on the following ratings: Nowhere (0%): score = 0; Localized (< 5%): score = 1; Scattered (5-15%): score = 2; Widespread (15-50%): score = 3; Throughout (> 50%): score = 4.

The overall magnitude of impacts resulted by multiplying the partial scores assigned for the three criteria, thus on a scoring scale ranging from 0 to 64, then classified from "negligible" (score 0) to "mild" (score 1-22) to "moderate" (score 23-43) to "severe (score 44-64). The calculated values for the magnitude of impacts are reported in Table 2.

Table 2. Impact magnitude estimates

Hazard	Receptor	Impact	Spatial scale	Temporal scale	Time of onset to impact	Overall magnitude
H1 Aquifer over-exploitation	R1 Aquifer	I1.1.1 Aquifer depletion	Throughout (> 50%) **4**	Medium term (5-20 years) **2**	Medium (1-10 years) **3**	Moderate **24**
		I1.1.2 Salinization by seawater intrusion	Throughout (> 50%) **4**	Medium term (5-20 years) **2**	Medium (1-10 years) **3**	Moderate **24**
	R2 Agricultural soils	I1.2.1 Salts build-up	Throughout (> 50%) **4**	Medium term (5-20 years) **2**	Immediate (0-1 year) **4**	Moderate **32**

7. ESTIMATION OF PROBABILITIES

The estimation of the overall probability of hazards is also based upon three criteria, respectively the probabilities of hazard occurring, of the receptors being exposed, and of harm resulting to the receptor. In each criteria, the probabilities were assessed and classified on a High (score 3) to Negligible (score 0) scale. The overall probabilities of hazards were obtained by combining the partial scores assigned in each criterion and then classifying the overall scores based on the following probability scale: Negligible (when score ~ 0); Low (when score = 1-9); Medium (when score = 10-18); High (when score = 19-27). The overall probabilities for the study area are those reported in Table 3

Table 3. Probability estimation

		H1	
Probability of hazard occurring	receptor independent	High (3)	
Probability of receptors being exposed	R1	High (3)	
	R2	High (3)	
Probability of harm occurring to receptor	R1	High (3)	
	R2	High (3)	
Overall probability		**H1.R1=27 (high)**	**H1.R2=27 (high)**

8. EVALUATION OF RISK SIGNIFICANCE

Risk significance is evaluated accounting for both the magnitude of consequences as well as the probability of effects occurring. In case of qualitative risk assessment, a simple two-ways entry matrix considering simultaneously probability and magnitude of consequences, such as the one reported in Table 4 can provide a consistent basis for decision-making.
The evaluated risk significances of the 3 analysed impacts for the case study are presented in Table 5. The results from the evaluation are then used to prioritize the most relevant risks and conduct options appraisal to identify feasible and consistent management solutions.
As for risk communication process, the results from the risk prioritization should be communicated to the technical staff and to the decision-makers of the WUO through thematic meetings. Also, outcomes from the evaluation of magnitude and probability and from the risk prioritization stage should be disseminated to farmers groups and to opinion leaders by means of extension service activities and specific field focus meetings.

9. OPTIONS APPRAISAL

Options appraisal consists in the identification of the most suitable risk management technique. This involves scoring, weighting and reporting the different risk management options, and comparing alternatives. There are various criteria for appraising the potential options. For the present study, the aspects considered in evaluating the risk management techniques are: a) social risk acceptability by stakeholders; b) technical feasibility; c) effectiveness in risk alleviation; d) duration of effects; e) costs for implementing the risk management options.
The results from conducting the options appraisal for the three major risks, namely aquifer quantitative depletion, aquifer degradation, and salts build-up in the agricultural soils are reported in Table 6a, 6b and 6c.

Table 4. Risk significance evaluation matrix.

Increasing acceptability →	Consequences			
	Severe	Moderate	Mild	Negligible
Probability				
High	high	high	medium/low	near zero
Medium	high	medium	low	near zero
Low	high/medium	medium/low	low	near zero
Negligible	high/medium/low	medium/low	low	near zero

Table 5. Risk Significance for the study area

Risk	Significance score
Risk (H1. R1.I1.1)	Moderate x High = **High**
Risk (H1. R1. I1.2)	Moderate x High = **High**
Risk (H1. R2. I2.1)	Moderate x High = **High**

Table 6a. Risk Management option selection matrix for aquifer depletion (Risk 1)

Risk 1 **Aquifer depletion**	**Timing** Instant result to progressive	**Social acceptability** (-- to ++)	**Feasibility** (-- to ++)	**Effectiveness in risk alleviation** (-- to ++)	**Duration**	**Cost** Low to high
Business as usual (zero option)	Never - -	Acceptable +	Very feasible ++	Very ineffective - -	Never - -	Very affordable ++
Limit water pumping from Groundwater	Immediate +	Unacceptable -	Feasible +	Very effective ++	Short term -	Unaffordable -
Improved rotation in water delivery	Medium +/-	Acceptable +	Feasible +	Effective +	Medium term +/-	Very affordable ++
Decrease water tariffs by WUO to compensate for pumping costs	Long term -	Very Acceptable ++	Feasible +	Effective +	Short term -	Unaffordable _
Water delivery on-demand	Medium +/-	Acceptable +	Feasible +	Very Effective ++	Medium term +/-	Affordable +

Table 6b. Risk Management option selection matrix for aquifer salinization (Risk 2)

Risk II **Aquifer salinization**	**Timing** Short term to permanent solution	**Social acceptability** (-- to ++)	**Feasibility** (-- to ++)	**Effectiveness in risk alleviation** (-- to ++)	**Duration** Instant result to progressive	**Cost** Low to high
Business as usual (zero option)	Never - -	Acceptable +	Very feasible ++	Very ineffective - -	Never - -	Very affordable ++
Stop groundwater pumping	Medium +/-	Very unacceptable - -	Feasible +	Very effective ++	Medium +/-	Unaffordable -
Limit groundwater pumping to safe yield of aquifer	Medium +/-	Unacceptable -	Feasible +	Effective +	Medium +/-	Unaffordable -
Rotation irrigation delivery + conjunctive use	Medium +/-	Acceptable +/-	Feasible +	Effective +	Medium +/-	Affordable +
Irrigation delivery on-demand	Medium +/-	Acceptable +	Feasible +	Effective +	Medium +/-	Affordable +
Artificial aquifer recharge	Immediate ++	Neither unacceptable nor acceptable +/-	Feasible +	Very effective ++	Medium +/-	Affrodable +

Table 6c. Risk Management option selection matrix for salts build-up in the agricultural soils (Risk 3)

Risk III **Salts build-up in the agricultural soils**	**Timing** Instant result to progressive	**Social acceptability** (-- to ++)	**Feasibility** (-- to ++)	**Effectiveness in risk alleviation** (-- to ++)	**Duration**	**Cost** Low to high
Business as usual (zero option)	Never - -	Acceptable +	Very feasible ++	Very ineffective - -	Never - -	Very affordable ++
Improved rotation delivery	Long term -	Acceptable ++	Very feasible ++	Effective +	Medium Term +	Affordable +
Improved rotation delivery + conjunctive use	Medium +/-	Very acceptable ++	Feasible +	Effective +	Medium Term +	Affordable +
Irrigation delivery on-demand	Medium +/-	Very acceptable ++	Feasible +	Very effective ++	Long term ++	Affordable +
Improved on-farm irrigation practices (leaching)	Medium +/-	Neither unacceptable nor acceptable +/-	Feasible +	Very effective ++	Long term +	Very affordable ++
On-demand delivery + leaching	Immediate +	Acceptable +	Feasible +	Very effective ++	Long term +	Affordable +

10. CONCLUSIONS

Selecting a preferred risk management option strongly depends on the weights attributed by the evaluator to the decision criteria for the different options with respect to the zero-alternative (business-as-usual). Some of the identified management options relate to alternative operation of the large-scale distribution network, whereas some others entail improved water management practices at the farm scale or mixed options.
As for the risk related to aquifer quantitative depletion, the preferred option could be to operate the distribution network by an improved rotation delivery, which could better match crop water requirements in terms of timing of delivery. This would require some accurate estimation of irrigation requirements and improved irrigation scheduling plan, as well as some extension service activities.
As for the risk of aquifer salinization, since it is tightly linked to the amount and concentration of groundwater pumping during the irrigation season, conducting artificial aquifer recharge would be very effective in reducing pressure over the aquifer. For mitigating the existing effects on aquifer salinity, a strong reduction in groundwater pumping should also be enforced along with artificial aquifer recharge. These two measures in conjunction will allow decreasing the existing level of salinity and inverting the trend of salinity increase in the whole study area.
Concerning the risk of salts build-up in the agricultural soils, the on-demand delivery in conjunction with improved irrigation practices (leaching) at the farm level would result as the best management options. These would entail some modernization works to the irrigation distribution network as well as intensive extension service activities for training farmers on the field on issues related to soil-water balance and salinity balance for the major crops grown in the area and for the prevailing farming practices and irrigation methods.
The selection of the most suitable risk management option would be a matter of strategic planning by the Regional Administration and by the Water User Organization as well as of the available funds, human resources and skills available and required for implementing the options.

Combining the risk management options for the above three risks would result in putting together conflicting objectives for different stakeholders that may be involved in the land planning and land use. Land users may in fact primarily or only be interested in mitigating the risk of salts build-up in the cropped soils, whereas land planners as well as actors responsible for sustainable use of natural resources would try addressing broader objectives with high priority, such as the reduction of aquifer depletion and of salinization.

ACKNOWLEDGEMENTS

This research was funded under INTERREG IIIB CADSES project STRiM (www.strim.eu). The authors would like to acknowledge the contributions of the entire STRiM consortium and participating stakeholders.

REFERENCES

Capaccionia B., Diderob M., Palettab C. and Diderob L., 2005. Saline intrusion and refreshing in a multi-layer coastal aquifer in the Catania Plain (Sicily, Southern Italy): dynamics of degradation processes according to the hydro-chemical characteristics of ground waters. Journal of Hydrology, 307: 1–16.

DEFRA, Department for Environment, Food and Rural Affaires, Guidelines for Environmental Risk Assessment and Management, HMSO, 2002.

Lamaddalena N., Ciollaro G., Pereira L.S., 1995. Effect of changing irrigation delivery schedules during periods of limited availability of water. J. Agric. Engineering Research, 61: 261-266.

Lamaddalena N., Sagardoy J.A., 2000. Performance Analysis of on-demand pressurized irrigation systems. FAO Irrigation and Drainage Paper n. 59, Rome, pp 132.

Merriam J.L., Styles S.W., and Freeman B.J., 2007. Flexible irrigation systems: concept, design and application. Journal of Irrigation and Drainage Engineering, Vol.133 No. 1, February 2007.

Paniconi, C., Khlaifi, I., Lecca, G., Giacomelli, A. and Tarhouni, J. (2001). *Modeling and analysis of seawater intrusion in the coastal aquifer of eastern Cap- Bon, Tunisia*. National Agronomic Institute of Tunisia (INAT) and Center of Advanced Studies, Research and Development in Sardinia (CRS4), Italy, Kluwer Academic publishers.Transport in porous media 43: 3 – 28.

Plusquellec, H.L. C.M. Burt & H.W. Wolter. 1994. Modern water control in irrigation. Concepts, issues, and applications. World Bank Technical Paper 246. Irrigation and Drainage series. World Bank.

Small, L.E., Svendsen, M. 1992. A framework for assessing irrigation performance. Working Papers on Irrigation Performance No. 1. Washinghton, D.C., International Food Policy Research Institute.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Geo-communication for risk assessment and catastrophe prevention of flood events in the coastal areas of Chennai

S. Glaser[a], R. Glaser[a], A. Drescher[a], C. Pfeiffer[a], E. Schliermann-Kraus[a], M. Lechner[a], J. Vencatesan[b]

[a] Department of Physical Geography, University of Freiburg, 79085 Freiburg, Germany (stephanie.glaser@geographie.uni-freiburg.de)
[b] Care Earth Research Institute, Chennai, India

Abstract: Risk assessment of environmental stressors in the context of information and communication technologies is an emerging field of scientific research, for example to gain an understanding of how floods will develop. Geo-information technologies are accepted as powerful tools in crisis situations. To be successful in risk assessment and catastrophe prevention, it is essential to follow an interdisciplinary approach combining natural scientific, socio-economic and technical knowledge. At the same time it is important to combine top down (planning, official institutions) with bottom up views of the affected people. Interconnecting computational resources, human interdisciplinary expertise and local knowledge, following a top down / bottom up approach, can yield significant benefit in catastrophe prevention. Once the scientific understanding has been gained, a second phase is necessary in which geo-visualisation is important: The message has to be communicated effectively. This article explains how geo-communication and geo-visualisation can foster successful catastrophe prevention to arbitrate between different stakeholders supporting decision making processes. With these policy goals in mind the research results of this study were merged in an ogc-conform, server-based geo-communication portal using several flash based visualisation techniques. Thereby, the use of web-based mapping systems and an interactive participatory mapping tool proved to be an effective technique for risk evaluation.

***Keywords**:* catastrophe prevention, participation, decision-making, geo-communication, geo-visualisation

1. INTRODUCTION

Flooding is one of the most widespread climatic hazards and poses multiple risks to humans. In the south Indian megacity Chennai, disastrous tropical monsoons linked with excessive precipitation frequently lead to wide-spread floods. Caused by rapid urbanisation, the population in urban and periurban areas is more and more affected by these events. Besides the marginalised population living in disfavoured areas, increasingly also the more wealthy population that settles in flood prone areas is affected. The situation is characterized by complex interrelations of anthropogenic and ecological factors that were analysed through an interdisciplinary risk and vulnerability assessment by an international team of scientists.

Much of the data, information and knowledge for risk assessment are geospatial in nature. Preventing, preparing for, responding to, and recovering from natural and human-induced disasters all require access to geographic, climatological, as well as socio-economic data. Increasing data heterogeneity, fragmentation and volume, coupled with complex connections among specialists, researchers, planers and affected people in disaster

response, mitigation, and recovery situations, demand new approaches for information technology to enable long-term catastrophe prevention [Tomaszewski, 2007]. For this, extracting relevant content and information visualisation techniques are necessary to facilitate visual exploration, analysis, synthesis and especially processes (e.g. of land use changes, impacts of urbanisation or changing adaptation mechanisms) and in order to interact with, to combine and to exchange complex data sets and information. To meet scientific as well as planning and social needs, geo-visualisation and geo-communication can be considered as an auspicious approach. It is a research field that develops visual methods and tools for information preparation and presentation. One aim is to encourage collaboration between researchers and to make research results relevant to the tasks at hand. But above all, geo-visualisation and geo-communication can encourage participative processes and foster communication, support knowledge exchange and decision making between different stakeholders required for successful catastrophe prevention.

2. GEO-COMMUNICATION AND GEO-VISUALISATION – CONNECTING PEOPLE AND DATA THROUGH INFORMATION EXCHANGE

Geo-communication is quite a new research field. It includes different disciplines, like geography and cartography, as well as information and communication sciences, visualisation techniques and multimedia design. One of its major interests is to communicate geographical knowledge and spatial information across to a recipient using methods of information transfer. To implement these specific methods of communication, visualisation plays an important role. Visualisation means “exemplification”, and generally stands for all techniques that process and visually represent abstract data and coherences to support an effective data analysis. Visualisations try to activate the visual skills and tendencies of humans. This leads to better and easier identification of correlations, and an extraction of essential information not immediately coherent out of crude data. Above all, visualisation in sciences is a powerful tool for exploration, analysis, synthesis and presentation of research results, and can support scientific research processes in visual-mental activity and communication. To get an ideal data presentation, visualisation includes research cognitions of human data processing. Visualisation is an especially promising approach to transferring and exchanging information and knowledge to laities, because it can scale down complex circumstances to clear structures.

Rapid advances in web-cartography and geographic information systems, as well as in information and communication technologies, have created a potential for dynamic and interactive visualization methods which are increasingly used by decision makers and planners. Especially using current web2.0-techniques in geo-communication portals enables decision-makers to make their decisions through exploring a broad range of data. Meanwhile a wide set of tools have been developed that allow for visual exploration. These allow for processes to interact, to combine and to exchange complex data sets and information. Providing effective tools for geo-communication is not only a technical challenge. In particular Lars Brodersen [2005, 2007] has developed an important geo-communication model based on established theories of communication science. These specific spatial communication models need to be developed further to provide effective geo-communication tools and portals.

Recognizing the importance of understanding between different pressure groups for successful catastrophe prevention, geo-visualisation was used as a communication interface for intermediation. Information and knowledge exchange between heterogeneous groups is always associated with problems, especially in developing countries. The communication and information transfer in both directions often fail, due to a lack of communication channels. The visualisation approach offers an auspicious opportunity to combine both local views, by applying a bottom-up approach, as well as institutional perspectives, following a top-down approach. Furthermore, expert information can be expressed in varying levels of abstraction to fit the needs of the respective user, and to make it more useable and intelligible. The multi-media based preparation and presentation of the research results combine information and attitudes of stakeholders on various levels to support mutual solution-finding processes. This proved to be an effective technique to support decisions-making in catastrophe prevention.

3. THEORETICAL FRAMEWORK AND METHODOLOGY

3. 1 Workflow: Flood risk assessment using an interdisciplinary approach

As mentioned, both flood risk assessment and prevention need a holistic approach regarding human and environmental aspects. The conceptual framework (Figure 1) shows the multi-layered approach of this study combining physio-geographic analysis, remote sensing and socioeconomic data.

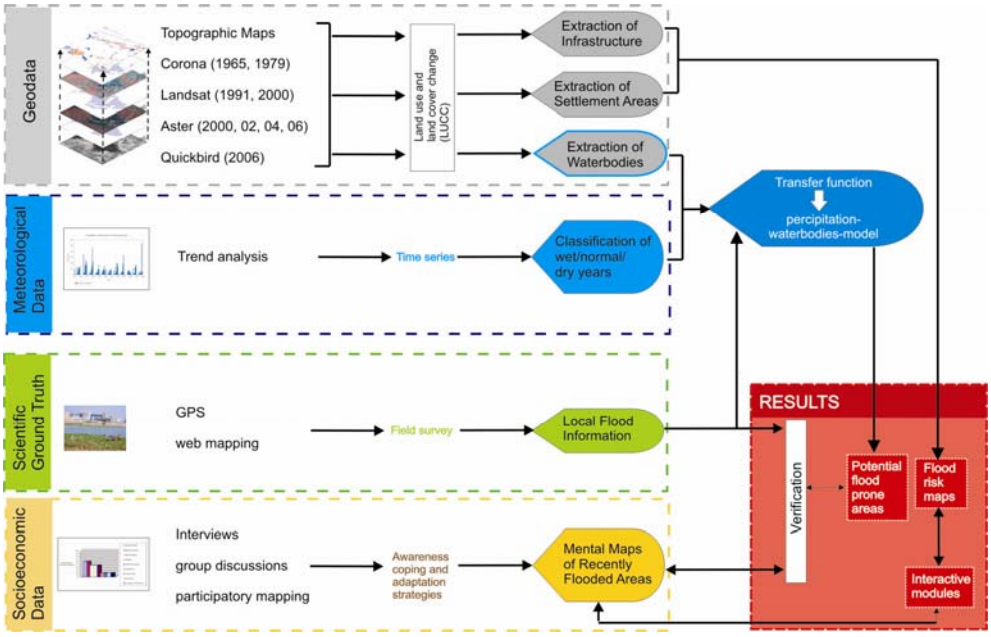


Figure 1: Workflow diagram - Interdisciplinary approach

Remote sensing data in different resolutions, (e.g. spectral and temporal ranges) were used to analyse the changes in settlement areas, infrastructure and the extent of water bodies from different periods of time since the 1960's up to now.
Early Corona-images from 1965 and 1979 were used to identify the state of natural water bodies and the dimensions of the former settlement. Multi-temporal Landsat images from 1991 and 2000, as well as repeated images from Aster, taken in 2000, 2002, 2004 and 2005, were used to receive time-dependent information about the dynamics of the main water body, the marshland in the South of Chennai. These were combined with long-term meteorological precipitation data to build a regression-based transfer function, showing the relations between rainfall and the extent of the water bodies in dry, normal, wet and extreme-rainfall years. Local flood information was gathered in transect walks and through a web-based mapping system, which also was used for verification.

In order to learn about risk perception and management, a multilevel approach was selected. This means focusing on quantitative macro-level information such as census data, but also on qualitative micro-level information for example, key-informant-interviews, group discussions and participatory mapping. An explorative, multi-methodological approach on different levels covers the complexity of attitudes and responses. Mental maps represent people's perception of their own environment and provide additional insights in local cognitions of flooding. The mapping was started by identifying community boundaries and infrastructure. In order to learn about the historic dimensions of floods, time-lines were selected. In doing so, outstanding floods of the last twenty years were

demarcated. Based on this information, discussion took place about response strategies used before, during and after floods. The mapping activities proved to be very helpful in communities with low literacy since the visualisation enabled the participants to overcome barriers of communication.

3. 2 Theory: Flood risk assessment using a risk and vulnerability framework

Flood risk assessment is theoretically embedded in a risk and vulnerability framework. Risk and vulnerability concepts are not new in disaster research. In the past, researchers acknowledged the fact that it is not sufficient to look at environmental exposure and stressors only, since this leaves out information about different human response strategies, as well as impacts [Bohle, 2001]. Turner *et al.* [2003] developed an expanded vulnerability framework, which sets the focus on the multifaceted synergies and linkages between human and environmental aspects. This framework takes into consideration that in case of a disasters, such as floods, vulnerability is not determined by exposure to risk only, but also by the sensitivity and resilience, as well as by coping and adaptation strategies of those affected. This interdisciplinary approach provides information about the complexity of human-environment systems.

Within the framework, a hazard, e.g. floods, is a thread to a system. The impact of a hazard depends on the exposure of a system. Exposure can be the frequency, magnitude and duration of floods. Sensitivity means the consequences for a population considering their human conditions – e.g. human capital and endowment - as well as environmental conditions – e.g. climate and soil. Resilience patterns are developed depending on the sensitivity. This includes coping as well as adaptation strategies. According to Turner, social vulnerability can be defined as the combination between exposure, sensitivity and resilience of a system [Turner *et al.*, 2003; Glaser *et al.*, 2006]. The interdisciplinary analysis of the complexity of risk and vulnerability is crucial for this study. The focus lies on the exposure to floods in PM, risk perceptions of those affected as well as their coping and adaptation strategies. The results are combined and visualised in the conclusion using a table based on Turners framework.

The Turner model visualises the linkages and dependencies between causes and consequences of floods in Tamil Nadu (Figure 2). The framework visualises the linkages and dependencies between causes (highlighted in red) and consequences (highlighted in black) of floods in Tamil Nadu. Hereby, three levels are covered: the international (green background), the national (yellow background) and the local level (blue background). Global socio-economic changes reflect in national population growth and labour-migration, environmental changes result in changing monsoon patterns and climate variability. These developments highly impact on local planning capacities of the government. The sensitivity of those exposed to floods varies. It reflects in the frequency and intensity of vector and water born diseases, loss of income, loss of schooling and psychological stress. Depending on the livelihood background (especially the financial and social resources) different coping and adaptation strategies are selected by the affected communities. However, on a household level strategies are limited. Although sustainable prevention measures and proper city planning would be needed, the government only offers rehabilitation services. The lack of national and local strategies increases the vulnerability to current environmental as well as socioeconomic changes, at the same time these changes put a burden on government authorities.

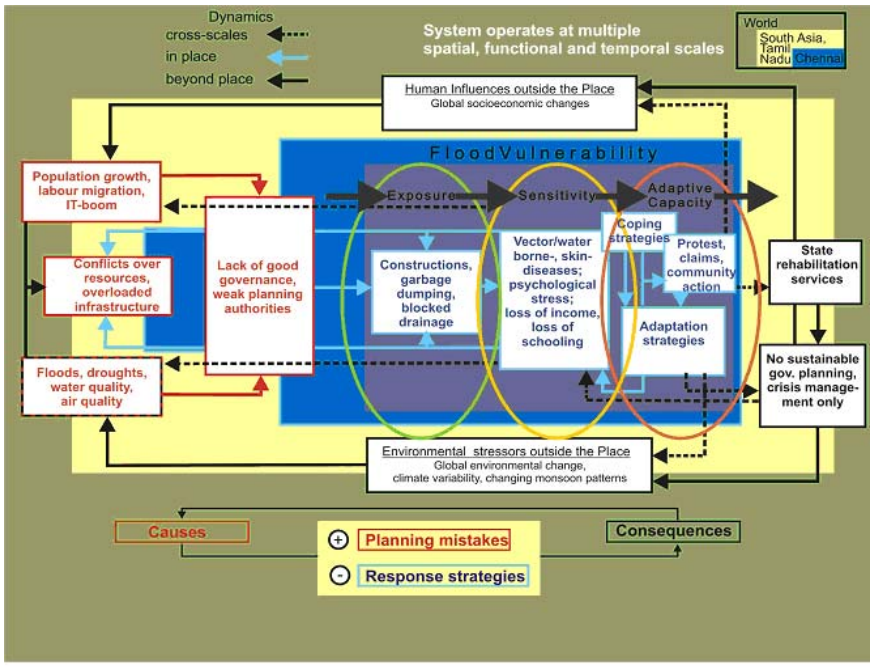


Fig. 2: Modified framework according to Turner *et al.* [2003]

4. GENERAL RESEARCH RESULTS OF THE STUDY

The analysis of the meteorological data clearly shows that there is an outstanding year-to – year variability, but also medium-term fluctuations on a decadal scale over the last 200 years on the eastern coast of India. But in the long-term perspective, there are no significant upwards nor downwards trends during the last 200 years [Walsh *et al.*, 1999]. The last 30 years are characterised by a decrease of annual precipitation, while the number of floods affecting the inhabitants of Chennai is increasing. The reason for these floods is the high monthly total precipitation that may exceed the total yearly rainfall. In some cases, even the daily total rainfall exceeds the average monthly amount. But most importantly, the statistically derived relation between the extent of the water body and the rainfall amount gives clear evidence that the flood risk in the south of Chennai is increasingly due to man-made mismanagement in the area of the Pallikaranai Marsh. Building activities in and around the Marshland cut off the water run off, and therefore aggravated the frequency and intensity of floods, even though the precipitation amounts are decreasing. Without a sustainable planning strategy, especially facing the natural run-off due to single strong rainfall events, Chennai will suffer from more devastating floods in the future.

The socio-economic analysis of risk perceptions and management strategies underline the interrelated reasons for floods, which are embedded in a broader context of economic globalisation, labour migration and rapid urbanisation. The pressure on cities and their planning authorities grows as the number of their inhabitants increases. The demand for developing new areas leads to the ecological destruction of the marshland close to Chennai. Planning authorities are not reacting accordingly. Illegal as well as legal garbage dumping into the wetland is also an increasing hazard. Existing infrastructure are not properly maintained. Canals are dysfunctional; storm water drains are not cleaned regularly; water bodies are not maintained. Uncontrolled construction of IT-companies and construction of private citizens, both legal and illegal, are in complete ignorance and violation of the norms and rules of the government. The combination of these various factors is resulting in a collapse of natural drainage systems.

5. VISUALISATION OF THE RESULTS - A KNOWLEDGE-BASED GEO-COMMUNICATION PORTAL FOR DECISION-MAKING IN CATASTROPHE PREVENTION

The results of this study were merged and visualized in an information portal that is publicly accessible, and can be expanded by the project partners in Chennai (compare Figure 3).
Technically it is based on a Flash-WebGIS- and a Mapbender-WebGIS-Client using the ogc-standards and services web-map-service (WMS) and transactional web-feature-service (WFS, WFS-T). ogc is a nonprofit, international standards organization that is leading the development of standards for geographic data related operations and services. By the use of these specifications the exchange of geo-information between different systems is assured. They also provide interfaces to web-services for data maintenance and supply. The consideration of ogc-standards guarantees a sustainable use and durable free access to data, as well as a transparent integration of heterogeneous, distributed geodata services. For data storage a PostGIS database is used.

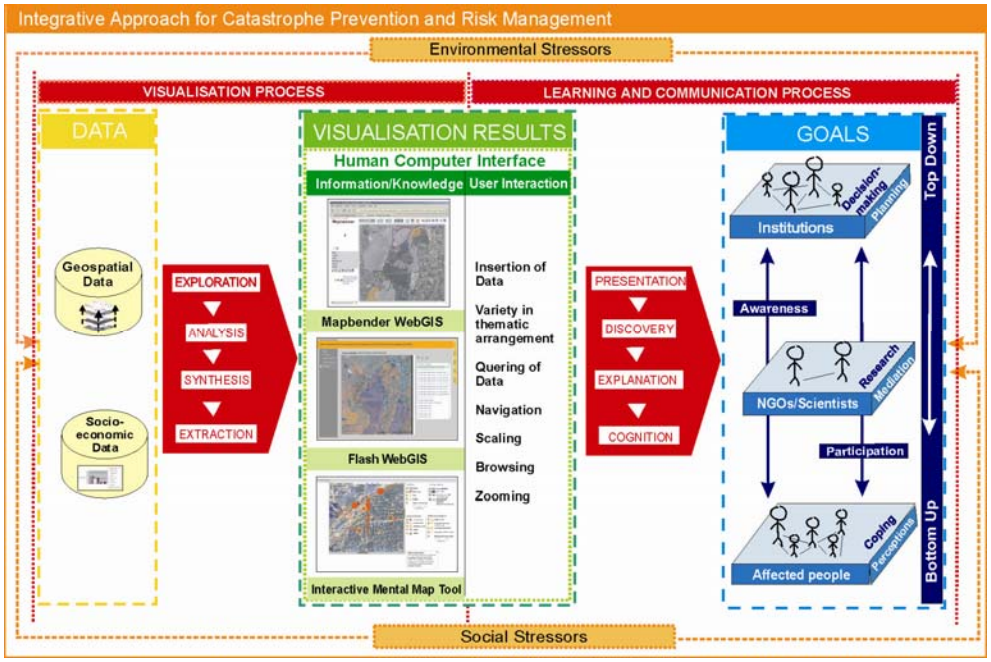


Figure 3: Visualisation process and resulting learning and communication processes

The visualisations are implemented with the authoring software Adobe Flash, which is specialised in developing interactive web content. The integrated programming language Action Script enables wide scope for designing interactive dynamic maps. The developed animations can be presented via the internet, or as offline solution. Flash is platform independent and can be used as browser plug-in, but also as a stand-alone solution. The Flash-Plug-in is available on 98% of users' computer desktops – more than any other similar plug-in technology. It is also easy and quick to install. Additionally, map files (.SWF files) are up to about 25% smaller than equivalent Java-based files, which contributes to a faster user experience. The widespread accessibility and the easy usability of the Flash plug-in turns Adobe Flash into an interesting visualisation tool, especially for use in development cooperation. It also features the use of vector and raster graphics, which is important for working with geographic data.

The geoportal can be divided in two parts: The first part is based on two WebGIS-Applications (5.1). The second part of the information portal offers different flash-based web-animation techniques, to provide an understanding of the processes and results of the

project. This part of the portal also includes a data page for accessing all acquired data, including original hand drawn mental maps, results of interviews and group discussions. In addition, a media page is integrated with an image database, with video and audio sequences of key informant and group interviews from the research area. A flash animation based on satellite images and geographic data gives an easy access to the quiet impressive land use and land cover change in the region. The interactive mental map tool (5.2) proved to be especially useful for stakeholder communication.

5. 1 Mapbender and Flash WebGIS-application

Mapbender is an open source WebGIS-application implemented in PHP and JavaScript for managing spatial data that are standardized following the ogc specifications. Data storage is organized with a PostgresSQL, PostGIS or MySQL database. The framework implements user management, authentication and authorization. Management interfaces, depending on authorization and demand of the user, are stored as configurations in the database. The Mapbender client-software can be used to display, overlay, edit and manage geodata.

The implementation of this tool is a response to the difficulty of obtaining satellite images that are taken at the time of flood events. To overcome the missing data, local stakeholders were familiarised with the program. The capacity building process enabled, and still enables, local partners to give their input of the exact area where floods are happening. After the transfer to the database in Freiburg, using ogc-compliant Web Services, the geodata can be accessed directly from expert tools, as well as any client-application qualified for working with GIS-web-services. So it can also easily be integrated in the research analysis. Through this process the acquisition and integration of the knowledge of local project partners were made possible.

For getting access to all acquired geodata, we provided a flash based mapping solution giving users the possibility to discover, explore and visualise different thematics, like for example: settlements, water bodies, infrastructures from different years, as well as land use and land cover change documented by the remote sensing data.

5. 2 Flash based interactive mental map tool

Geo-information technologies have long been accepted as being important for catastrophe prevention. Although the advantage of working with participatory maps is known, in the past visualisation techniques were rarely used for information preparation und presentation of hand drawn mental maps. In our project we developed an interactive mental map tool, by combining multiple sources of hand drawn maps with various geographical and socioeconomic data. In order to get insights into local views and understandings of flooding, mental maps drawn by affected community members were used. Hereby, existing infrastructure, as well as flood related time lines were demarcated. The time lines aimed at illustrating the historical dimensions of outstanding floods within the last twenty years. During the mapping activity, discussions about response strategies used before, during and after floods were initiated. Mental maps proved to be particularly helpful for illiterate community members who could visualise, reflect and communicate their local knowledge and experiences. The obtained hand drawn mental maps were then transferred in a digital version for more useful and precise analysis. Following cartographic rules, a comprehensive and generally understandable legend was defined by translating information from group discussions and the hand drawn information into icons, symbols, lines, polygons and points. The digital mental maps proved to be a very accurate transformation of the original information into digital readable elements. To compare the mental maps with the topographic reality we used a Quickbird satellite image as an overlay. The combination with geo-referenced highly resolved satellite imagery showed an astonishing accordance. The ground-truthing functioned as an additional source of verification of the basic information. This successful translation from raw data, such as hand drawn mental maps, digital mental maps and satellite images, did result in an interactive mapping tool that merges interdisciplinary data on infrastructure, local perceptions, coping and

adaptation strategies, with remote sensing data and modern map making technology. The integration of manifold information in one interactive mapping tool led to a more realistic presentation of the situation. The interactive character of this mapping tool proved to be very efficient and effective in communicating research results. It initiated bottom-up information exchange on various levels, from community members up to city planners, and therefore helped to improve and support sustainable planning and decision making processes.

6. CONCLUSION

These results, especially the geo-communication portal, had been presented at two policy workshops in Chennai, India and in Freiburg, Germany. The aim was to bring together different stakeholders, representatives of planning authorities, administrative managers, resident welfare organisations, as well as community members who are personally affected by the problems that have been identified throughout the research, to address the issue of the Pallikaranai Marsh.
The visualisation of the research results helped enormously in expanding the awareness of many participants of the extent of local problems, especially for politicians and city planners. Additionally, they provided new insights and arguments for human rights activists and conservationists. Generally, it was pointed out during the workshop that a better accessibility of reliable data was crucial for the improvement of planning all over India. In this context, the information portal with publicly available data is a very useful tool that could serve as an example for future projects.
A second important achievement of this study was the interdisciplinary approach that has succeeded in offering new possibilities of looking at land use conflicts, their possible resolutions and catastrophe prevention. Another goal of the project was to establish a sustainable tool for flood risk assessment in Chennai, which allows a continuous integration of new data, even after the research had been finished. The chosen digital setting gives the involved experts, under the supervision of the NGO Care Earth, the chance to monitor future flood events, and to suggest appropriate action on site and provide in-time information to relevant authorities. In order to guarantee sustainable processes, continuous capacity building of the local partners was initiated and implemented during the application of the Mapbender client software.

In order to understand and communicate scientific approaches more effectively, and to enhance the level of understanding between different stakeholders, geo-communication is a fundamental approach and geo-visualisation an essential technology.

ACKNOWLEDGEMENTS

We would like to thank the participants of the mapping activities and the interviews in the three study areas in Chennai. Additionally we thank the team of Care Earth for supporting us greatly during the data collection in Chennai. The underlying research was funded by the German Federal Ministry for Economic Cooperation and Development (BMZ) and commissioned by the Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH - Advisory Project “Disaster Risk Management in Development Cooperation” on behalf of the BMZ.

REFERENCES

Bohle, H.-G., Vulnerability and Criticality. Perspectives from Social Geography, *IHDP-Update (2)*, 1-5, 2001.

Brodersen, L., Semiotik i geokommunikation – fra virkelighed til handling, Tankegang. Frederikshavn, 2005, English Abstract.

Brodersen, L.: Geokommunikation. Tankegang. Frederikshavn, 2007.

Brodlie, K., Fairbairn, D., Kemp, Z., Schroeder, M., Connecting People, Data and Resources – Distributed Geovisualization, Dykes, J., MacEachren, A.M., Kraak, M.-J., *Exploring Geovisualization*, Elsevier. London, 425-443, 2005.

Cai, G., Sharma, R., MacEachren, A.M. and Brewer, I., Human-GIS Interaction Issues in Crisis Response, *Int. J. of Risk Assessment and Management 2006,* Vol. 6, No.4/5/6, 388-407, 2006.

Downs, R. and Spea, D., Kognitive Karten, Die Welt in unseren Köpfen. New York: Harper & Row, 1982.

Dreger, R.: Geokommunikation, Die Welt im virtuellen Raum. dieuniversitaet-online. *Sommerserie, Reisen und Wissenschaf,*. 1. September 2004, Wien, 2004. http://www.dieuniversitaet-online.at/dossiers/beitrag/news/geokommunikation-die-welt-im-virtuellen-raum/178.html (13.02.2008).

Few, R., Ahern, M., Matthies, F., Kovats, S., Floods, health and climate change: a strategic review, Norwich Tyndall Centre for Climate Change Research, 2004.

Glaser, R. *et. al*., Wasserkrisen in der Megacity Chennai in Südindien. In: Gebhardt, H., Glaser, R., Radtke, U., Reuber, P. (Ed.), *Geographie - Physische Geographie und Humangeographie*, Heidelberg: Spektrum, 1009-1014, 2004.

Jobst, M.: Interpretierte Realität – 3D-Kartografie als Hilfsmittel der Geokommunikation. In: Strobl, J., Blaschke, T. and Griesebner, G. (Ed.), *Angewandte Geoinformatik 2004. Beiträge zum 16. AGIT-Symposium Salzburg.* Wichmann. Heidelberg. 280-285, 2004.

MacEachren, A.M., Gahegan, M., Pike, W., Brewer, I., Cai, G., Lengerich, E., Hardistry, F., Geovisualization for knowledge construction and decision support, *Computer Graphics and Applications*, IEEE, Volume 24, Issue 1, 13-17, Jan-Feb 2004.

Nestle, M. *et. al.,* Chennai – Umweltkrise und Gesundheitsrisiken in einer indischen Megacity, Glaser, R & K. Kremb (Ed.), *Planet Erde*. WBG: Darmstadt, 209-214, 2005.

Stelzl, H., Raggam, H., Sacherer, O. and Almer A., Automatisierte 3D-Visualiserung von Geoinformation zur Unterstützung der Feuerprävention im mediterranen Raum. In: Strobl, J., Blaschke, T. and Griesebner, G. (Ed.), *Angewandte Geoinformatik 2004. Beiträge zum 16. AGIT-Symposium Salzburg.* Wichmann. Heidelberg. 644-649, 2004.

Tomaszewski, B.M., Weaver, C., Robinson, A.C., Stryker, M., MacEachern, A.M., Geovisual Analytics and Crisis Management, Van de Walle, B., Burghardt, P., Nieuwenhuis, C., *Proceedings of the 4th International ISCRAM Conference,* Delft 173-179, 2007.

Turner, B. L. *et al*., A framework for vulnerability analysis in sustainability science, *Proceedings of the National Academy of Sciences* (USA), 100 (14), 8074-8079, 2003.

Walsh, R.P.D., Glaser, R. and S. Militzer, The Climate of Madras During the Eighteenth Century.- *International Journal of Climatology*, 19: 1025-1047, 1999.

WHO, Public Health mapping and GIS, http://www.who.int/health_mapping/en/, (13.04.2008).

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Simulating earthquakes effects: the SHAKYGROUND simulation tool

M. Cristaldi[a], U. Delprato[a], C.Lupo [a], S. Gresta[b]
[a]IES Solutions Srl, Italy, www.i4es.it, www.react-ist.net, info@i4es.it
[b] Università degli Studi di Catania - Dipartimento di Scienze Geologiche, Italy, gresta@unict.it

Abstract: In this paper we describe SHAKYGROUND, a software tool that has been designed with the aim of providing the user a friendly environment allowing the simulation of strong ground motion by means of synthetic simulation of acceleration seismograms and the estimation of seismic load in a given area.

***Keywords**:* Earthquake effects, simulation code, seismic code, ground motion, wave propagation, urban planning.

1. INTRODUCTION

In a classical sense, earthquakes originate from the rapid release of elastic energy caused by the rupture of rock in the earth. Damaging effects due to an earthquake arise (i) from a source outcropping directly at the earths surface, (ii) from the radiated seismic waves. In practice the second effect is the most important. It occurs more frequently and the area shattered by seismic waves is by far larger than the zone which could be hit by an outcropping source.

Strong ground motion is seriously affected by wave propagation effects caused by changes due to absorption, reflection and refraction at the boundaries of the geological structures. In particular the subsurface geological structure is of principal importance in this context. Since there is no simple way to account for wave propagation effects we have calculated the transfer function of the propagation medium using Haskell (1953) matrices for a 1D-model.

SHAKYGROUND has been created with the scope to provide a reliable, easy to use, and versatile tool for the estimation of seismic strong ground motion. It forms the seismological base for the design of earthquake-resistant buildings and for land use planning. Moreover, it allows the identification of sites where the expected ground shaking is severe; this can be useful for emergency planning, as well as for enhancing earthquake awareness.

The bespoken software tool incorporates concepts of the seismic source process and wave propagation effects which are well confirmed by numerous scientific studies and practical applications. The problem of seismic load estimation is approached by means of synthetic simulation of acceleration seismograms, that gives the possibility to account for the specific geological site conditions and the characteristics of the seismotectonic zone.

The user's experience is exploited for establishing the model parameters rather than for the search of a "suitable" experimental example seismogram. The mentioned simulation tool allows to vary the geotechnical and source parameters randomly disclosing the possibility to evaluate the stability of the results statistically.

Finally the program produces a report of important simulated engineering seismological signal parameters together with response spectra which can be directly compared to standard seismic codes (for ex. EC8, US Building Code).

SHAKYGROUND is a software that runs on Microsoft Windows operating systems (95, 98, NT, 2000, XP, VISTA), and a database engine for the models and results storage. The user interface allows to enter the input parameters on "sheets". There are sheets related to the seismic source parameters, others concern the geotechnical parameters, the calculation of response spectra and some general information concerning the operation of program and its output. This system protects the user from errors due to misspellings and the use of improper parameters.

Accessing inputs from an external data base is a special feature that facilitates the use of the application in seismic zoning projects where a large number of sites and different seismic scenarios have to be considered. Earthquake parameters and typical sets of geotechnical parameters (called "stratigraphies") can be stored and accessed in "catalogs". The models (either derived or not from this catalogs) can be saved with the simulations results and they can be used by different users in a network and can be easily exported on the Internet/Intranet. It is possible to pass the afore mentioned sheets to the data base and to recombine them creating a new model and new sheets. This makes the setup of a model easy and fast, and avoids many errors occurring during the creation of suitable input parameters.

The results of the software tool are both numerical and graphical: response spectra, the simulated seismograms allow the user an immediate evaluation.

2. STRONG GROUND MOTION SIMULATION

Ground motion estimation for the purpose of seismic hazard and loading analysis may be eventually carried out in some empirical manner or on the basis of theoretical models. The empirical approaches essentially consist in the use of observed relations of strong ground motion parameters derived from various earthquakes zones. In a more refined manner the analyst tries to cull from a vast data set of strong ground motion records those accelerograms which duplicate the source, travel path and local site conditions. The difficulties coming along with this strategy are quite well-known: the empirical relations derived from records of other earthquake zones with differing seismotectonic setting, characteristics of wave propagation effects and subsurface site conditions can hardly be adjusted to the situation of actual interest.

Given the large variety of possible earthquake scenarios with respect to source parameters, effects of wave propagation and site conditions, the choice of "suitable" example seismograms poses severe questions as to the significance of these approaches. The drawbacks of the use of some empirical strategy is certainly exacerbated in earthquake zones with long quiescent periods as, e. g., Italy where the return periods of large damaging earthquakes range from tens to hundreds of years.

A possible way out of these problems, which was adopted for SHAKYGROUND, is the estimation of strong ground motion by means of synthetic simulation of acceleration seismograms. As shown for various earthquake zones and sites, synthetic simulation according to such concepts gives indeed a reasonable match of observed data, provided there is a sufficient knowledge of the source and geotechnical parameters (see, e. g., Langer et al., 1999; Gresta et al., 2004).

Global source parameters are accounted for by applying a band-pass filter to the Gaussian white noise, i.e. $C\ M_0\ S(f,f_0)\ P(f, fmax)$ where C is a constant for geometrical spreading and radiation pattern, M_0 the seismic moment of the event, f_0 the corner frequency, $S(f,f_0)=f^2/(1+(f/f_0)^2)$, $P(f,fmax)=(1+(f/fmax)^{2q})^{-1/2}$, q the parameter of the steepness of the high frequency decay (here q=4). The corner frequency f_0 can be related to the size of the source (its radius r_0) after Brune (1970) by: $f_0=0.372\ c/r_0$, where c is the shear-wave velocity. Finally the seismic stress drop is computed as $7M_0/(16r_0^3)$.

The strength of the synthetic approach resides in the possibility to account for the specific geological site conditions, the effects of wave propagation and the characteristics

of the seismotectonic zone. According to the strategy chosen in SHAKYGROUND the user's experience is exploited for establishing the model parameters rather than for searching of a suitable, experimental example seismogram. Given the stochastic nature of the source model, The program performs a number of simulations and calculates average, standard deviations and peak hold values of engineering seismological parameters.

Furthermore, it allows to randomly change the geotechnical and source parameters, thus enabling the statistical evaluation of the stability of the results. Finally SHAKYGROUND produces a report of relevant simulated engineering parameters of seismological signal together with response spectra which can be directly compared to standard seismic codes.

3. HOW IT WORKS

3.1 Data Input

The users can start a project typing in its name and, if needed, the customer's name (see figure 1).

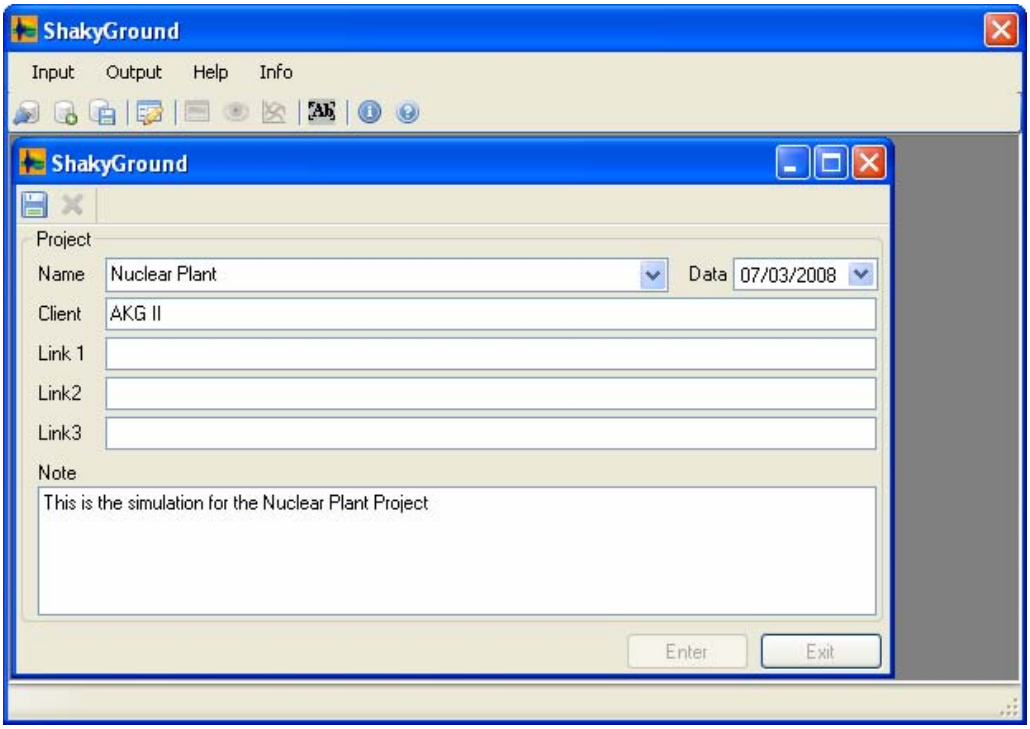


Figure 1. The Project Screen

After saving the data, the project is created and the user can start the data entry. Each project can be made by many different "models". In SHAKYGROUND, "models" are made up of a series of "sheets", and each of them is related to a specific item of the model, e.g. the seismic source, the geological structure, the response spectra and the general conditions for the calculus. A further sheet supports the configuration of the results. Such "sheets" concept has been adopted to provide a safe input, to avoid either misspellings or the use of incorrect parameters. The sheets can be stored in the program catalog and re-used for the creation of new models.

The first sheet focuses on the seismic source parameters and their position with respect to the receiver. The random fluctuation of the parameters may also be specified. The user can re-use a previously defined source by scrolling the sources presented in the catalog ("Sources in Catalog" box) and accessing the corresponding source parameters. Otherwise one can specify a new name in the box "Source Name:", enter a new set of parameters, and finally save the new source in the catalog.

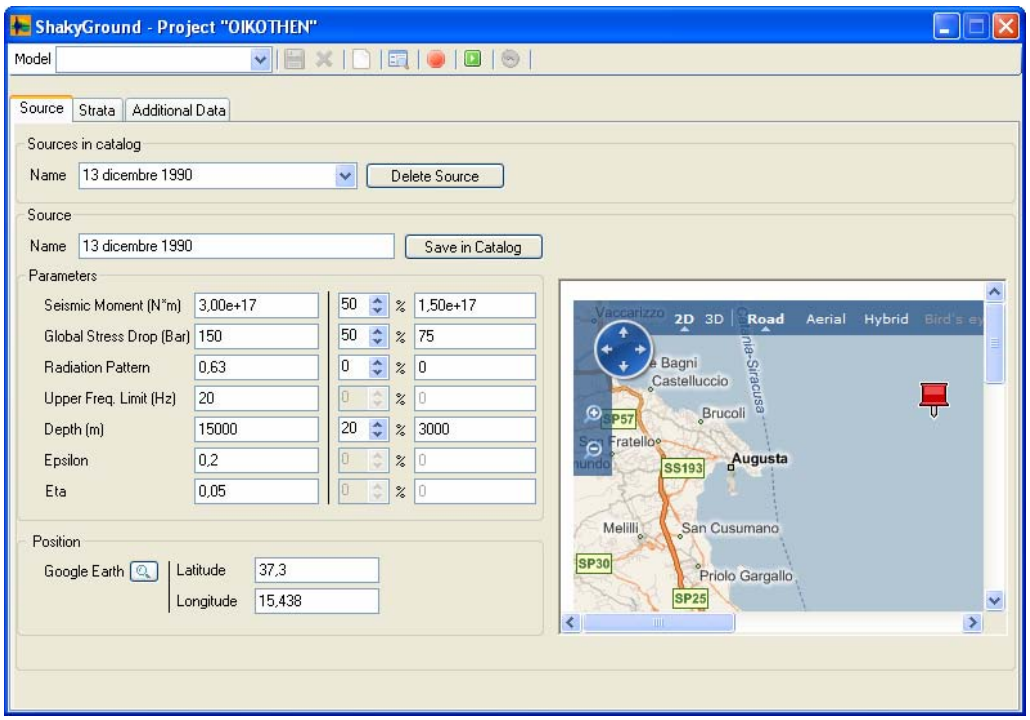


Figure 2. The Source Sheet

The software works together with Google Earth to allow for an easy location of an earthquake. The map of the event is presented in a browser screen on the right hand part of the "Source Sheet" window.

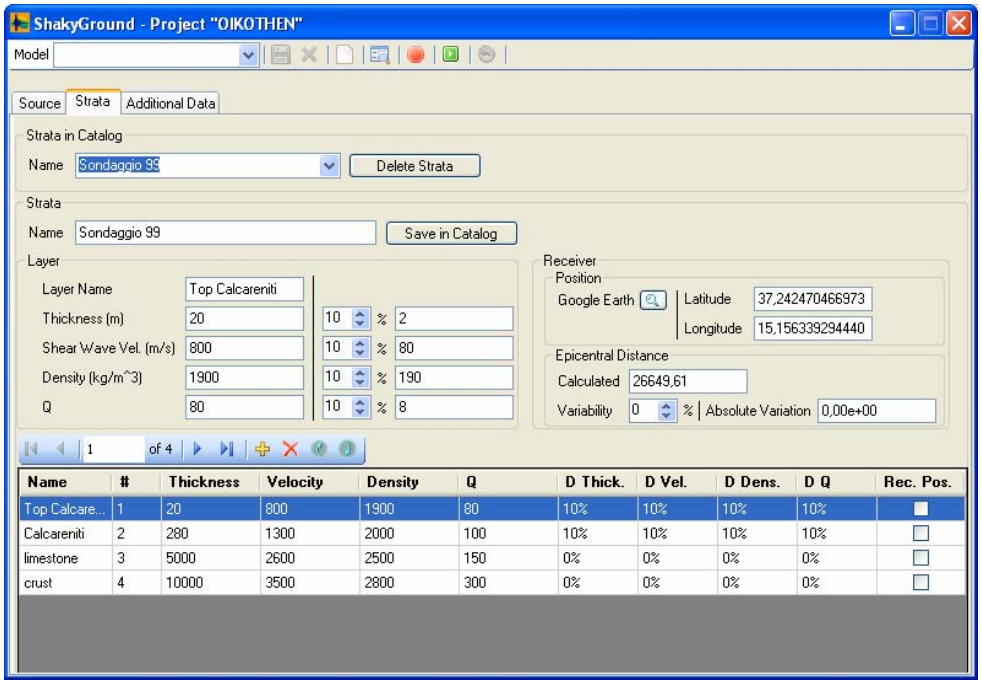


Figure 3. The Strata Sheet

The "Strata" sheet (Figure 3) is made up of two levels. In the box "Stratigraphy" one can select an already existing geological model consisting of a number of layers. The user can

define a new layer by using the box “Layer name” or the table below. A already explained for the source parameter sheet, the user can here specify the random fluctuation of the layer parameters. By clicking on the “Insert” button, the new layer is inserted in the layer stack. Any layers can be moved within the layer stack by selecting it in the box at the bottom and using the “Up” and “Down” buttons.

The created geological model can be saved in the catalog with a new name or overwriting an existing one. For zoning projects, the strategy of using pre-defined stratigraphies is particularly effective and safe, since with little effort the user can recombine typical geological conditions at the surface and in the structure underneath.

SHAKYGROUND allows to place the receiver either at the surface or at some interface within the layer stack. With this option the tool allows analysing the case of a building whose base is supposed to be situated at some depth. The default position is set at the ground surface.

The “Additional Data” sheet (Figure 4) contains three different areas: Response Spectrum, General and Output.

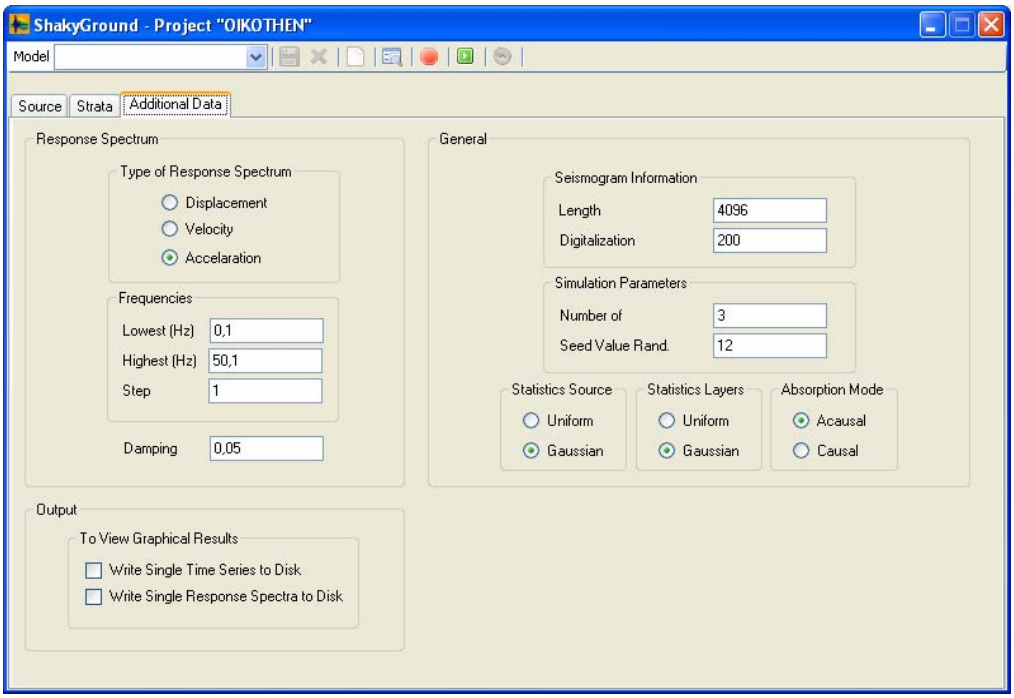


Figure 4. The “Additional Data” Sheet

In the “response spectrum” section the user can specify some settings used by the model for calculating the response spectra (type of the response spectrum, frequency bounds and resolution). The default response spectrum is “Acceleration” as required by most seismic regulations, but when needed velocity or displacement response spectra can be selected. The default value for the damping is set to 0.05, i.e., 5% of the critical damping. This damping value is usually considered in standard seismic regulations, but occasionally damping values of 1%, 2%, or 10% have also to be used.

The “General” section of the sheet contains basic information about how to perform the calculations. In the box “Seismogram Information” the user must specify the length of the seismogram and the sampling frequency. In the box “Simulation Parameters” the user selects the number of simulations carried out for the model and the seed value for the generator of the pseudo random numbers. Two further items (box “Statistics Source” and “Statistics Layers”) concern the mode how the program varies the model parameters randomly within the ranges selected in the source parameter and layer sheets. It is possible

to configure the model to calculate source and geotechnical parameters according to a uniform or a Gaussian distribution.

The "Output" section is used to specify how SHAKYGROUND should handle the output. Usually a short output, with a report about calculated signal parameters and response spectra, is sufficient. By ticking on "Write single Time series to disk", the program is configures to write each single synthetic seismogram is on the computer hard disk: such dataset and can be visualized at the end of the run. The same applies for the Response spectra. These features enable the user to access the single time series and response spectra, thus verifying the good choice of the parameters and the performances of the model. Since the time series and response spectra are recorded as plane ASCII files, they can be easily imported by other sophisticated data processing and visualization packages.

3.2 Data Output

SHAKYGROUND offers several levels of output which enable the user to immediately check and evaluate the results of a simulation. After completing the numerical processing, the standard output is generated (example displayed in Figure 5).

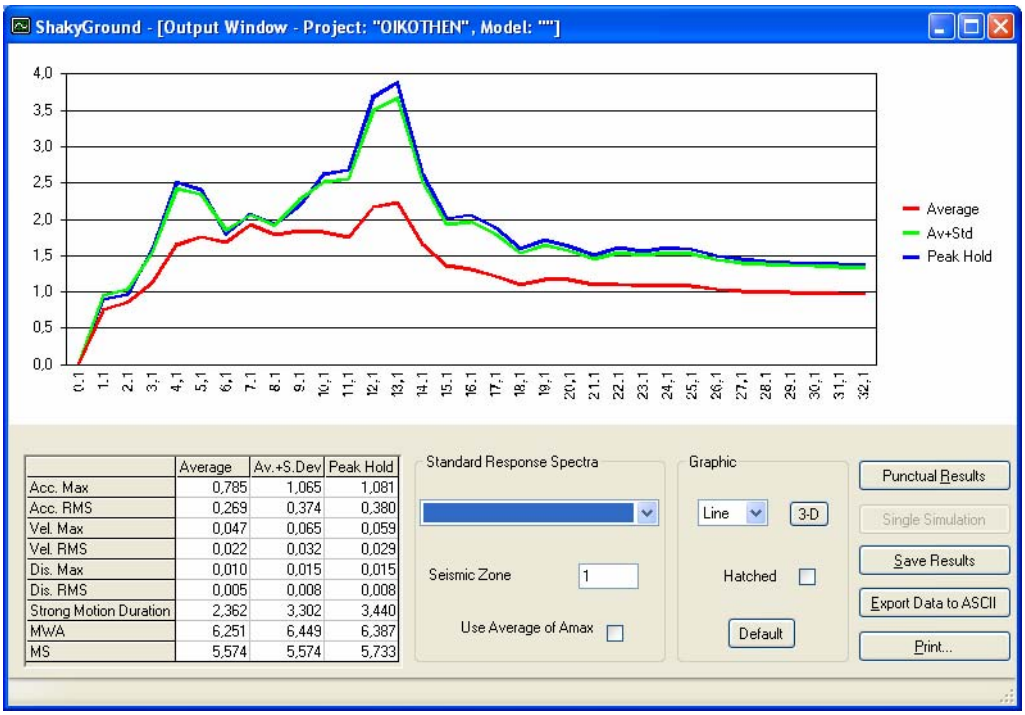


Figure 5. Output Window for SHAKYGROUND

It consists of a plot of the response spectra and their statistical variation, and of a table with important signal parameters, i.e., peak ground acceleration, velocity and displacement, RMS-values of the three types of signal, the duration of strong ground motion (calculated following the concept of HUSID-plots), finally corresponding magnitude values (MS deduced from the seismic moment, MWA as obtained from the simulated signals). All these results can be saved to a georeferenced data base, where they can be easily accessed for further processing with GIS-applications.

By choosing "View Results" the program generates a screen displaying the graphical visualization of the single simulated seismograms and response spectra (Figure 6), where the user can scroll the recorded seismograms and response spectra one by one. The accelerograms can be visualized together with velocity, displacement and WOOD-ANDERSON seismograms, as shown in the figure. By choosing a seismogram length different than the one proposed as default, plots can be visualized in their total length or be zoomed.

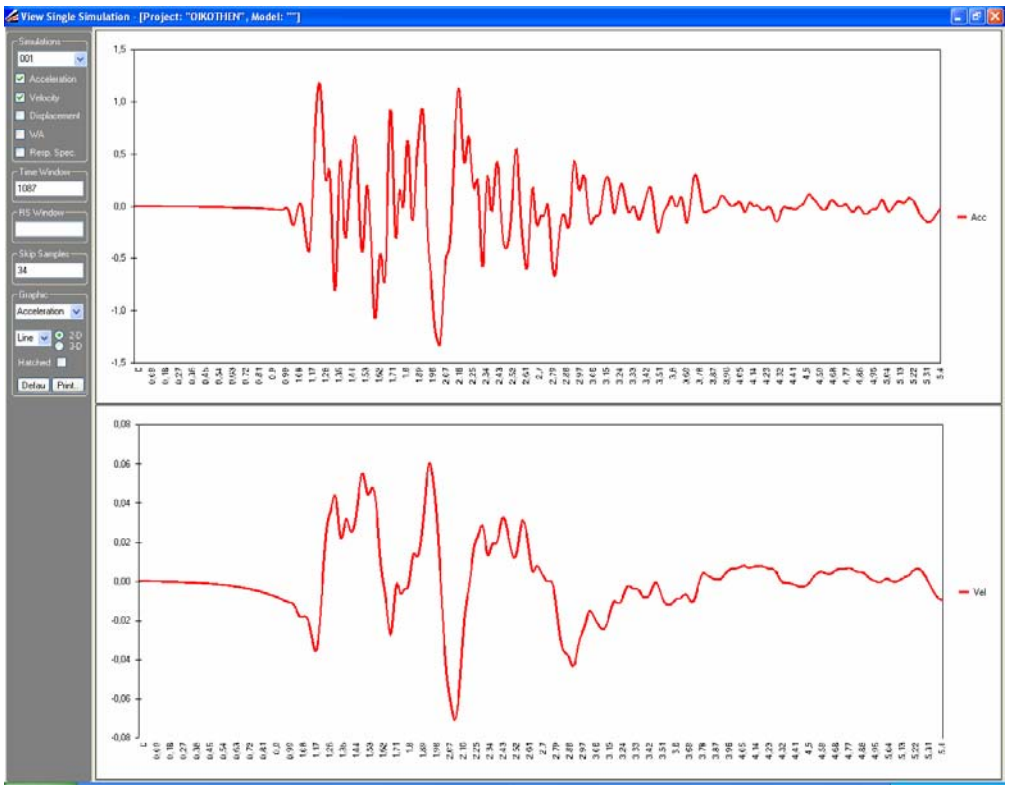


Figure 6. Single Results

The user may decide which signal types are shown by ticking on/off the selection boxes on left-hand side. The graphical interface allows the editing of the figures so that they can be easily imported in other documents, an extremely useful feature for the preparation of publications and reports.

4. SHAKYGROUND USERS

Public Agencies (local, regional, national) interested to territory planning, emergency planning and seismic microzoning

Public Agencies may use SHAKYGROUND to identify critical points in a territory, spotting where damages would be grater if an earthquake should occur. The model can help the identification of areas and sites where severe ground shaking is expected, so that competent public authorities can adjust their territory planning (emergency, territorial, microzoning etc.).

The program is designed to support seismic microzoning project with many earthquakes and sites examples as input. It has been positively used in seismic microzoning projects for the cities of Catania, Ragusa, Sarno, Nocera (Italy).

Civil, mechanical and electrical engineers, geologic and geotechnical consultants

These users may use SHAKYGROUND to obtain numerical data of earthquake ground motion at the level of the foundation of a building.

At the same time they can obtain the prevision of the seismic impact in a specific site where a building is planned to be erected or reconstructed or reinforced. Moreover they can insert the schemes of the synthetic seismogram resulted by the run of the simulations as input for their calculations.

Civil Protection

The necessities of these users are the same of those of the Public Agencies. The critical sites location allows the territorial planning and the emergency planning optimization. It will be possible to identify zones where earthquake damages will be particularly serious.

People living in these areas will have greater need of aid, and at the same time these areas would not have to be chosen for the deployment of first aid structures.

Insurance and re-insurance Companies

Companies that want to propose insurances contracts covering damages to buildings caused by earthquake, can use SHAKYGROUND to classify all sites on the basis of the estimated effects for several earthquake scenarios. The so obtained results of the seismic zoning would help the insurance companies to develop a specific marketing strategy for each of these areas.

CONCLUSIONS

Based on sound and proven methodologies, SHAKYGROUND provides the user with a user friendly interface for tackling the complex problem of the assessment of seismic hazard. The program has already been used in many projects in Italy, Germany, USA, Japan, Corea and has proven to be a reliable and useful software for estimating strong ground motion. More information on SHAKYGROUND can be found at www.shakyground.biz.

REFERENCES

Aki, K., Characterization of barriers on an earthquake fault. *Journal of Geophysical Research, 84*, 4160-4168, 1979.

Boore, D.M., Stochastic simulation of high frequency strong ground motion based on seismological models of the radiated spectra. *Bulletin of the Seismology Society of America, 73*, 1865-1894, 1983.

Brune, J., Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research, 75*, 4997-5009, 1970.

Gresta S., Langer H., Mucciarelli M., Gallipoli M.R., Imposa S., Lettica J., Monaco C. (2004). The site response in the city of Ragusa-Ibla (Sicily) by using microtremors and strong ground motion simulations. In: Risk Analysis IV (C.A. Brebbia ed.), WIT Press, Southampton, pp. 93-101.

Haskell N.A., The dispersion of surface waves in layered media. *Bulletin of the Seismology Society of America, 43,* 17-34, 1953

Langer H., Catalano S., Cristaldi M., De Guidi G., Gresta S., Monaco C., Tortorici L., 1999. Strong ground motion simulation in the urban area of Catania on the basis of a detailed geological survey. Proc. 2nd Int. Sympos. Earthquake Resistant Engineering Structures, Catania, Italy, June 15-17, 1999, pp. 343-352, ISBN 1-85312-689-6

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Thunderstorm Strike Probability Nowcasting, a New Algorithm

S. Dance [a], **B.Ebert**[a] **and D.Scurrah**[a]

[a]*Centre for Australian Weather and Climate Research, GPO Box 1289K, VIC, 3001, Australia (S.Dance@bom.gov.au)*

Abstract: Thunderstorm nowcasting is an important function for the severe weather forecaster. The usual method is for radar to detect thunderstorms, and for downstream systems to generate 'threat areas' for issuing warnings. However, these threat areas are difficult for other systems to work with as they are not mathematically well defined. In this paper we develop an algorithm, THESPA (THunderstorm Environment Strike Probability Algorithm), to translate a radar thunderstorm detection into a 'strike probability', a new concept for thunderstorm nowcasting. This represents the probability a given point will be affected by a thunderstorm in a given period. Being well defined, this can be combined with other strike probabilities or can be utilized by other systems or forecasters to produce a more valuable warning product. The paper includes studies that justify assumptions made about the statistics of thunderstorm motion, and results demonstrating how effective the algorithm is at thunderstorm nowcasting.

Keywords: thunderstorm; strike probability; radar; nowcasting;

1 INTRODUCTION

Nowcasting[1] thunderstorms is an important societal and economic activity. The Australian Bureau of Meteorology has a number of nowcast systems in place to detect thunderstorms (TSs), determine their speed, direction and severity, and make predictions about their location in the future. However, the current systems do not make probabilistic nowcasts, rather they delimit a homogeneous threat area without distinguishing high or low threat gradations. One such system, TIFS (Thunderstorm

[1] A nowcast is a short term weather forecast.

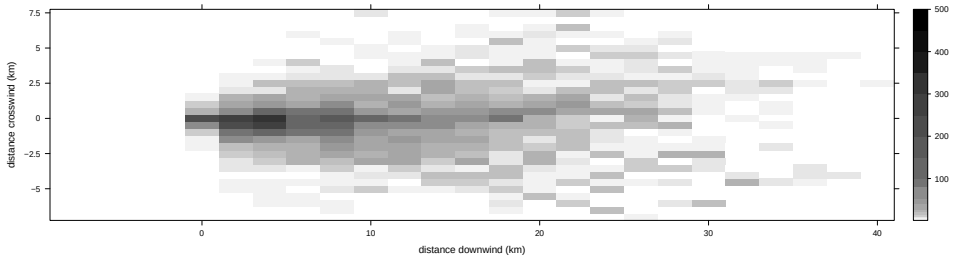


Figure 1: *Histogram of detected TS positions compared with their nowcasts, which have been remapped to a nominal direction and speed of* $90\ degrees$ *and* $50\ kmh^{-1}$*, with detections remapped accordingly. Approximate counts for a given bin are indicated by the scale on the right.*

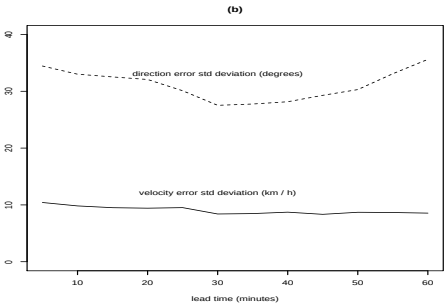


Figure 2(a) *TS nowcast verification statistics by lead time t (minutes) showing standard deviations of errors in V speed ($km\,h^{-1}$) (line) and θ direction ($degrees$) (dotted line)*

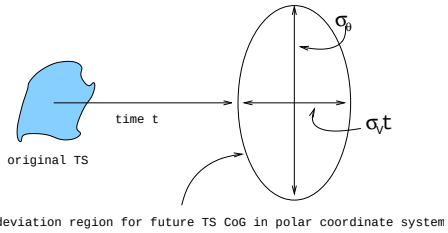


Figure 2(b) *Shows the original TS and the 1 standard deviation ellipse for its CoG after time t given speed V, and uncertainty in velocity and direction of σ_V, σ_θ.*

Interactive Forecast System, Bally [2004]), displays TSs detected by a radar processing system TITAN (Thunderstorm Identification, Tracking, Analysis And Nowcasting, Dixon and Wiener [1993]) in a graphical user interface (GUI) that allows forecasters to manipulate the nowcast threat area and issue warnings. Currently, this system initially generates an ad hoc threat area based on TS velocity and size. We wish to improve on this by producing a mathematically soundly based probabilistic threat area that allows the output to be combined with probabilistic products from other systems, and forecasters to issue more meaningful warnings. In this paper we present an algorithm, THESPA (THunderstorm Environment Strike Probability Algorithm), to generate such probabilistic nowcasts. In this, we borrow the concept of the 'strike probability' from cyclone forecasting, which gives the probability a given place will be affected by the phenomenon over the nominated period.

In Section 2 we talk about the antecedents of the term "strike probability". In Section 3 we describe TS motion studies which result in an empirical strike probability which we wish to capture with THESPA, and an analysis of how the standard deviation of thunderstorm motion measures vary through the lead time (how far in the future the nowcast is predicting). In Section 4 we describe the assumptions the algorithm depends upon. We then describe the reasoning behind the algorithm by dividing the problem into two stages: firstly in Section 5 we derive a probability density function giving the chance a TS centre of gravity (CoG) touches a point at any time over the nowcast period (ie, 30 or 60 minutes). Then in Section 6 we go on to establish the probability distribution that a point is affected by *any part* of the TS at any time in the nowcast period, ie, the strike probability. In Section 7 we show examples of the distribution, together with statistical results.

2 PREVIOUS WORK

The phrase 'strike probability' has been borrowed from tropical cyclone forecasting. Here the two main approaches are based on numerical weather prediction (NWP) which model the cyclones explicitly. One approach to introducing probability is to employ ensembles of numerical models, the weighted percentage of the ensemble that puts the cyclone within a set distance of a given point is the strike probability for that point [Weber, 2003].

Another approach [Templeton and Keenan, 1982] to the strike probability problem is to use a single NWP forecast given as a number of time steps. The strike probability is calculated by interpolating between the steps, assuming the cyclone path is given by a polynomial, the shape of the cyclone is circular, the probability distribution of each step is given by an *a priori* bivariate normal distribution, and there is statistical independence between steps.

Our approach is closer in spirit to the latter, where probabilities are introduced through an explicit

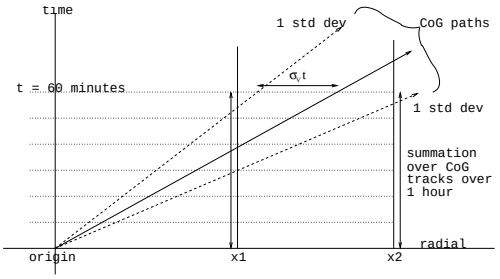


Figure 3(a) *Time-radial slice with two points x1, x2 and indicating the 1 standard deviation subset of CoG paths that intersect those points.*

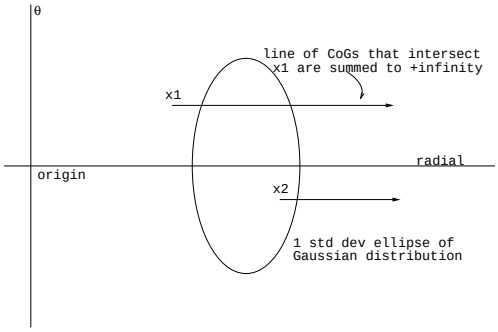


Figure 3(b) *Final (ie, 60 minutes) 2D Gaussian showing lines of projected trajectories of the CoG that intersect the points x1 and x2, in polar coordinates.*

assumption of a bivariate normal distribution, although our approach assumes a linear path, a storm shape given by the initial detection, and complete dependence along potential storm tracks and mutual exclusivity between tracks.

3 THUNDERSTORM MOTION STUDIES

To determine the statistics of the thunderstorm motion a database of 1682 thunderstorm tracks[2] was collected from the Kurnell radar near Sydney, Australia, with 8302 TITAN nowcasts of up to one hour and verified against actual motion [Ebert et al., 2005]. It is likely that in places other than Sydney the results found below may vary. The data consists of a forecaster selected set of TITAN output.

We aggregated all TS nowcasts and subsequent detections into one composite histogram, as an empirical "strike probability" chart, where the velocity of the first member of each storm track was taken as the nowcast. The detected locations of the storms in each track, in storm-relative coordinates (nowcasts were remapped to a nominal $50\ km\,h^{-1}$ and $90\ degrees$, with detections for each timestep along the TS track remapped accordingly), were recorded in a 25×30 spatial grid by incrementing a counter at the detected position of the TS along its track (each gridbox was $2\ km \times 0.5\ km$). This produced the 2D histogram of storm detections shown in Figure 1.

It can be seen that the detected TSs produced a histogram with an elliptical core and triangular envelope. Few TSs last the full hour (which would put the detections on the $50\ km$ mark). In fact most TSs last only a short time or are slow compared with their initial speed, as evident from the large spike in the histogram less than $10\ km$ from the origin. This may be related to the use of TITAN 35 dBZ data, which would include weak storms with a short life.

Another analysis (Figure 2(a)) showed that the standard deviation for speed (V) and direction (θ) errors between nowcast and detection (σ_V and σ_θ), are roughly constant (around $10\ km\,h^{-1}$ and 30 $degrees$ respectively) over all lead times. This is consistent with the findings of Bellon and Austin [1978] where velocity error and direction error is roughly constant over lead time. Thus we can use these standard deviations for predicting TS motion as THESPA requires (see Section 4).

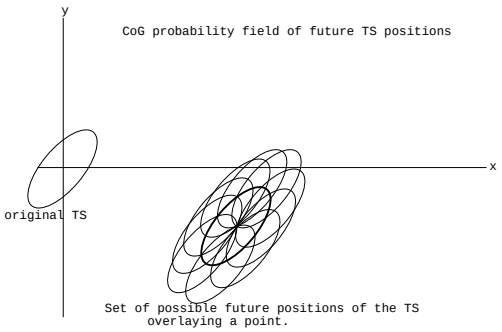


Figure 4(a) *Schematic probability density field of possible positions of the CoG of the TS, with for one particular point, the set of possible future positions of the TS that overlay it (TS overlay set). Note that those shown (light ellipses) are examples of the extremes of the TS overlay set, with the particular point being right on their boundaries, and the collection of their CoGs laying on the dark solid ellipse (the TSA boundary). The TS overlay set also includes all those ellipses whose CoGs lie within the TSA boundary.*

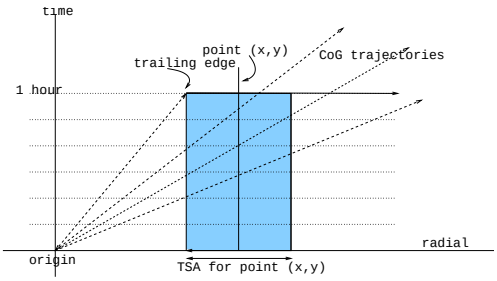


Figure 4(b) *Time-radial slice through the set of potential TS CoG trajectories that intersect a point (x, y) (TSA) at any time during the nowcast period (say 30 or 60 minutes). The probability density associated with the trailing edge position already accounts for the contributions from all trajectories passing through it (ie, all trajectories below that point in the figure), and thus that point is the only point one needs to consider to compute the probability density for that radial through the TSA.*

4 ASSUMPTIONS

In this paper we present a foundation for the calculation of the probability that a TS is over (affects) some point (x, y) at some time in the nowcast period, given an existing TS at time $t = 0$ defined by its shape and size (here given as an ellipse), speed V and direction θ. For mathematical simplicity it is assumed the TS moves at constant velocity, lasts for the full nowcast period, and does not change shape or size. Of course, these are not realistic assumptions, but are necessary for the development of the algorithm. For instance, it is not clear whether the TS will die early, or become stronger, split, or be replaced by a new TS. We also assume that the distribution of TS speed and direction errors (difference between actual speed and direction and nowcast speed and direction) is a 2 dimensional bivariate normal distribution [Brunk, 1965], and that the standard deviations of this distribution, σ_V and σ_θ, are constant over the nowcast period (as justified by the study of TS statistics in Section 3). The forecaster should be made aware of these assumptions when using the system, although the assumption of linearity corresponds closely to current practice.

5 PROBABILITY DENSITY FUNCTION OF POINTS MEETING THE CENTRE OF THE TS

In this section we calculate the probability density function that a given point meets the CoG of the TS at some time in the nowcast period. This is based on the assumption (above) that the future positions of the TS CoG is given by a two dimensional Gaussian distribution [Brunk, 1965], with standard deviations σ_V, σ_θ (Figure 2(b)).

Below, without loss of generality, we assume the velocity is oriented along the x axis. In the implementation these results can be rotated to the actual direction of the thunderstorm. The formula for the above 2D Gaussian distribution of time-dependent joint speed and direction errors in (r, θ) polar space[3] is:

$$f(r,\theta) = (1/2\pi\sigma_V t\sigma_\theta)e^{-(((r-Vt)/\sqrt{2}\sigma_V t)^2+(\theta/\sqrt{2}\sigma_\theta)^2)} \quad (1)$$

where t is the lead time, and which is normalized so that integrating over all r, θ gives a probability

[2] A track records the position of a single identified storm for each time step during its life.
[3] To convert from Cartesian to polar coordinates,
use $r = \sqrt{(x^2 + y^2)}$ and $\theta = atan(x/y)$.

of 1. Note that the dimension of the probability value this function returns is $angle^{-1}distance^{-1}$.

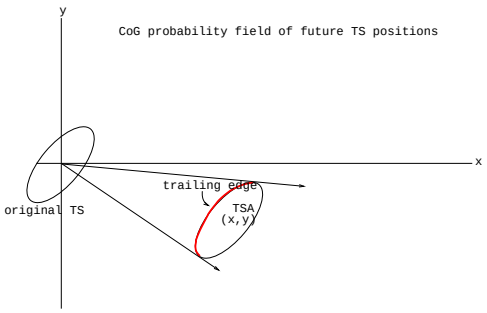


Figure 5: *Shows a plan in Cartesian coordinates of the original TS, the probability field of places that meet the CoG of the TS, the TS area (TSA) for a position (x, y), and the "trailing edge" that contributes all the probability to the sum for that position.*

One problem here is that the 2D Gaussian distribution is not designed for a restricted space like the polar plane. We deal with this by remapping the polar plane so that polar angles $\theta > 90\ degrees$ are converted to $180 - \theta$, $\theta < -90$ are converted to $-180 - \theta$ and polar distance r is converted to $-r$ in both these cases. This allows us to use the same probability density function (PDF) in front of and behind the storm with respect to its velocity, with the caveat that there will be a discontinuity in the PDF around $\pm 90\ degrees$.

To compute the probability that any given point (r, θ) meets the CoG of the thunderstorm during time t, say 60 minutes, consider Figure 3(a), which shows time as the vertical axis, and the radial dimension as the horizontal axis. For a given point $x1$ or $x2$, one has to integrate the probabilities of all the possible locations of the CoG that pass through that point up to time t. These CoG positions map onto a line on the 2D Gaussian shown in Figure 3(b).

The probabilities along these lines of CoG intersections are summed to positive infinity for positive radial distance and negative infinity for negative radial distance to produce the probability per degree of θ that the CoG of the thunderstorm met the point $x1$ or $x2$ at any time in the nowcast period. That is, for $x1$ in Figure 3(b) one sums over most of the Gaussian, and for $x2$, which is further from the origin, only the tail end of the Gaussian distribution. Since each CoG trajectory is mutually exclusive of all others, it is possible to combine the probabilities additively:

$$r > 0 : probCoG(r, \theta) = \int_{r}^{+\infty} (1/2\pi\sigma_V t\sigma_\theta) e^{-(((s-Vt)/\sqrt{2}\sigma_V t)^2 + (\theta/\sqrt{2}\sigma_\theta)^2)} ds \tag{2}$$

After some manipulation,

$$r > 0 : probCoG(r, \theta) = (1/\sqrt{2\pi}\sigma_\theta) e^{-(\theta/\sqrt{2}\sigma_\theta)^2} (1 - erf((r - Vt)/\sqrt{2}\sigma_V t))/2 \tag{3}$$

and

$$\begin{aligned} r < 0 : probCoG(r, \theta) = \int_{-\infty}^{r} (1/\pi\sigma_V t\sigma_\theta) e^{-(((s-Vt)/\sigma_V t)^2 + (\theta/\sigma_\theta)^2)} ds = \\ (1/\sqrt{2\pi}\sigma_\theta) e^{-(\theta/\sqrt{2}\sigma_\theta)^2} (1 + erf((r - Vt)/\sqrt{2}\sigma_V t))/2 \end{aligned} \tag{4}$$

where erf is the **error function** [Fristedt and Gray, 1997], s and ds are radial distance measures, and (r, θ) is a point in polar coordinates. Note that at $r = 0$ the values returned by Equations 3 and 4 are not equal. In fact, we never evaluate them at $r = 0$ as that occurs for points within the detected TS, where of course the probability that the point is affected by the TS is 1.

6 PROBABILITY DISTRIBUTION FOR POINTS AFFECTED BY THE TS

In the last section we found a probability density function for points meeting the CoG of the TS in the nowcast period. In this section, we derive a probability distribution that points will be *affected*

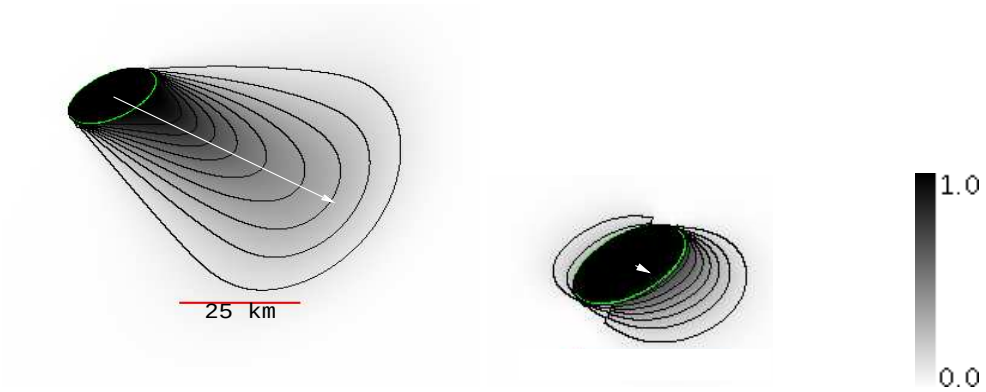


Figure 6: *Examples of output from algorithm with typical values derived from TITAN storm tracking. On the left,* $\theta = 115\ degrees$, $\sigma_\theta = 30\ degrees$, $V = 54\ km\,h^{-1}$, $\sigma_V = 10\ km\,h^{-1}$, *on the right,* $\theta = 115\ degrees$, $\sigma_\theta = 30\ degrees$, $V = 4\ km\,h^{-1}$, $\sigma_V = 10\ km\,h^{-1}$, *with probability contours shown at 0.1 intervals. The original storms are shown as ellipses with nowcast velocity as an arrow. A scale marker shows* $25\ km$. *Black shading represents a strike probability of 1, white shading represents a strike probability of zero, as indicated by the greyscale bar on the right.*

by any part of the TS over the nowcast period. It is assumed that the size and shape of the TS does not change through this period.

To compute the required probability given a specific shaped TS, we need for each point (x, y) to combine the probabilities of all possible future positions of the TS that overlay that point (dubbed TS overlay set, and whose set of CoGs is dubbed TS area or TSA), see Figure 4(a). Thus for each such TS position overlaying (x, y), we find the probability at its CoG, and for the whole TSA, need to combine such probabilities for all ellipses within the TS overlay set. This requires an operation similar to convolution in image processing, in which the possible TS positions are "raster scanned" over the point.

However, we cannot simply add the probabilities, as they are correlated. That is, some potential positions of the CoG entail other positions, namely those of TS trajectories that lie in the same direction but are faster (which have already passed over the same point earlier in the nowcast period), and which now (at the end of the nowcast period) lie further away from the origin, see Figure 4(b). These "downstream" points should not be included in the probability sum. Since the probability density of its CoG position is in units of $angle^{-1}$, we can find the total probability that the point (x, y) was affected by the TS ($strikeProb(x, y)$) by integrating along the "trailing edge" of the TSA (points of the TSA closest to the origin, Figure 5) for that point. Let TE_{xy} be a function from polar angle to radial distance defining the shape of the trailing edge for the given TSA, ie, $TE_{xy}(\theta) \rightarrow r$, with domain θ_1 to θ_2 being the angular limits of the trailing edge, then replacing r in $probCoG(r, \theta)$ (see Equations 3 and 4) with $TE_{xy}(\theta)$ gives:

$$strikeProb(x, y) = \int_{\theta_1}^{\theta_2} probCoG(TE_{xy}(\theta), \theta) d\theta \tag{5}$$

Note that this equation is far cheaper to evaluate than a full convolution due to the much smaller number of points along the trailing edge compared with the full area of the TSA. One then repeats this operation for each point in the geographic region to generate the complete probability field. Note, this is not a probability **density** function, each point now has an associated **probability**. To deal with the case of multiple TSs in the same region, the strike probability distribution can be calculated for each TS separately, and then pointwise combined with the usual rules of probability (assuming independence).

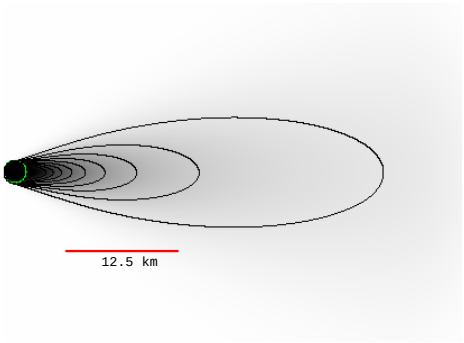


Figure 7(a) *Example of algorithm output with values chosen to emulate data in the histogram in Figure 1 ($\theta = 90$ $degrees$, $\sigma_\theta = 30$ $degrees$, $V = 50$ $km\,h^{-1}$, $\sigma_V = 10$ $km\,h^{-1}$), with probability contours shown at 0.1 intervals, and a scale marker showing 12.5 km. Black shading represents a strike probability of 1, white shading represents a strike probability of zero, as in Figure 6.*

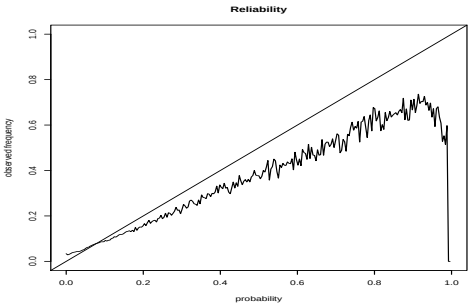


Figure 7(b) *Reliability chart for thunderstorm strike probability versus observations: for each probability value the observed frequency of thunderstorms at pixels with that nowcast probability is plotted, compared with the ideal curve, the straight diagonal line.*

7 RESULTS

The THESPA algorithm described above was coded in Java. The results for some typical thunderstorm parameters are shown in Figure 6. The original TS shape is shown as an ellipse, whose interior of course has a probability of 1 (black) that the point is affected by the TS. For high speed relative to speed standard deviation (σ_V), the probability field shows a smoothed "triangular" envelope centered around the velocity vector (left side of Figure 6). For slower moving TSs there is a significant chance the TS will move opposite to its initial velocity vector, which shows up as finite probability on the back side of the TS (with respect to velocity). This is clear in the contoured output shown to the right in Figure 6. One can also see in this instance, around the 'ends' of the ellipse (at 90 $degrees$ from the velocity vector), the discontinuity due to the remapping of the polar coordinate system described in Section 5.

To compare the output of the algorithm with actual storm behaviour, we ran our algorithm using parameters chosen to match the histogram shown in Figure 1, producing the probability field shown in Figure 7(a). It can be seen that there is indeed some similarity between the two figures, suggesting that our algorithm is generating plausible strike probabilities.

To further compare THESPA with actual storm behaviour, we processed the TITAN database (see Section 3) so that for each storm track, we generated a strike probability nowcast from the first detection that included a nowcast velocity (this is the second detection, as TITAN requires 2 detections to compute velocity), provided the storm was strong enough (ie, VIL, or vertically integrated liquid, above 30 $kg\ m^{-2}$) to last a reasonable time. For each increment in predicted strike probability we counted the number of observed TS occurrences and non-occurrences to calculate an observed frequency. We ignored locations where there were no nowcasts. The resultant reliability chart (in which the frequency of a phenomenon is plotted against its nowcast probability [Jolliffe and Stephenson, 2003]) is shown in Figure 7(b). Overall, higher nowcast strike probabilities are associated with higher observed frequency of TS detection. However, at very high probabilities (over 0.9) the low number of cases gives erratic results. In general, THESPA has a tendency to overestimate the probabilities. This is likely to be related to the assumption that storms last for the full nowcast period and do not experience growth or decay.

8 CONCLUSION

The THESPA algorithm presented here has been designed for mathematical tractability and algorithmic simplicity. However, algorithmic simplicity cannot capture the variety of real phenomena. For instance, If a storm track lasts less than the nowcast period, it is either a brief storm, it split or merged, or TITAN lost track of it, whereas our algorithm assumes each storm lasts, and is tracked, for the full nowcast period. Secondly, the observations take place at most every 5 minutes, whereas our algorithm assumes continuous observation. Despite these caveats, THESPA performed reasonably well with good reliability.

Future developments include dealing with storm lifetimes of less than the nowcast period, and performing the calculations within Cartesian coordinates, thus removing the discontinuities resulting from the use of polar coordinates. In this paper we have adapted the concept of strike probability to thunderstorms, and described an algorithm for consistently and defensibly generating them. The algorithm uses a number of simplifying assumptions to allow it to be run in an interactive package (TIFS) in real time. Our claim is that the resultant distribution is nonetheless useful, especially in light of the results given in Section 7.

ACKNOWLEDGEMENTS

The authors wish to thank the anonymous reviewers of this paper who made many useful suggestions.

REFERENCES

Bally, J. The thunderstorm interactive forecast system: Turning automated thunderstorm tracks into severe weather warnings. *Weather Forecasting*, 19:64–72, 2004.

Bellon, A. and G. L. Austin. The evaluation of two years of real-time operation of a short-term precipitation forecasting procedure (SHARP). *Journal of Applied Meteorology*, 17(12):1778–1787, 1978.

Brunk, H. D. *An Introduction to Mathematical Statistics*. Blaisdell Publishing Company, 1965.

Dixon, M. and G. Wiener. TITAN: thunderstorm identification, tracking, analysis and nowcasting - a radar based methodology. *J. Atmos. Oceanic Tech.*, 10(6):785–797, 1993.

Ebert, E., T. Keenan, J. Bally, and S. Dance. Verification of operational thunderstorm nowcasts. In *World Weather Research Program Symposium on Nowcasting and Very Short Range Forecasting*, page 8.12, Toulouse, France, 2005.

Fristedt, B. and L. Gray. *A Modern Approach to Probability Theory*. Birkhaeuser Boston, 1997.

Jolliffe, I. T. and D. B. Stephenson. *Forecast Verification*. John Wiley and Sons, 2003.

Templeton, J. I. and T. D. Keenan. Tropical cyclone strike probability forecasting in the Australian region. Technical Report 49, Bureau of Meteorology, GPO Box 1289K, Melbourne, VIC, 3001, Australia, 1982.

Weber, H. C. Hurricane track prediction using a statistical ensemble of numerical models. *Monthly Weather Review*, 131(5):749–770, 2003.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Environmental Informatics, Environmetrics, Chemoinformatics, Chemometrics: Integration or Separation!?

Kristina Voigt
Helmholtz Zentrum Muenchen, German Research Center for Environmental Health (GmbH), Ingolstaedter Landstr. 1, 85764 Neuherberg,Germany,
e-mail: kvoigt@helmholtz-muenchen.de

Abstract: Several interdisciplinary scientific disciplines have been established in the past 30 years focussing on the computer science, mathematical and statistical aspects of environmental protection issues. An important society in this respect is the International Environmental Modelling and Software Society iEMSs who is organizing this conference. However, iEMSs will not be treated in this paper. Other disciplines concerning the evaluation of environmental and chemical data are: Environmental Informatics, Environmetrics as well as Chemoinformatics and Chemometrics. The aim of the research of these societies is to get more information out of the data and to arrive at some kind of decision support to improve the status of the environment. The introduced disciplines have all a different history, focus, grade of internationalisation and background of members. This paper aims at informing the members of the societies and interested working groups in the scientific vicinity as well as those generally interested in environmentally-related scientific disciplines.

Keywords*: Environmental Informatics; Environmetrics; Chemometrics; Chemoinformatics;* ranking methods*, environmental databases, interdisciplinary environmental reserach*

1. INTRODUCTION

The tremendous increase of environmental and chemical data, results in an increasing need for a new generation of software tools that assist the extraction of information from data and transforming the results into knowledge. In front of this background scientific disciplines evolved world-wide which are of great importance to solving the huge environmental data proliferation problems.
The topics of Environmental Informatics, Environmetrics as well as Chemoinformatics and Chemometrics have been initiated in order to support the evaluation of these huge amounts of data.
Around 40 years ago, both environmental science and computers were in the starting blocks of a world-changing evolution. In the status nascendi, these two fields seemed to be completely independent. Computer science with its technical and technological origin was and is opposed to the diverse scientific, ecological, social, and political environmental movements in many respects. Nevertheless, the increasing performance of computers and the need for solutions that respond to environmental stress let these different areas amalgamate. In the early till mid-1980s time was ripe for merging computer science with environmental science (Pillmann, 2006, 2007). The scientific discipline of Environmental Informatics was born. Concerning the data-analysis of environmental and chemical data the scientific field of environmental statistics also named Environmetrics was established in the late 1980s (El-Shaarawi, 2002). With respect to chemistry it was realized some decades ago

that the amount of information accumulated by chemists, in the long run, be made accessible to the scientific community only in electronic form, in other words, it had to be stored in databases. This new field, which deals with the storage, the manipulation, and the processing of chemical information, was emerging (Gasteiger, 2003). This is now one important field in Chemoinformatics. Since then, the application of computer science to the solution of chemical problems has ventured into many more areas of chemistry also into the field of environmental chemistry and ecological chemistry. The International Chemometrics Society was founded in 1974 and was rapidly developing in the 1980s (Einax, 2004).

Several different disciplines were emerging some twenty to thirty years ago which recognized the urgent need to combine aspects of computer science, statistics and mathematics with environmental sciences in general and environmental chemistry in particular. Although the need for these interdisciplinary and applied disciplines was recognized, collaborations between these groups are still in the initiating phase.

This article should support and trigger further understanding and hence cooperation among the disciplines Environmental Informatics, Environmetrics, Chemoinformatics and Chemometrics and of course trigger the interplay with the iEMSs.

2. ENVIRONMENTAL AND CHEMICAL DISCIPLINES WITH RESPECT TO COMPUTER SCIENCE AND STATISTICS

2.1 Introduction

As indicated above, the International Environmental Modelling and Software Society iEMSs who is organizing this conference is not treated explicitly in this paper but those societies and groups which are working in similar areas are discussed.
In the following section the environmental and chemical disciplines are introduced with their societies, WebPages, background literature as well as some journals in which the fields are covered. It goes without saying that only examples are given concerning the background literature and journals.
Table 1 gives an overview on the disciplines and societies.

Table 1. Overview on Disciplines, Societies

Name of Discipline	Society	URL	Background material Examples
Environmental Informatics	TC Informatics for Environmental Protection	http://www.iai.fzk.de/Fachgruppe/GI/welcome.eng.htm l	Proceeding volumes of annual conferences, Umweltinformatik (Page 1995), Environmental Informatics (Avouris, 1995)
Environmetrics	TIES	http://www.environmetrics.org/	Encyclopedia of Environmetrics (El-Shaarawi, 2002)
Chemoinformatics	No society defined yet	http://www.gdch.de/vas/tagungen/tg/5316__e.htm	Handbook of Chemoinformatics (Gasteiger, 2003)
Chemometrics	International Chemometrics Society?	http://www.chemie.uni-jena.de/institute/ac/einax/ak_chemo/ak_chemo	Chemometrics in Environmental Analysis (Einax, 2004)

2.2 Environmental Informatics

Environmental problems in the 1980s resulted in an increased ecological awareness worldwide, which considerably enhanced political and scientific activities for the protection

of the environment. In the science sector, a rapidly growing community conceived new computer applications for decision making and information exchange in the field of environmental protection. Activities in the field of Computer Science applied to environmental research and protection are denoted by the term "Environmental Informatics" (EI). This is the global definition of EI. The EI community continually addresses new and challenging topics in the interface between the fields of Informatics (Computer Science) and Environmental Sciences. The nucleus of the upcoming field of "computer application for environmental protection" was the 1st Symposium held at the Karlsruhe Research Center in 1986 and the foundation of the Technical Committee "Informatik im Umweltschutz" (Environmental Informatics). The new extended title of the Technical Committee is now Environmental Informatics, Informatics for Environmental Protection, Sustainability and Risk Management. More than 22 years later, 22 EnviroInfo conferences have been held with approximately 3.000 papers by 7.500 authors published in proceedings' volumes since the first event. From a German initiative the conference has largely expanded into the European field.
The history and the development are outlined in Figure 1.

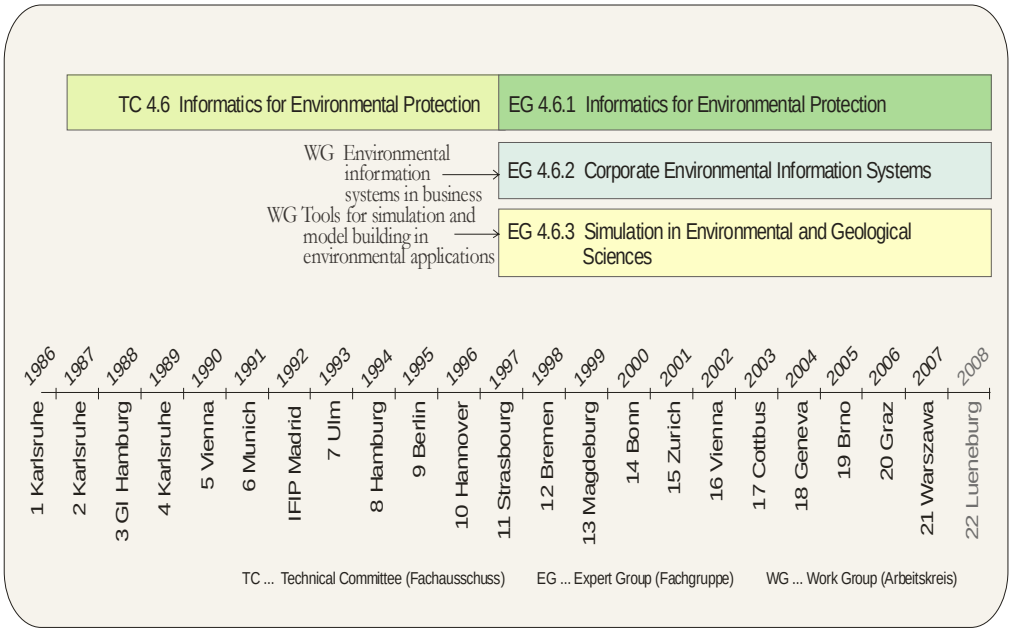


Figure 1: Development and Places of EnviroInfo Conferences 1986-2008 (Pillmann, 2007)

The scope of the Technical Committee in particular consists in the dissemination and application of modern methods and techniques of computer science in the environmental sector and to lend expert assistance in the development of effective, application specific informatics tools. The collaboration of partners from universities, research institutions, agencies and industries etc. is performed. Apart from the organization of conferences and proceedings' volumes the Committee publishes a newsletter twice a year. In this publication news on the current work in the society, upcoming events like conferences, workshops are announced, conference reports are given. In the section scientific contributions new developments are explained. The newsletter which is written in German and partly in English is available on the Web page of the society (see Table 1). An important aspect of the work of Environmental Informatics is the establishment of special interest groups and working groups which take a closer and more in depth look at specific issues, e.g. Environmental Information Systems in Business, Simulation in Environmental and Geological Sciences, Environmental Databases, etc. These working groups meet regularly, publish their results in proceeding volumes, and report to the society.
Apart from these basic activities, several textbooks have been written by members of the EI community over the past two decades. One of the first basic textbooks was entitled "Umweltinformatik" edited by Bernd Page and Lorenz Hilty in 1994. A second edition was published soon after the first one (Page and Hilty, 1995). An international textbook edited by Nicholas Avouris and Bernd Page followed soon after (Avouris and Page, 1995). Two

important books concerning one of the basic topics of EI, namely environmental databases and environmental information systems were published by Guenther (1998) and Rautenstrauch and Patig (2001), respectively. A recent attempt to extend the focus from environment to sustainable development in general was made by Hilty (2004). An approach to quantify the impact of information and communication technologies on sustainability was developed in a recent project (Hilty, 2006a). This is only a small sample of available background material concerning EI.
The Environmental Informatics initiative does not have a scientific journal of its own. However, a Special Issue entitled "Environmental Informatics" of the scientific journal Environmental Modelling and Software was published recently (Hilty, 2006b). The history, current research and trends in environmental informatics are given in this issue.

2.3 Environmetrics

TIES (The International Environmetrics Society) is a non-profit organization aimed to foster the development and use of statistical and other quantitative methods in the environmental sciences, environmental engineering and environmental monitoring and protection. To this end, the Society promotes the participation of statisticians, mathematicians, scientists and engineers in the solution of environmental problems and emphasizes the need for collaboration and for clear communication between individuals from different disciplines and between researchers and practitioners. The Society further promotes these objectives by conducting meetings and producing publications, and by encouraging a broad membership of statisticians, mathematicians, engineers, scientists and others interested in furthering the role of statistical and mathematical techniques in service to the environment. The first TIES Conference took place in Cairo, Egypt in 1989. TIES is an international society and its conferences are being held at several continents of the world. The 2007 conference was in Mikulov, Czech Republic, while the TIES 2008 conference will be held in Kelowna, BC, Canada.

Abdel El-Shaarawi the founder of the TIES conferences is optimistic about the future of the field of Environmetrics as quoted in an interview article (Esterby, 2003). Environmetrics has been around for a reasonably long time and has been quite successful expanding in terms of its methods and their range of application. Prof El-Shaarawi is of the opinion that without any doubt, statistics is the broadest field of science in terms of the use of its methods. The same formula for advancement as used by the great statisticians in the past, if used by the current generation, will guarantee a bright future for the field. This formula requires that statistical methods should be developed to address real problems in science and technology, and statisticians should be taught how to effectively communicate with scientific disciplines.
Currently there are ongoing activities that TIES will become a section of the International Statistical Institute (ISI).
An important background work of TIES is the Encyclopedia of Environmetrics, edited by Abded El-Shaarawi and Walter Piegorsch (El-Shaarawi, 2002). The publication of this major reference work, which includes more than 500 detailed articles, is a landmark for environmetrics. It really defines the subject and indicates its future directions. As can be seen, there is no shortage of real problems for scientists to work on, where their investigations have direct impact on improving the quality of life for the current and future generations.
The official journal of the International Environmetrics Society is named Environmetrics, published by Wiley-Interscience. The Editor-in-chief is Abdel El-Shaarawi (2008). In this journal selected papers from the TIES conferences are published.
TIES also publishes a newsletter which is available via the Webpage (see Table 1).

2.4 Chemoinformatics

Chemoinformatics is a fairly new name for a discipline that has been around for quite a while. Different people sometimes give rather different definitions of chemoinformatics.

The rather broad and general definition by Johann Gasteiger the editor of the Handbook of Chemoinformatics (Gasteiger, 2003) is: Chemoinformatics is the application of informatics methods to solve chemical problems. Chemoinformatics is the use of computer and informational techniques, applied to a range of problems in the field of chemistry. These in silico techniques are often used in pharmaceutical companies in the process of drug discovery.
The field of Chemoinformatics was not founded, nor was it formally installed. It slowly evolved from several, often quite humble, beginnings. Scientists in various fields of chemistry struggled to develop computer methods in order to manage the enormous amount of information and to find relationships between the structures and properties of a compound. In the 1960s some early developments became evident which led to a flurry of activities in the 1970s (Gasteiger, 2003). These days a big issue in this respect is the topic Chemoinformatics and Drug Development.
In Germany conferences on Chemoinformatics have been initiated and held by the CIC (Chemistry-Information-Computer - A division of the German Chemical Society). These conferences have international speakers as well as an international audience http://www.gdch.de/strukturen/fg/cic__e.htm. Last year's conference was the 3. German Conference on Chemoinformatics / 21. CIC-Workshop.
Important aspects of Chemoinformatics are published in the Handbook as well as in the Textbook of Chemoinforamtics (Gasteiger, 2003). The field of chemometrics has seen considerable growth in 2006. This growth is evidenced by the fact that significant advances are no longer found in the pages of a few specialty publications but across a wide range of mainstream chemistry and general science journals, even in the Journal of the American Chemical Society JACS (Agrafiotis, 2007). So, no specific journal on chemoinformatics exists. However, in a variety of chemistry journals the aspects of chemoinformatics are published, e.g. Journal of Chemical Information and Modelling published by ACS (American Chemical Society).

2.5 Chemometrics

Chemometrics is the application of mathematical or statistical methods to chemical data. The International Chemometrics Society (ICS) offers the following definition: Chemometrics is the chemical discipline that uses mathematics, statistics, and formal logic:

- To design or select optimum experimental and measurement procedures
- To provide maximum relevant chemical information by analyzing chemical data
- To obtain knowledge about chemical systems.

Chemometrics, however, is a relatively young science that was born at the end of the 1960s and became an independent working field in the 1970s. The name "chemometrics" was first coined by the Swedish scientist Svante Wold in the early 1970s. His cooperation with the American analytical chemist Bruce R. Kowalski, who at the time was working on pattern recognition methods in chemistry, resulted in the foundation of the International Chemometrics Society in 1997. The first German working party on chemometrics was founded in 1984. Around that time chemometric working groups were also established in many other countries (Einax, 2004). The International Chemometrics Society has no active WebPage at the moment. In Germany the working group "Chemometrics and Laboratory Data Processing" within the division of Analytical Chemistry in the German Chemical Society (GDCh) emerged from the former working groups of Chemometrics and Laboratory Automation, which united in 1994. This group focuses on the combination of data acquisition in the laboratory, their statistical evaluation, plausibility checks and the appropriate quality assurance as good analysis practice (http://www.chemie.uni-jena.de/institute/ac/einax/ak_chemo/ak_chemo).
Current developments in chemometrics are characterized more by applications than by fundamental work in the statistical and/or mathematical field. In addition to classical problems of analytical chemistry, like the optimization methods, interdisciplinary works are gaining more and more interest. The investigation of the environment (Einax, 1997), of food and pharmaceuticals (Voigt, 2004, 2007), and chemometric modelling of technological processes and process analytical chemistry should be mentioned. A further important topic is the application of chemometrics to analytical quality assurance.

Concerning environmentally-directed chemometrics the reference work Chemometrics in Environmental Sciences (Einax, 1997) should be mentioned. A special issue of the journal Analytical and Bioanalytical Chemistry was recently issued (Einax, 2004) which describes the state of the art as well as current trends in chemometrics.
Important journals in this field are: Chemometrics and Intelligent Laboratory Systems published by Elsevier, Journal of Chemometrics published by Wiley-Interscience, and Journal of Computational Chemistry published by Wiley-Interscience.

3. INTERACTION AND COOPERATION BETWEEN THE APPLIED ENVIRONMENTAL DISCIPLINES

All the discussed disciplines, as well as the organization hosting this conference, the International Environmental Modelling and Software Society iEMSs, are organized in international societies and working groups who work in theoretical fields and aim at making contributions to the better understanding of environmental problems and finally aiming at improving the environmental conditions of our planet. All groups are also aware and prepared to do some interdisciplinary work.
Although there are a variety of overlapping activities, limited knowledge about the work of other societies and hence little cooperation between the societies has yet been performed.
Very few attempts have been made so far to initiate common sessions at conferences or common workshops. At the EnviroInfo 2003, 17th International Conference Informatics for Environmental Protection EnviroInfo in Cottbus, Germany, a workshop on chemometrics was organized (Gnauck, 2003)). At the EnviroInfo in Vienna, Austria a TIES Session took place (Pillmann, 2002). These were single events and did not have the impact on the two societies and groups of experts as expected. So unfortunately these kind of cooperative activities have been discontinued so far. Of course quite a few experts attend several of the mentioned interdisciplinary conferences and present their results.
A new attempt is made at the iEMSS in Barcelona while organizing a session on Environmental Informatics, History and Current Trends by the author of this paper.
Figure 2 should demonstrate the current status and the hopefully future trend in the co operations between the named groups and societies.

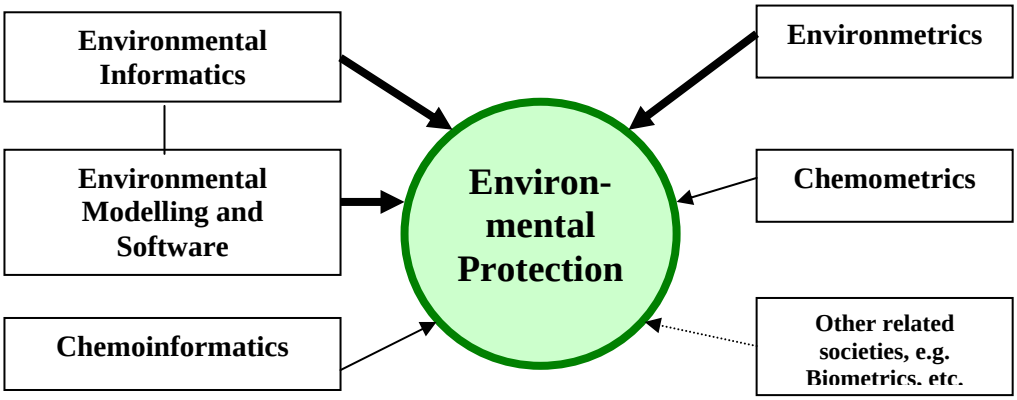


Figure 2: Societies which Focus on Informatics, Mathematical and Statistical Aspects of Environmental Protection Issues.

Currently the societies have more or less one direction: the improvement and protection of our environment by applying their theoretical knowledge. For the chemical societies this is only the case in those research areas where environmental topics, like environmental chemicals are treated. The aim could be reached more effectively if the cooperation between the groups were intensified.

It is envisaged that the societies and experts continue to expand their efforts in collaboration and taking advantages of other theoretical aspects and ideas concerning environmental protection.

4. APPLICATION EXAMPLE OF APPLYING ALL DISCIPLINES

This current research approach should demonstrate an example where environmental informatics, environmetrics on one side are combined with chemometrics and chemoinformatics on the other side. Furthermore, this example can also be regarded as an environmental modelling approach.
In this approach we evaluate 17 Internet databases (objects) by 16 pharmaceuticals (attributes/criteria). The method used for ranking databases is the Hasse Diagram Technique (HDT) which originates in discrete mathematics. It is a multi-criteria evaluation method which can well be used as a tool to rank objects, in this case databases, and is hence also applicable for decision making. The applied methodology reveals the best (maximal objects) and the worst (minimal objects) databases and conflicts among them, due to different information content.

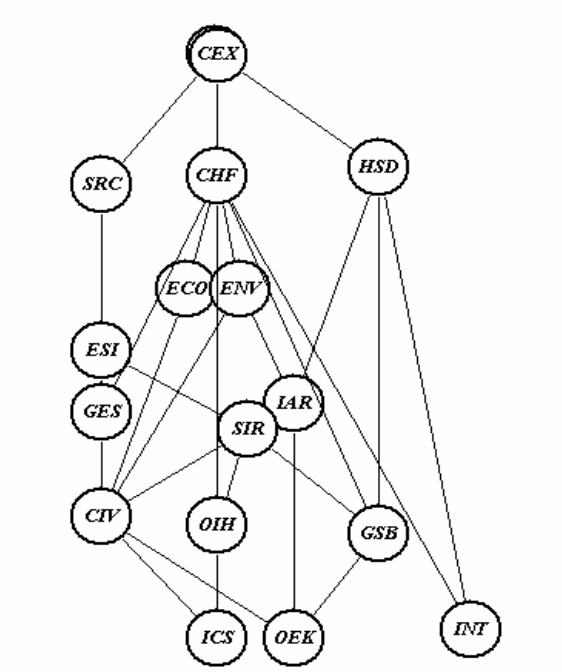


Figure 3: Hasse Diagram of 17 Databases x 16 Pharmaceuticals (equivalent object: {CEX;RXL})

In Figure 3 we can explicitly see 16 out of the 17 discussed databases. The database CEX is equivalent to the database RXL. There exists only one maximal object, the equivalence class {CEX;RXL} which means that these two databases are better with respect to their availability on data of the 16 selected pharmaceuticals than all the other 15 databases. The minimal objects are ICS, OEK and INT. They are the worst databases in this evaluation approach (Voigt, 2007).

5. OUTLOOK

Interdisciplinary work, like given in the example which combines environmental informatics (databases) with environmetrics (mathematical method: Hasse Diagram Technique) and also lies in the chemometrics and chemoinformatics area (pharmaceuticals) helps to support the collaboration between the different societies named in this paper. Additionally the author is of the strong opinion that differences in the approaches and

scientific backgrounds should be used and cultivated in order to get a better understanding of the situation of our environment and hence being able to improve environmental conditions.

REFERENCES

Agrafiotis, D.K., D. Bandyopadhyay, J.K. Wegner, H. van Vlijmen, Recent Advances in Chemoinformatics, *J. Chem. Inf. Model.* 2007 (47), 1279-1293, 2007.

Avouris, N.M., and B. Page, Environmental Informatics, Methodology and Applications of Environmental Information Processing, Kluwer Academic Publishers, Dodrecht, 1995.

Einax, J., Chemometrics in Analytical Chemistry, *Analytical & Bioanalytical Chemistr*y, 380 (3), 368-369, 2004.

El-Shaarawi, A., and W. Pigorsch, Encyclopedia of Environmetrics, Wiley, Weinheim, 2002.

El-Shaarawi, A., (Editor in Chief), Environmetrics, the official journal of the International Environmetrics Society (TIES), Wiley-Interscience, 2008.

Esterby, S., A Conversation with Abdel H. El-Shaarawi, *Statistical Science*, 18 (4), 481-488, 2003.

Gasteiger, J., and T. Engle, Chemoinformatics, A Textbook and Handbook, Wiley-VCH, Weinheim, p. 4, 2003.

Gnauck, A., and R. Heinrich, EnviroInfo 2003, The Information Society and the Enlargement of the European Union, 17th International Conference Informatics for Environmental Protection, Part 2: Applications, Workshops, Posters, Metropolis Verlag, Marburg, 2003.

Guenther, O., Environmental Information Systems. Springer-Verlag, Berlin, 1998.

Hilty, L.M., Seifert, E., Treibert, R. (Eds.), Information Systems for Sustainable Development. Idea Group Publishing, Hershey (PA), 2004.

Hilty, L.M., P. Arnfalk, L. Erdmann, J. Goodman, M. Lehmann, P. Waeger., The relevance of information and communication technologies for environmental sustainability: a prospective simulation study, *Environmental Modelling & Software* 21 (11), 1618-1629, 2006a.

Hilty, L.M., B. Page, J. Hrebicek (Guest Editors), Special Issue Environmental Informatics, *Environmental Modelling & Software*, 21 (11) 1517-1629, 2006b.

Page, B., and L.M. Hilty, Umweltinformatik, Informatikmethoden fuer Umweltschutz und Umweltforschung, 2. Auflage, R. Oldenbourg Verlag, Muenchen, 1995.

Pillmann, W., and K. Tochtermann, EnviroInfo 2002, 16th International Conference, Informatics for Environmental Protection, Environmental Communication in the Information Society, International Society for Environmental Protection, Vienna, 2002.

Pillmann, W., W. Geiger, K. Voigt, Survey of Environmental Informatics in Europe, *Environmental Modelling & Software*, 21 (11), 1519-1527, 2006.

Pillmann, W., W. Geiger, R. Isenmann, Informatics for Environmental Protection, Sustainable Development, and Risk Management, in: Geiger, Schreiber (Eds.), Umweltinformatik, Informatik für Umweltschutz, Nachhaltige Entwicklung und Risikomanagement, Rundbrief des FA 4.6 Informatik im Umweltschutz, Nr. 40, January, pp. 7-11, 2007.

Rautenstrauch, C., and S. Patig, Environmental Information Systems in Industry and Public Administration. Idea Group Publishing, London, 2001.

Voigt K., R. Brüggemann R., S. Pudenz, Chemical Databases Evaluated by Order Theoretical Tools, *Analytical and Bioanalytical Chemistr*y, 380, 467-474, 2004.

Voigt, K., Brüggemann, R., Pudenz, S., A multi-criteria evaluation of environmental databases using the Hasse Diagram Technique (ProRank) software, *Environmental Modelling & Software*, 21 (11), 1587-1597, 2006.

Voigt, K., and R. Brüggemann, Data-Availability of Pharmaceuticals detected in water: An evaluation study by order theory (METEOR), in: Hryniewicz Olgierd, Studzinski Jan, Szediw Anna, EnviroInfo Warsaw 2007, pp. 201-208, Shaker Verlag, Aachen, 2007.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Informatics and the Life Cycle of Products

Lorenz M. Hilty, Roland Hischier, Thomas F. Ruddy, Claudia Som

Technology and Society Lab, Empa – Swiss Federal Laboratories for Materials Testing and Research, St. Gallen, Switzerland (lorenz.hilty@empa.ch)

Abstract: Informatics can make a relevant contribution to sustainable development, if the effects of ICT applications are systematically assessed from a life-cycle perspective and the results of life-cycle assessment (LCA) studies are taken into account by decision makers. The basic scheme of a product life cycle includes the three phases *production*, *use* and *end of life*. In the production phase, raw materials are transformed into the product. In the use phase, the product delivers the service it has been intended for. After the service life of the product ends, parts of the product may be reused or recycled. The rest leaves the system for final disposal or to be recycled in other product systems. Only if life-cycle thinking is applied both to ICT products and to products influenced by ICT applications, is it possible to decide whether a potential ICT application will have a positive or negative environmental impact on the bottom line. With life-cycle thinking, it will be possible to make substantial steps toward sustainable development. Informatics, and in particular environmental informatics as a specialized sub-discipline of it, can contribute to life-cycle thinking by supporting the modelling and data collection process in LCA studies. In addition, dynamic simulation models are useful in prospective technology assessment where LCA methodology reaches its limits.

Keywords: Life Cycle Assessment (LCA); ICT and sustainability; ICT and energy efficiency; life cycle inventory database; life cycle modelling; environmental informatics.

1. INTRODUCTION

The discipline of Informatics (the analysis and design of information processing systems) and its technology (Information and Communication Technology, ICT), have various relationships to the field of Life Cycle Assessment (LCA), i.e. cradle-to-grave accounting of the energy and material flows into and from the environment:

1. Each ICT hardware product has a life cycle which can be assessed with standard LCA methodology and tools. *Example:* The life-cycle of a PC can be modelled using an LCA tool (such as Umberto [ifu, 2008]) and a life cycle inventory database as a data source (such as ecoinvent 2.0 [ecoinvent Centre, 2008]).
2. Each application of an ICT product has various impacts on the life cycle of other products. *Example:* MP3 players have an impact on the demand for audio CDs, for CD players, for CD racks, on headphones, on the use of PCs for Internet access (music downloads), etc.
3. ICT applications support life-cycle thinking and thus contribute to sustainable development. *Examples:* Software tools for life-cycle modelling as well as life-cycle inventory databases help to understand product life cycles and bring the effort needed to do an LCA study down. Such tools can be viewed as applications of environmental informatics [Hilty et al., 2006c].

The following sections give a conceptual overview of each of the three aspects, including examples of existing studies and tools.

2. THE LIFE CYCLE OF ICT HARDWARE

Figure 1 shows the life cycle of an ICT hardware product, which does in principle not differ from the life cycle of any other type of product: It consists of the three phases production, use and end of life. The design phase has been added here pro forma to have this phase in place as a point where the life cycle can be substantially influenced (see section 3). The design phase is generally not viewed as a part of the life cycle in LCA methodology. Each process along this chain consumes materials and energy and uses some infrastructure. Both the materials and energy transformed and the infrastructure have their own life cycles, which have to be accounted for as well. This idea would lead to infinite recursion in life cycle modelling [Hilty and Schmidt, 1997], unless rules for drawing system boundaries are applied. Such rules are part of standard LCA methodology [Frischknecht and Rebitzer, 2005; ISO, 2006a, 2006b]. According to ISO, "the product system should be modelled in such a manner that inputs and outputs at its boundary are elementary flows". In the case of a life-cycle study of an ICT hardware product, this means, e.g., that the primary production of the metals used in production, the supply chain for the energy used in each phase as well as the final disposal activities are traced through the exchange of chemical elements with the environment.

Viewed from this perspective, it becomes apparent that the current discussion about the energy consumption of ICT and "green data centers" has quite a narrow focus, looking only at energy consumption in the use phase. This is considered an oversimplified perspective because

- the way the electrical energy is generated (electricity supply mix) may be relevant [Frischknecht et al., 2007];
- other life cyle phases may be more relevant than the use phase;
- other environmental impacts may be more relevant than energy consumption.

Although many LCA studies have been done in the ICT field and results are available, this simplistic view is still common in industrial and political initiatives. Life cycle thinking should be part of any Integrated Impact Assessment (IIA) methodology [Ruddy and Hilty, 2007]. Life cycle perspectives are also useful in technology assessment, where risks of prospective applications of new technologies are anticipated [Bauer et al., 2008].

The following studies give typical examples of how the LCA methodology can be applied in the ICT field:

- Several LCA studies on PCs [Atlantic Consulting & IPU, 1998; von Geibler et al., 2003; Hikwama, 2005; Eugster et al., 2007]
- A study on computer screen technologies [Socolof et al., 2001, 2005]
- Two LCA studies on mobile phone networks [Scharnhorst et al., 2005, 2006a, 2006b; Emmenegger et al., 2006] and one on fixed-line networks [Takahashi et al. 2003]
- Several LCA studies on components, such as ICs [Tekawa et al., 2002; Andrae et al., 2004; Andrae, 2005] and batteries [Rydh et al., 2002; Fisher et al., 2006]
- Two LCA studies on the benefits of electronic waste recycling [Hischier et al., 2005; Choi et al. 2006]

Some of the results will be referred to in the following subsections.

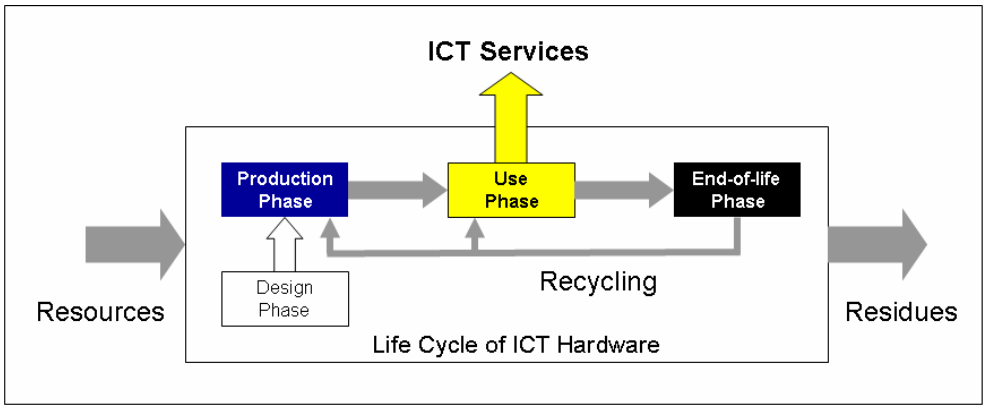


Figure 1. The Life Cycle of ICT hardware. All phases cause (direct or indirect) resource consumption and the release of residues to the environment.

2.1 Environmentally relevant aspects of ICT production

Early studies on the ecological impacts of computers (such as that compiled by Hilty et al. [2000]) showed that the production phase matters. The impact of the production phase of computers and other ICT products cannot be neglected in comparison with the use phase.

The most recent study on desktop PC production [Eugster et al., 2007] showed that the Cumulative Energy Demand (CED) of PC production in China roughly equals the CED of three years of use under average conditions.[1] If all relevant environmental impacts are aggregated using the Eco-Indicator'99 methodology, it turns out that the PC even had to be used for roughly six years until the use phase countervails the production phase. If an energy supply mix with a sizable share of non-fossil energy is assumed (i.e. the PC is used in a country with a lot of hydro-, wind- or solar power), the comparison between production and use becomes still better for the use phase.

Which aspect of the production of a PC has the highest environmental impact? According to Eugster et al. [2007], the production of integrated circuits and other parts containing precious metals[2] is the most energy-intensive part. Assembling the parts is almost irrelevant in comparison.

2.2 Environmentally relevant aspects of ICT use

In the use phase, an ICT product delivers its service to the user. LCA studies define a functional unit (such as "1 hour of PC use") to which the environmental impact of the whole life cycle is related.

Given the relatively high environmental impacts of ICT production, the use of ICT hardware has obviously one environmentally relevant parameter: the length of its useful life. If a device is used for six years instead of three years, it can deliver double the amount of functional units, which implies that the share of the production impact attributed to a functional unit is cut by half.

As a rule of thumb, the length of the useful life of ICT devices is more important than their power consumption during use. Short software innovation cycles with increasing hardware

[1] However, in the case of servers in data centers, which are run 7x24 hours a week, energy consumption in the use phase may dominate the life cycle. This is also because indirect energy demand often occurs for air conditioning and for uninterruptible power supply (UPS) units.

[2] Extracting precious metals, in particular gold, from the earth crust creates high environmental impacts. This contributes a relevant part to the overall environmental impact of electronics.

requirements have therefore a negative effect: They shorten the useful life of the hardware, which is ecologically disastrous [Hilty et al., 2006b].

2.3 Environmentally relevant aspects of the end-of-life phase of ICT

Electronic waste recycling is a field of great interest in both research and politics [Hilty, 2005; Widmer et al., 2005]. Increasing amounts of Waste Electrical and Electronic Equipment ("WEEE" or "e-waste" for short) make this type of waste and its treatment a rapidly growing problem in many parts of the world.

Although there are methodological challenges involved in assessing the environmental impacts of metal supply [Althaus and Classen, 2005], it can be shown that e-waste recycling clearly pays off in terms of environmental impacts because primary production can be avoided when metals are recovered. Some 20-25% of the energy consumed in the production phase can theoretically be saved by recycling [Hischier et al, 2005]. Hence, it is mainly the reuse of metals that makes WEEE recycling environmentally beneficial.

Unfortunately, this is only true for recycling processes that can handle the pollutants contained in WEEE properly. This is not the case in the so-called informal recycling which is very common in emerging economies. There is a huge informal industry in Asia which recycles WEEE mainly to recover gold and copper, an activity which is very profitable as long as the health and environmental impacts are not accounted for. As an example, Figure 2 shows the extraction of copper from printed wiring boards in a typical backyard company with 12 workers in Delhi, India. This company recovers 1-2 tons of copper per month [Widmer et al., 2005].

Many countries are currently implementing WEEE recycling systems. Unfortunately, ever-larger fractions of electronic scrap may escape these systems if the trend towards embedded ICT systems continues. According to the technological visions of *pervasive computing*, *ubiquitous computing*, or *ambient intelligence*, most ICT products will be integrated into every-day objects and no longer be perceived as discrete ICT devices [Som et al., 2004]; this will create new challenges for recycling [Hilty, 2005]. The diffusion of RFID transponders (also known as smart labels), which is currently just beginning, can be viewed as a forerunner of pervasive computing. Tagging papers, food packages etc. may have consequences for the efficiency of established recycling processes and for the availability of scarce metal resources in the future [Wäger et al., 2005].



Figure 2. Manual recovery of copper from electronic waste in India [Widmer et al., 2005]

3. EFFECTS OF ICT APPLICATIONS ON THE LIFE CYCLE OF PRODUCTS

In order to classify the effects of ICT applications on other products' life cycles, we will adopt a conceptual framework that was originally created to categorize ICT impacts on traffic:

- Induction: ICT applications can induce traffic. *Example:* People who get to know each other via the Internet may want to meet some day in person.
- Optimization: Traffic processes are optimized by specific ICT applications. *Example:* Intelligent transport systems (ITS) may avoid congestion.
- Substitution: ICT-based processes replace traffic. *Example:* Virtual meetings are substituted for physical meetings.

Although created for the impact of ICT on traffic, this conceptual framework can easily be generalized to other applications fields.

Figure 3 shows how the service provided by using an ICT product can have an effect on the life cycle of another product. There are three types of effects: induction (straight arrow), optimization (dotted arrows), and substitution (bent arrow).

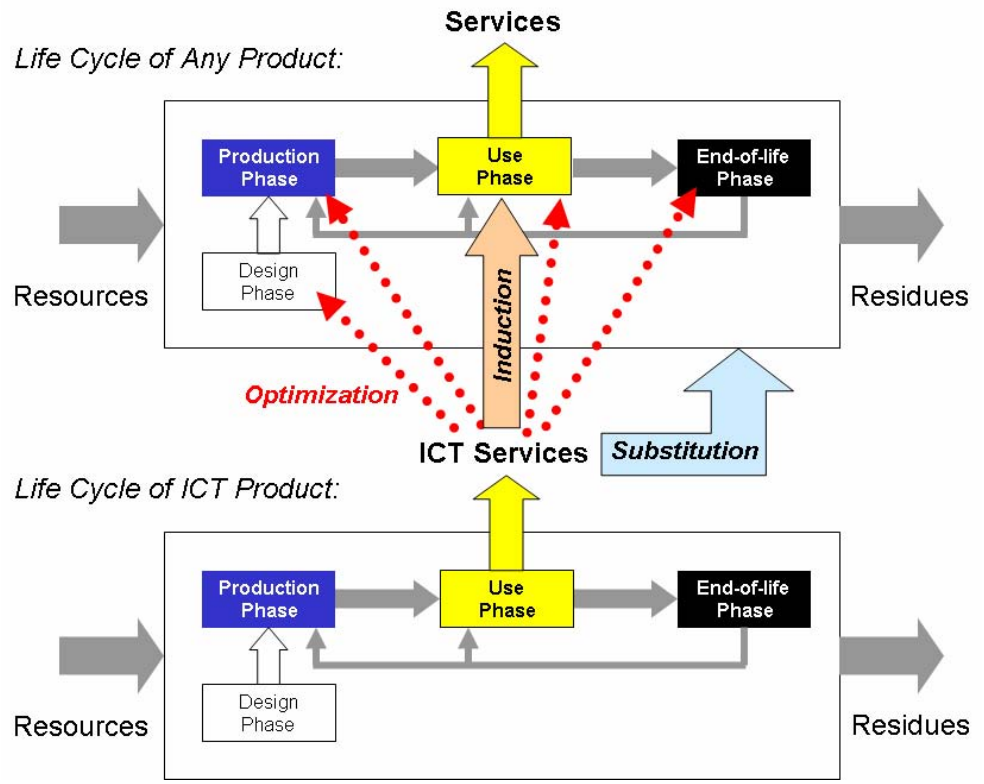


Figure 3. Types of impacts of ICT services on the life cycle of another product or service.

Induction effects occur when an ICT service stimulates the use of the other product, i.e. more functional units per unit of time are consumed (e.g. the text-processing service provided by a PC system with a printer may stimulate paper consumption).

Optimization effects may occur in all phases of the life cycle, as well as in the design phase. CAD tools, for example, can be used to optimize a product for environmental criteria (eco-design). Design has a strong impact of the life cycle because it constrains the optimization potentials that will exist in the production, use and end-of-life phases. For example, if the

variety of materials or the complexity of the product can be reduced in the design phase, it will be possible to reach a higher efficiency level in end-of-life treatment.

Substitution effects occur when an ICT service replaces the use of a physical product, e.g. when e-mail replaces the use of conventional letters.

With LCA methodology, it is possible to quantify the potential environmental benefit of specific optimization and substitution effects. For example, a given conference can be hypothetically virtualized and the difference in environmental impacts assessed. We did so for *EnviroInfo 2001*, which was held at ETH Zurich, assuming that all travel of participants would have been replaced by Internet connections for video streaming, online discussions, upload and download of presentations, etc., including the estimated environmental impact of the Internet connections. It turned out that the virtual conference would only have caused 2-3% of the actual environmental impact of the physical conference. The main part of this could be attributed to the 6% of the participants who had to take a long-range flight: Their flights accounted for 60% of the overall environmental impact of the conference [Hischier and Hilty, 2002].

4. TOOLS AND DATABASES FOR LIFE CYCLE ASSESSMENT

Life Cycle Assessment (LCA) involves modelling and the collection of inventory data. The life cycle is viewed as a system of processes (exchanging energy and materials) that is modelled in an assumed stationary state. Inventory data specifies the inputs and outputs of each process. In order to do an LCA efficiently, the connection between the modelling system and a comprehensive inventory database is essential.

Two examples of tools which fulfil this condition are *Umberto* [ifu, 2008; see Figure 4] and *SimaPro* [PRé consultants, 2008]. Both integrate the life cycle inventory database *ecoinvent*, which is one of the most advanced collections of life cycle inventories world-wide [Frischknecht et al., 2005a, 2005b, 2005c]. Version 2.0 which was released in November 2007, includes data on electronics, which makes it possible to do LCA studies in the ICT field [ecoinvent Centre, 2008]. The US National Renewable Energy Laboratory provides the *U.S. Life-Cycle Inventory Database*. The data can be imported into major LCA tools [NREL, 2008].

The *Umberto* example (Figure 4) shows part of an oil refinery modelled as a network of processes (boxes in the notation used by Umberto) connected by flows. Each process transforms incoming flows into outgoing flows according to a set of linear equations. However, more complex algorithms may also be applied to model a process. The life cycle of any product or service is such a network, albeit much larger than the cutout shown.

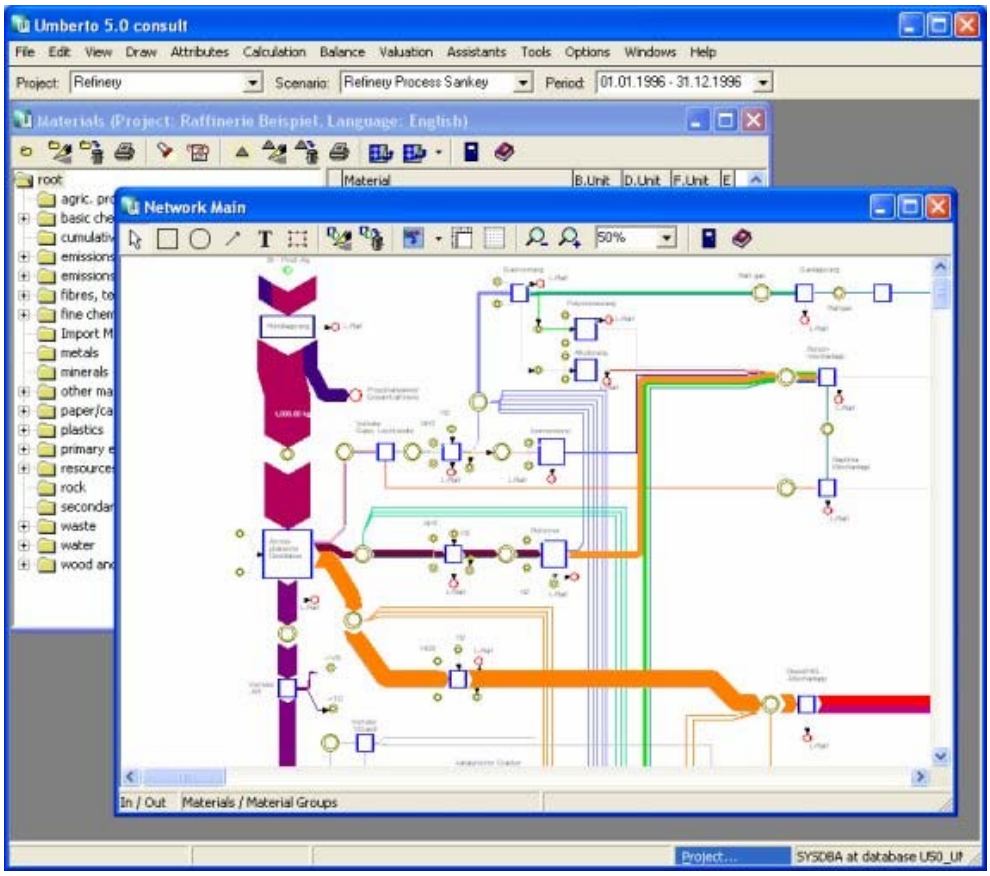


Figure 4. Example of a material flow network modelled with *Umberto*. The Sankey-type visualization show here supports the user in recognizing the relevance of potential measures to improve the system [ifu, 2008].

The static nature of LCA is not always compatible with the dynamic technological development in the ICT field. In order to do *prospective* studies of the positive or negative environmental effects of ICT, dynamic models may be needed, as demonstrated in the study "The Future Impact of ICT on Environmental Sustainability" commissioned by the Institute for Prospective Technological Studies (IPTS) of the European Commission [Hilty et al., 2006a]. The System Dynamics model used in this study considered that ICT consumer products will become smaller and more numerous in the future. Dynamic models have proven to be more adequate than LCA when waste generation and recycling are key issues, because stocks of scrap may play an important role in such contexts [Krivtsov et al., 2004; Streicher-Porte et al., 2005].

5. CONCLUSIONS

The life cycle of ICT hardware has a considerable environmental impact. The main causes are the use of scarce metals and toxic components. The former account for a considerable part of life-cycle-wide energy consumption; the latter cause pollution whenever the end-of-life-treatment is inadequate. The software update cycle with its increasing hardware requirements contributes to the overall impact by shortening the useful life of the hardware.

However, the environmental impact of the ICT hardware life cycle is still much smaller than that of many materials- and energy-intensive processes (such as traffic or space heating) that can be optimized, and in some cases replaced, by ICT services.

As a tool for systems analysis and decision support, ICT has an important role to play in enabling and promoting *life cycle thinking.* Here is where environmental informatics [Hilty et al., 2006c] can make a substantial contribution. Environmental informatics can provide

tools that efficiently and effectively support decision makers who want to improve the ecological life cycle of products and services.

REFERENCES

Althaus, H. J., & Classen, M.: Life Cycle Inventories of Metals and Methodological Aspects of Inventorying Material Resources in Ecoinvent. International Journal of Life Cycle Assessment, 10 (2005) 43-49

Andrae A. S. G., Zou G. and Liu J.: LCA of Electronic Products. An environmental assessment of Gallium Arsenide Monolithic Microwave Integrated Circuit System-In-a-Package (SIP) Switch Product. In: Int J LCA, 9/1 (2004), pp. 45 - 52.

Andrae A. S. G.: Environmental life-cycle assessment in microelectronics packaging. Ph.D. thesis at the Department of microtechnology and nanoscience, Chalmers University of Technology, Göteborg (2005).

Atlantic Consulting and IPU: LCA Study of the Product Group Personal Computers in the EU Ecolabel Scheme, (1998).

Bauer, C., Buchgeister, J., Hischier, R. et al.: Towards a Framework for Life Cycle Thinking in the Assessment of Nanotechnology. Journal of Cleaner Production, 16 (2008) 910-926

Choi B.-C., Shin H.-S., Lee S.-Y. and Hur T.: Life Cycle Assessment of a Personal Computer and its Effective Recycling Rate *In: International Journal of Life Cycle Assessment*, 11/2 (2006), pp. 122-128.

ecoinvent Centre: The ecoinvent Centre Portal. http://www.ecoinvent.org/ (accessed 01.03.2008)

Emmenegger, M. F., Frischknecht, R., Stutz, M. et al.: Life Cycle Assessment of the Mobile Communication System UMTS – Towards Eco-Efficient Systems. International Journal of Life Cycle Assessment, 11 (2006) 265-276

Eugster, M., Hischier, R., Huabo, D.: Key Environmental Impacts of the Chinese EEE-Industry – A Life Cycle Assessment Study. Empa and Tsinghua University, St.Gallen and Bejing (2007)

Fisher K., Wallén E., Laenen P. P. and Collins M.: Battery Waste Management Life Cycle Assessment. Environmental Resources Management ERM, Ltd. (2006).

Frischknecht, R., & Rebitzer, G.: The Ecoinvent Database System: A Comprehensive Web-Based LCA Database. Journal of Cleaner Production, 13 (2005a) 1337-1343

Frischknecht, R., Althaus, H. J., Doka, G. et al.: Selected Modelling Principles Applied in the Ecoinvent Database. Journal of Life Cycle Assessment, Japan, 10 (2005b) 112-122

Frischknecht, R., Jungbluth, N., Althaus, H. J. et al.: The Ecoinvent Database: Overview and Methodological Framework. International Journal of Life Cycle Assessment, 10 (2005c) 3-9

Frischknecht, R., Althaus, H. J., Bauer, C. et al.: The Environmental Relevance of Capital Goods in Life Cycle Assessments of Products and Services. International Journal of Life Cycle Assessment, 12 (2007) 7-17

Hikwama B. P.: Life Cycle Assessment of a Personal Computer. University of Southern Queensland, Southern Queensland (2005).

Hilty, L. M., Schmidt, M.: Der fraktale Produktlebenszyklus. Umweltwirtschaftsforum 4, 1997, 52-57

Hilty, L. M., Ruddy, T., Schulthess, D.: Resource Intensity and Dematerialization Potential of Information Society Technologies. University of Applied Sciences Northwestern Switzerland, Olten 2000 (Discussion Papers; 2000-01)

Hilty, L. M., Arnfalk, P., Erdmann, L. et al.: The Relevance of Information and Communication Technologies for Environmental Sustainability - A Prospective Simulation Study. Environmental Modelling and Software, 21 (2006a) 1618-1629

Hilty, L. M., Köhler, A., Von Schéele, F. et al.: Rebound Effects of Progress in Information Technology. Poiesis & Praxis: International Journal of Technology Assessment and Ethics of Science, 4 (2006b) 19-38

Hilty, L. M., Page, B., Hřebíček, J.: Environmental Informatics. Environmental Modelling and Software, 21 (2006c) 1517-1518

Hilty, L. M.: Electronic Waste - an Emerging Risk? Environmental Impact Assessment Review, 25 (2005) 431-435

Hischier, R., & Hilty, L. M.: Environmental Impacts of an International Conference. Environmental Impact Assessment Review, 22 (2002) 543-557

Hischier, R., Wäger, P., Gauglhofer, J.: Does WEEE Recycling make Sense from an Environmental Perspective? the Environmental Impacts of the Swiss Take-Back and Recycling Systems for Waste Electrical and Electronic Equipment (WEEE). Environmental Impact Assessment Review, 25 (2005) 525-539

ifu – Institut für Umweltinformatik Hamburg: Umberto – know the flow. http://www.umberto.de/en/product/index.htm (accessed 1.3.2008)

ISO – International Organization for Standardization: Environmental Management - Life Cycle Assessment - Principles and Framework. European Standard EN ISO 14'040, Geneva (2006a).

ISO – International Organization for Standardization: Environmental Management - Life Cycle Assessment - Requirements and Guidelines. European Standard EN ISO 14'044, Geneva (2006b).

Kim S., Hwang T. and Overcash M.: Life Cycle Assessment Study of Color Computer Monitor. In: International Journal of Life Cycle Assessment, 6/1 (2001), pp. 35-43.

Krivtsov, V., Wäger, P. A., Dacombe, P. et al.: Analysis of Energy Footprints Associated with Recycling of Glass and Plastic - Case Studies for Industrial Ecology. Ecological Modelling, 174 (2004) 175-189

NREL – National Renewable Energy Laboratory: U.S. Life-Cycle Inventory (LCI) Database. http://www.nrel.gov/lci/ (accessed 05.04.2008)

PRé consultants: The ecoinvent database. http://www.pre.nl/ecoinvent/default.htm (accessed 1.3.2008)

Rydh C. J. and Karlström M.: Life cycle inventory of recycling portable nickel-cadmium batteries. In: Resources, Conservation and Recycling, 34 (2002), pp. 289 - 309.

Ruddy, T., & Hilty, L. M.: Impact Assessment and Policy Learning in the European Commission. Environmental Impact Assessment Review, corrected proof (2007)

Scharnhorst, W., Althaus, H. J., Hilty, L. M. et al.: Environmental Assessment of End-of-Life Treatment Options for a GSM 900 Antenna Rack. International Journal of Life Cycle Assessment - Online First, (2005) 12pp

Scharnhorst, W., Hilty, L. M., Jolliet, O.: Life Cycle Assessment of Second Generation (2G) and Third Generation (3G) Mobile Phone Networks. Environment International, 32 (2006a) 656-675

Scharnhorst, W.: Life Cycle Assessment in the Telecommunication Industry: A Review. International Journal of Life Cycle Assessment, (2006b) 1-12

Socolof M. L., Overly J. G. and Geibig J. R.: Environmental life-cycle impacts of CRT and LCD desktop computer displays. In: Journal of Cleaner Production, 13/13-14 (2005), pp. 1281-1294.

Socolof M. L., Overly J. G., Kincaid L. E. and Geibig J. R.: Desktop Computer Displays: A Life-Cycle Assessment. Volume 1. EPA-744-R-01-004a. U.S. Environmental Protection Agency (US-EPA), Washington D.C. (2001).

Som, C., Hilty, L. M., Ruddy, T. F.: The Precautionary Principle in the Information Society. Human and Ecological Risk Assessment, 10 (2004) 787-799

Streicher-Porte, M., Widmer, R., Jain, A. et al.: Key Drivers of the e-Waste Recycling System: Assessing and Modelling e-Waste Processing in the Informal Sector in Delhi. Environmental Impact Assessment Review, 25 (2005) 472-491

Takahashi, K.I. , Nakamura, J., Maeda, T., Origuchi, T., Kunioka, T., Harada, H., Miyamoto, S., Fujimoto, J.: Environmental effects of information telecommunication networks in Japan. IEEE International Symposium on Electronics and the Environment, 2003.

Tekawa, M., Fujimoto, J., Sakurai, T., Hashimoto, S.: Life Cycle Assessment for Electronic Device and Energy Reduction of Products Using Electronic Devices. NEC Technical Journal 55/4 (2002) 152-155

von Geibler J., Ritthoff M. and Kuhndt M.: The environmental impacts of mobile computing. A case study with HP. Wuppertal Institute, Wuppertal (2003).

Wäger, P. A., Eugster, M., Hilty, L. M. et al.: Smart Labels in Municipal Solid Waste - A Case for the Precautionary Principle? Environmental Impact Assessment Review, 25 (2005) 567-586

Widmer, R., Oswald-Krapf, H., Sinha-Khetriwal, D. et al.: Global Perspectives on e-Waste. Environmental Impact Assessment Review, 25 (2005) 436-458

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Current Trends in *e*Environment and its Role in *e*Democracy

J. Hřebíček[a], R. Legat[b], and M. Nagy[b]

[a] *Institute of Biostatistics and Analyses, Masaryk University, Kamenice 126/3, 625 00 Brno, Czech Republic (hrebicek@iba.muni.cz)*

[b] *Umweltbundesamt GmbH,Spittelauer Lände 5, 1090 Wien, Austria (rudolf.legat&michael.nagy@umweltbundesamt.at)*

Abstract: The Ad hoc Committee on *e*Democracy of the Council of Europe (CAHDE) was established and started its work at September 2006. Its second plenary meeting was held in Strasbourg on October 2007, where the *White paper* of *eEnvironment* (Electronic access to Environmental information) was presented. Its fundamental ideas were prepared during discussion on the workshop *Seamless Access to Environmental Information in the EU - Building an Integrated Information Space for the Environment* of the conference EnviroInfo 2007 in Warsaw. The *e*Environment basis is the Aarhus Convention, which is implemented in the European Community and supported by the EU Directives: 2003/4/EC *(Public Access to Environmental Information)*; 2003/35/EC *(Public Participation)*; 2003/98/EC *(Re-use of Public Sector Information)* and 2007/2/EC *(Infrastructure for Spatial Information in the European Community)*. The *e*Environment is going to be one of the fundaments of *e*Democracy. The Communication COM(2008) 46 final *(Towards a Shared Environmental Information System (SEIS))* will provide the basis for any citizen to be informed about environmental matters and to use this information for active participation in decision making. Further, it will be develop a *Single Information Space for Europe in the Environment (SISE)* specified in the Experts Consultation Workshop at Brussels on 15 February 2008. The SISE together with the SEIS will provide some sort of integrated information space in which environmental data and information will be combined with knowledge for a decision support of environmental protection and sustainable development. This will enable a deeper vision of *e*Democracy and it allows for different types of environmental data and information processing that is not currently possible. Thus, *e*Environment fully supports the principles of *e*Democracy and its new trends will be further discussed in detail in the paper.

Keywords: *e*Environment; Environmental information; Aarhus Convention; *e*Democracy; *e*Government; INSPIRE; GMES; SEIS; SISE.

1. INTRODUCTION

The *Ad hoc Committee on eDemocracy of the Council of Europe* (CAHDE) was established by decision of the Committee of Ministers of EU and started its work at September 2006. The committee CAHDE is primarily an intergovernmental body, whose members are delegated by the 46 Member States of the Council of Europe, and by relevant international organisations (e.g. EU, OSCE) and authors of the paper participate in the expert group of CAHDE.

The four themes set out by the Council of Ministers Recommendation of *e*Democracy Rec(2004)15 establishes the main goals for *e*Democracy as being:

- strengthening the participation, initiative and engagement of citizens in national, regional and local public life;
- improving the transparency of the democratic decision-making process and the accountability of democratic institutions;
- improving the responsiveness of public authorities;
- fostering public debate and scrutiny of the decision-making process.

The political decision making process can be expressed by the following iterative stage model, see Figure 1. This model also shows how *e*Democracy applications can be applied to the different stages.

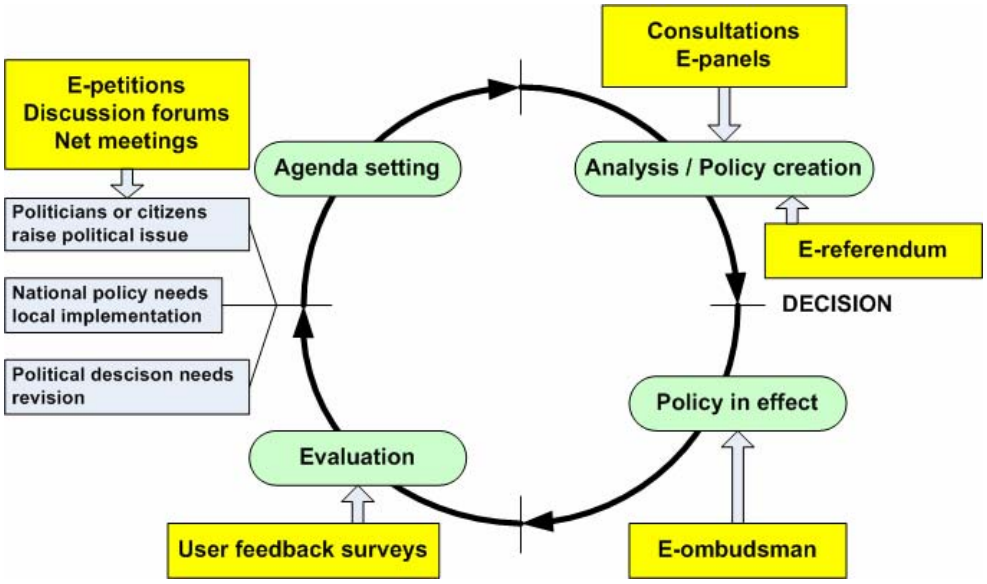


Figure 1. The model of political decision making process in E-Democracy.

*e*Democracy strives to simplify processes between public institutions, the legislative, citizen and businesses regarding exchange information, communication and transactions in many sectors using information and communication tools and technologies (ICT). It simplifies bureaucracy, makes decision and law implementation more transparent and supports public participation in many ways.

The *White paper* of *eEnvironment* (Electronic access to Environmental information) [Nagy, Legat, and Hrebicek, 2007] was presented at the second CAHDE plenary meeting, which was held in Strasbourg on 8 and 9 October 2007. Its fundamental ideas were prepared during discussions on the workshop *Seamless Access to Environmental Information in the EU - Building an Integrated Information Space for the Environment* (Chair: Dr. Thomas Pick) of the 21. International conference EnviroInfo 2007 in Warsaw. They are based on the following considerations [Nagy, Legat, and Hrebicek, 2007]:

- *Control effect*: An effective legal protection for citizen requires that the decision relevant information is available for them. Access to environmental information provides each individual with the possibility to control the compliance with environmental law and to point out deficits in the implementation. Hence, the right to access environmental information leads to a decentralized and effective control of governmental activities by the public.
- *Participation effect*: The right to access the environmental information increases transparency and allows a better public participation in governmental decisions. Therefore, access to environmental information is an important step to participation and democratisation of the environmental legislation.
- *Education effect* (awareness effect): Knowledge regarding the state of the environment is not limited to public authorities; hence this leads to an increased

public acceptance of measures (policies, legislation, plans, programs, etc) for the protection of the environment. This leads to a better awareness of environmental issues in the general public.

- *Prevention effect*: The general right of publication of environmental information should discourage potential polluters of the environment, because this bears the risk of publication of their activities.
- *Standardization effect*: The International Conventions and European Directives dealing with access to environmental and spatial information provide EU-wide comparable principles regarding access to environmental information. This facilitates pan European activities to protect the environment from pollution and prevents competitive distortion.

2. *e*ENVIRONMENT IMPLEMENTATION

2. 1 Legal background

*e*Environment is going to be one of the fundaments of *e*Democracy. It issues also from the Aarhus Convention [UNECE, 1998], which grants the public rights and imposes on Parties and public authorities obligations regarding access to information and public participation and access to justice (the 3 pillars of the Aarhus Convention, see Figure 2).

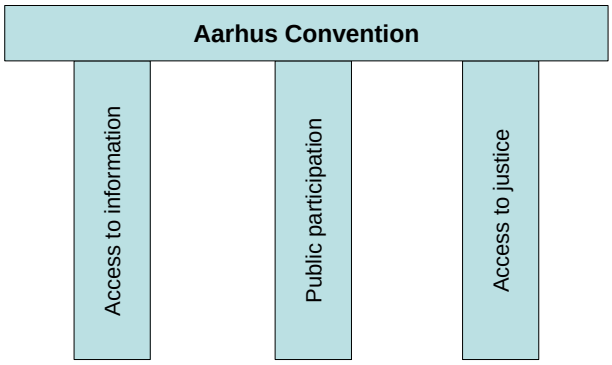


Figure 2. The Aarhus Convention and its 3 pillars.

This provides a good opportunity for the *e*Environment implementation to be an integral part of pan European *e*Government services. Next to legal implementation there are lots recent developments in technically implementing *e*Environment and the 3 pillars of the Aarhus Convention [Nagy, Legat, and Hrebicek, 2007]. With it's 3 pillars the Århus Convention establishes a number of rights of the public (individuals and their associations) with regard to the environment which in the last 5 years have been cast into legal rights for all citizens of Europe.

On European level there is following legislation, which enables to implement the White paper of *e*Environment ideas:

- Directive 2003/4/EC on *Public Access to Environmental Information.* It requires Member States to progressively make available the environmental information held by public administration bodies and disseminate it to the public. Thus, this Directive represents the implementation of the two pillars of the Aarhus Convention (access to information and access to justice).
- Directive 2003/35/EC *providing for public participation in respect of the drawing up of certain plans and programmes relating to the environment and amending with regard to public participation and access to justice Council Directives 85/337/EEC and 96/61/EC.* It provides members of the public with opportunities to participate on the permitting and ongoing regulation of certain categories of activities within Member States. Such opportunities are provided through access

to information, justice, and through consultation on certain key documents. This Directive is the legal implementation of the second and third pillars of the Aarhus Convention (public participation and access to justice), acknowledging that effective public participation in decision making enables the public to express, and the decision-maker to take account of opinions and concerns which may be relevant to those decisions, thereby increasing the accountability and transparency of the decision-making process and contributing to public awareness of environmental issues and support for the decisions taken.

- Directive 2003/98/EC on *Re-use of Public Sector Information.* According to the Directive, public sector bodies collect, produce, reproduce and disseminate documents to fulfil their public tasks and that making this information available to the public should e.g. allow European companies to exploit its potential and contribute to economic growth and job creation. The Directive establishes a minimum set of rules governing the re-use and the practical means of facilitating re-use of existing documents held by public sector bodies of the Member States. It shall ensure that these documents shall bee re-usable for commercial or non-commercial purposes without competitive distortion.
- Directive 2007/2/EC *Infrastructure for Spatial Information in the European Community (INSPIRE).* It is related to the first pillar of the Aarhus Convention (access to environmental information) and lays down general rules aiming at the establishment of an infrastructure for spatial information in the European Community. It is designed to assist policy-making in relation to policies and activities that may have a direct or indirect impact on the environment. The INSPIRE Directive is addressed to public authorities and the general public as users of spatial environmental information. In the context of this Directive *infrastructure for spatial information* means metadata, spatial data sets and spatial data services, network services and technologies as well as agreements on sharing, access and use of data. It furthermore means coordination and monitoring mechanisms, processes and procedures, established, operated or made available in accordance with this Directive. The Directive applies to spatial data held by or on behalf of public authorities and to the use of spatial data by public authorities in the performance of their public tasks.
- Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions COM(2008) 46 final *Towards a Shared Environmental Information System (SEIS).* It sets out an approach to modernise and simplify the collection, exchange and use of the environmental data and information required for the design and implementation of environmental policy. The current, mostly centralised systems for reporting will be progressively replaced by systems based on access, sharing and interoperability. The overall aim is to maintain and improve the quality and availability of information required for environmental policy, in line with better regulation, while keeping the associated administrative burdens to a minimum.

2.2 Barriers of *e*Environment in data and information management

There are main barriers of *e*Environment regarding availability of the collected environmental data and information are due to a range of underlying drivers of a legal, financial, technical and procedural nature:

- *Lack of interoperability and thus "connect ability" and sharing potential between the data and information systems*: many of the collections of data, information and documents, reports etc are sitting in isolated databases and systems, whose existence is not widely known and cannot talk to each other (i.e. are not interoperable). This is often due to the fact that they are mostly custom-built based on protected company standards, designed in isolation without provisions for interoperability with an integrated framework and platform. It is clear that this situation entails a considerable waste of resources to "reinvent the wheel" as available data, information and functional elements and tools are not shared and

accessible to others [Ráček and Hřebíček, 2006] [Štefaník, Hřebíček and Ráček, 2007], [Vögele, 2007].

- *Electronic access to environmental information and electronic Reporting* (the use of Internet) *not yet a reality*: Although there are currently several EU and Member State activities to modernise environmental reporting towards electronic-Reporting systems seeking to benefit from the capabilities offered by the modern ICT [Büchele and Ennőckl, 2005], much monitoring and reporting data is still not reported electronically, while built-in delays in the reporting systems often result in a paper-trail of data that are several years old and obsolete when published or submitted in reports. This decreases enormously the possibility to use efficiently this data for policy decision making and formulation [Vogel, 2008].
- *Lack of comparability and quality of data*: This is a particularly serious problem for both Member States and EU Institutions and bodies, since it limits the usability of the data, prohibits efficient assessment of data and their efficient conversion to information tailored to the users needs, increases considerably the time and the labour needed for duality assurance as well as re-engineering of data and information in order to render them comparable [Vogel, 2008]. The cause of these shortcomings have been identified as resulting from diverging definitions and data collection practices and methodologies, a lack of common standards, insufficient inter-operability between monitoring systems and lack of coordination of monitoring programmes between levels of government and across borders [Hlaváček, Hejč, and Hřebíček, 2007].
- *Data access and sharing restrictions:* The existence of policies and businesses models prohibiting wider accessibility and sharing of environmental related data and information owned by public authorities (but also by research institutions and consortia) and their impact for policy makers, citizens and businesses has been recognised in many forums, reports and media (e.g. GMES, GEO, INSPIRE, Aarhus convention) [Pick, 2007], [Krammer and Legat, 2008].

In order to overcome the above barriers and meet the EU policy objectives of the Lisbon and the European i2010 strategy [i2010, 2005], the IDABC Programme (Interoperable Delivery of European *e*Government Services to public Administrations, Business and Citizens) has shifted its priorities not only to equip organisations of the public sector with ICT but also to promote building such common infrastructures and pan-European *e*Government services [Pillmann, Geiger, and Voigt, 2006], [Hřebíček, Hlaváček, and Zálišová, 2007].

2.3 The role of SEIS in *e*Environment

According to the Communication COM(2008) 46 final the SEIS will provide the knowledge base required to design, implement and evaluate the environmental and other policies that are needed to achieve sustainable development and are underpinned by reliable, scientifically sound, up-to-date and generally *fit-for-purpose*, data and information covering all elements of the DPSIR (Driving forces, Pressures, State of the environment, Impacts and Responses) framework. The general objectives of the SEIS are:

- to secure a clear political agreement around a set of principles on which the SEIS is to be based;
- to continue rationalising the *knowledge base* through the assessment and streamlining of existing reporting requirements within environmental legislation while implementing information and communication technology solutions for electronic reporting;
- to establish and implement data and information sharing agreements in addition to an efficient ICT infrastructure to facilitate the discovery, assessment, access and sharing of environment-related data and information;

- to reinforce and, where necessary, establish monitoring infrastructures and surveys for the collection and archiving of *fit-for-purpose* environment-related data that are cost effective and flexible but can be sustained over the long term.

2.4 The role of EEA and EIONET in *e*Environment

The *European Environment Agency* (EEA) together with the Commission will play the most important role in overcoming above barriers of *e*Environment. EEA particularly plays a crucial role in collecting and providing environmental information, with the help of its *European environment information and observation network* (EIONET). EIONET is a network of some 900 experts from over 300 national environment agencies and other bodies dealing with environmental information in 37 European countries, as well as five European Topic Centres (ETCs) of EEA working on specific environmental themes. Based on input from the EIONET partners the EEA has identified a set of priority annual data flows. These data in the areas of air quality, air emissions, inland waters, marine and coastal waters, contaminated soil, nature conservation and land cover, are used to update the core set of environmental indicators, which form the basis of EEA reports and assessments.

REPORTNET is EIONET's infrastructure for supporting and improving data and information flows and has helped modernise the collection of information and reporting systems [Pillmann, Geiger, and Voigt, 2006]. The system integrates different web services and allows for distributed responsibilities. REPORTNET has initially been mainly used for reporting environmental data to EEA, but is now also hosting some of the Commission's environmental reporting information. The open system also allows for deliveries to be made to other national and International organisations. However to date the information is often uploaded in the form of reports in word processing or spreadsheet format and not structured in connected databases. This makes it difficult to efficiently retrieve, integrate and analyse the information in order to efficiently inform policy and decision making.

As part of the REPORTNET, the EEA has created since 1998 a Reporting Obligations Database (ROD) and a Central Data Repository (CDR). There are our experiences with REPORNET tools: Hřebíček, Pitner and Ráček [2004] developed environmental data models for reporting obligations of the Czech Republic to CDR, as well as they analysed how these obligations are met in practice ROD information tools of the EEA. However ROD has a delay with respect to its development plan and also large dynamic changes, which bring the necessity to change developed environmental data models for reporting obligations of the Czech Republic. REPORTNET will therefore need further development for it to be able to serve as an fully effective tool for a coherent simplified reporting system and provide timely availability of data, information and documents, and this will require agreement on its use by all Member States.

2.5 The role of FP7 and development of SISE

The 7th Research Framework Programme (FP7) objective "*ICT-2007.6.3: ICT for Environmental Management and Energy Efficiency*", [FP7, 2006], introduced a *Single Information Space for the Environment in Europe* (SISE) in which environmental institutions, service providers and citizens can collaborate or use available information without technical restraints. [Schouppe, 2008]. This will be a platform also for *e*Environment. The aim of the SISE is: an ICT research vision for real-time connectivity between multiple environmental resources which would allow seamless cross-system search; as well as cross-border, multi-scale, multi-disciplinary data acquisition, pooling and sharing. Furthermore, it would allow for service-chaining on the Web, thereby stimulating data integration into innovative value-added Web services. Typical topic for this session will include ICT research for Environmental application and services, open semantics and standards, data interoperability and Web communities, data visualisation and modelling including risk assessment, SISE deployment models, and the consolidation of a existing networks to build a European Research Area in the field of ICT for environmental sustainability.

3. CONCLUSIONS

The SEIS together with new FP7 research projects for the SISE [Schouppe, 2008] will provide some sort of an integrated information space in which the connection of *e*Environment with *e*Democracy will support of environmental protection and sustainable development.

The cost and benefits of implementing a connection of *e*Environment with *e*Democracy will depend on the timescale over which it happens, and the precise measures that are taken to achieve it. However, the potential benefits of such a system can nevertheless be expected to be considerable. Improving the mechanisms for collecting, exchanging and using the data can be expected to significantly increase the use that is made of such data, together with a significant reduction in cost for the users. [Vögele at all, 2007]. Improvements in the access and interoperability of data systems will also reduce the need for reporting requirements, leading to a streamlining of data requirements and data flows, including the phasing out or repeal of outdated or redundant reporting requirements.

Other benefits include better legislation, more efficient EU-level analyses, efficiency gains in relation to the achievement of international policy commitments and evaluation obligations, empowerment of citizens through the widespread availability of information, increased availability of data for use by researchers, and better profile for the EU in various global forums. Updating the standardised reporting directive will bring about immediate simplification benefits, will in addition help to trigger further streamlining within Member States and compliance with the SEIS and SISE principles. Citizens in particular stand to benefit from the increased transparency and availability of information that makes full use of the rapidly evolving ICT for the support of *e*Democracy.

ACKNOWLEDGEMENTS

The paper is supported by the Ministry of Environment of the Czech Republic (project No. SP4/i2/26/07), further by the Ministry of Education and Youth (project INCHEMBIOL, No. MSM0021622412).

REFERENCES

Büchele, K. T., and Ennőckl, D., UIG. Umweltinformationgesetz. Kommentar.NWV-Verlag, Wien-Graz. 170 pp. 2005.

FP7, <http://cordis.europa.eu/fp7/ict/>, 2008/03/01

E-Government Arbeitsgruppe UmweltInformation, <http://reference.e-government.gv.at/UW-UI_Umweltinformation.1024.0.html>, 2008/03/01

Hlaváček, M., Hejč, M., and Hřebíček, J., eGovernment Services in Environment - Automate Data Quality Assessment - Czech Republic Approach. In EnviroInfo 2007. 21st International Conference on Informatics for Environmental Protection. Environmental Informatics and System Research. Volume 2. Workshop and application papers. Warsaw, Poland: Shaker Verlag, pp. 159-166, 2007.

Hřebíček, J., Hlaváček, M., and Zálišová, I., eGovernment Services for Environment Protection of the Czech Republic. In Best Practice and Innovation: 5th International Eastern Europe e|Gov Days. Praha, Czech Republic: Österreichische Computer Gesellshaft, pp. 58-65, 2007.

Hřebíček, J., Pitner, T., and Ráček, J., Analysis of Environmental Information Management in Czech Republic. In 6th International Symposium on Environmental Software Systems (ISESS'05). Sesimbra, Portugal: International Federation for Information Processing, pp. 26 -30, 2005.

i2010, COM/2005/0229 final: "i2010 – A European Information Society for growth and employment" <http://ec.europa.eu/information_society/eeurope/i2010/index_en.htm>, 2008/01/03

Krammer, H. J., and Legat, R., E-Government Working Group for Environmental Information. 6th International Eastern Europe e|Gov Days: Result and Trends, Prague, Czech Republic: Österreichische Computer Gesellshaft, pp. 230-235, 2008.

Nagy, M., Legat, R., and Hrebicek, J., Electronic Access to Environmental Information – an Important Fundament for E-Democracy and Environmental Protection, CAHDE (2007) 23 E. <http://www.coe.int/t/e/integrated_projects/democracy/02_activities/002_e-democracy/>, 2008/01/03

Pick, T., From Aarhus to INSPIRE: Putting Environmental Information on the Map. In EnviroInfo 2007. 21st International Conference on Informatics for Environmental Protection. Environmental Informatics and System Research. Volume 2. Workshop and application papers. Warsaw, Poland: Shaker Verlag, pp. 159-166, 2007.

Pillmann, W., Geiger, W. and Voigt K., Survey of environmental informatics in Europe, *Environmental Modelling & Software*, 21(11), 1519-1527. 2006.

Ráček, J., and Hřebíček, J., Interoperability Framework of the Czech Republic in Environmental Information Exchange. In 20th International Conference on Informatics for Environmental Protection. Managing Environmental Knowledge. Graz, Austria: Shaker Verlag, 2006, pp. 509-512, 2006.

Štefaník, M., Hřebíček, J., and Ráček, J., Interoperability in eGovernment Services for Environment Protection. In Best Practice and Innovation: 5th International Eastern Europe e|Gov Days. Praha, Czech Republic: Österreichische Computer Gesellshaft, pp. 68-71, 2007.

Schouppe, M., Towards a Single Information Space for the Environment in Europe. Concept and purpose. Expert Consultation Workshop, European Commission Information Society & Media Directorate-General, Unit ICT for Sustainable Growth. Brussels, 15 February 2008. <ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/sustainable-growth/workshops/workshop-20080215-mschouppe_en.pdf>, 2008/03/01

Vogel, W. R., E-Government and Environmental Protection in Austria. 6th International Eastern Europe e|Gov Days: Result and Trends, Prague, Czech Republic: Österreichische Computer Gesellshaft, pp. 219-224, 2008.

Vögele, T., Kruse F., Klenke M., Lehmann H., Giffei C., Management of Environmental Metadata in Germany - Tools and Experiences. 13th EC-GI & GIS Workshop, Porto, 2007. < http://www.kst.portalu.de/pub/konf/2007/paper_isess_2007.pdf>, 2008/03/01

Vögele, T., An effort to achieve organizational interoparability of Environmental Information - PortalU® and the Environmental Information Infrastructure in Germany. International Symposium in Environmental Software Systems (ISESS) Prague, 22.-26.05.2007. <www.kst.portalu.de/pub/konf/2007/paper_isess_2007.pdf>, 2008/03/01

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Material Flow Analysis for Eco-Efficiency with Material Flow Network Reference Models – Concepts and Case Study

B. Page[a], V. Wohlgemuth[b] and M. Raspe[b]

[a] University of Hamburg, Department of Informatics, Vogt-Koelln-Str. 30, D-22527 Hamburg, Germany (page@informatik.uni-hamburg.de)

[b] FHTW Berlin, Study Program Industrial Environmental Informatics, Blankenburger Pflasterweg 102, D-13129 Berlin (volker.wohlgemuth@fhtw-berlin.de)

Abstract: Material Flow Analysis is a method for determination of material and energy flows in companies for ecobalances and process optimization. Ecobalancing provides a basis for efficient utilization of material resources and energies and for reduction of environmental harmful emissions and waste. Reference models in material flow analysis (e.g. with Material Flow Networks) serve as templates for the development of more refined models in the same domain. They facilitate material flow analyses by reducing effort in time and costs in the model building process by reutilization of predefined structures. In the context of a Thai-German cooperation project a reference model for the Thai agro-industrial branch of native starch production has been designed which reduces the entry level for material flow network modelling in Thai starch industry. Material flow analyses should help to reduce environmental impacts such as water pollution as well as to improve production processes on an economic level. Process improvements for eco-efficiency are targeted on a reduction of starch losses during starch production while a better yield of starch is allowed. The reference model has been developed with the software tool Umberto® which is a graphical modeling tool for Material Flow Networks. The software supports a modular design of the reference model in order to make it easily adaptable to specific production processes in different starch factories. The general, well structured reference model approach for material flow analysis, developed to reduce environmental burdens and financial losses in Thai starch industry, is easily transferable to other industrial branches beyond Thai agro-industry as well as to other regions.

Keywords: Material Flow Analysis; Reference models; Eco-efficiency; Native starch production.

1. INTRODUCTION INTO MATERIAL FLOW ANALYSIS

Material Flow Analysis is an important constituent of Environmental Management Information Systems (EMIS). As modeling approach in material flow analyses so called Material Flow Networks are applied making use of special application software. Material Flow Networks are a method from Environmental Informatics and support computation of material and energy flows. Material Flow Networks, representing relevant energy and material flows and their transformations within a company, have been originally developed at the University of Hamburg as a powerful Material Flow Analysis method; see Moeller [1995] and Moeller [2000]. Material Flow Networks are a graphical modelling notation based on the popular Petri Net methodology for specifying concurrent systems in Computer Science and have been transferred into the Environmental Sciences. They provide both detailed information on environmental effects of a company or on ecological optimization of products as well as the influence of the associated material flows on the costs. In Material

Flow Networks the *transitions*, represented in diagrams by squares, stand for the location of material and energy transformations. Transitions play a vital role in Material Flow Networks, because material and energy transformations are the source of material and energy flows. Another defining characteristic of Material Flow Networks is their concept of *places*. Places separate different transitions. This allows a distinct analysis of every transition. Beyond that places can describe inventories for materials. Circles are used in diagrams to represent places. Arrows show the paths of material and energy flows between transitions and places. They constitute a combined flow and inventory analysis:

Notation used in Network Model of the starch production process:

- The central components of Material Flow Networks are the **Transitions**, which are marked as blue **boxes** in the network. They specify energy and material transformation processes in a production system. There are predefined rules for transformation of input into output materials which can be based on parameters, mathematical functions, standard transitions from libraries, or user defined transformation programs.
- **Circles** are so-called **Places**, i.e. inventory and connections without material transformation. They describe system boundaries (Input-/Output), take care of branching and fusing in the network or denote inventories.
  Green circles marked with a vertical line on the left are **Input Places** for materials and energy; **red circles** are **Output Places** for products and emissions. **Yellow circles** are connection places for material flow branching and storage.
- **Connections** shown as directed **lines** represent material and energy flows from Places to Transitions and from Transitions to Places. Several materials can flow in one arrow.
- Embedded blue boxes within another box denote **Submodels** which include another partial network in order to reduce model complexity.

Figure 1: Graphical Notation used in Material Flow Networks [Pimpisut et al., 2007]

Based on this early innovative Environmental Informatics idea a suitable software tool has been developed first as a prototype at the university institute and afterwards as a commercial software product under the name of Umberto® (s. www.umberto.com) in the nineties. Meanwhile Material Flow Networks and the related modelling software Umberto® are an established method and tool in the context of EMIS for building up accounting systems for relevant material and energy flows and transformations (including cost data) within a production process in a company in order to achieve Eco-Efficiency [Moeller et al., 2001]. Today the software is widely used in many industries on the national as well as on the international level.

Since material flow analyses can be very time consuming and costly it would be an advantage if a template for Material Flow Networks in a special industrial branch was at hand from the very beginning. Reference models offer such templates for the development of Material Flow Networks and allow a reduction of effort in material flow analyses.

One recent international application of Material Flow Networks has been a joint Thai-German project for establishing an EMIS in Thai native starch industry presented in this paper as well as in palm oil industry published elsewhere [Page et al., 2007] .

The aim of this work is the development of a reference model for material flow analyses in order to facilitate them. In addition an indicator system is defined to evaluate material flows. The main motivation of this approach beyond the facilitation of material flow analyses is an environmental and economic improvement of production processes.

This paper is organized as follows: After this short introduction into Material Flow Networks as a modelling approach in material flow analysis and to the Thai-German

cooperation project we discuss the general ideas of reference modelling in material flow analysis. In Chapter 3 we introduce a case study in Thai starch industry where a production reference model as subtask of a joint Thai-German development project for an EMIS has been deployed. The background of the EMIS project is disclosed and the system components briefly described. The different steps of the design process of the starch production reference model, the model structure, the Key Performance Indicators and computational results are explained. Finally some concluding remarks with a short summary of the findings in the project as well as with some hints on the generality of the approach beyond the special application domain round up this paper.

2. REFERENCE MODELS IN MATERIAL FLOW ANALYSIS

Since we cannot rely on a generally accepted definition for *Reference Models* in literature we understand Reference Models as information models for supporting the construction of other models in a given domain or industrial branch. They represent an idealized mapping of systems onto models in a special domain and serve as template for model building in this area. Therefore the main asset of reference models is its benefit in generating more specific models with less effort. In Material Flow Analysis the effort in modeling as well as data collection can be significantly reduced by predefined model structures and data sets. Also an adaptation of given model structures is straightforward for achieving a good match with the real world production system.

In order to achieve the aim of providing a model template with reference models they have to meet a number of requirements. First we gear towards *generality*, which means that a reference model has to be suitable for the deduction of a range of specific models. A*daptability* of reference models to special needs as well model *extensibility* are required. *Usability* means that a reference model should be specified to the extent that it can be already applied without modifications. In the context of material flow analyses reference models should support all modeling steps completely. With the delivery of characteristic structures and generalized data in a reference model the requirement of *completeness* is fulfilled. Beyond that, modifications of reference model components should be supported by user instructions and hints. The greatest assets of reference models lay in the model building phases of model design, data collection and model implementation. In *model aggregation* a specific model is assembled from different model modules. Therefore a *modularization* of reference models is useful where sections of a production process or plant are defined and compiled as partial reference models or sub models, respectively. The user might concentrate on a few production sections with separate partial reference models which have to be relatively independent from each other.

In order to carry out a material flow analysis efficiently a comprehensive model *documentation* is required which describes model application step by step and explains each part of the reference model in an understandable manner. Since the modeling process is carried out in a software environment data and model structures are reusable. Finally *transparency* of all modeled structures, collected data and of computational model results is required in the reference model building process.

3. CASE STUDY: REFERENCE MODEL FOR NATIVE STARCH PRODUCTION IN THAILAND

The Department of Industrial Works (DIW) in Bangkok in cooperation with the German Society for Technical Cooperation (GTZ) has carried out a development project “Management Information Systems for Industrial Pollution Prevention and Control” from 2005 until 2007, where the authors have been acting as project consultants. The main result of this project has been the development and implementation of a special Environmental Management Information System for Eco-Efficiency in different agro-industrial sub-sectors including native starch in Thailand. An integral part of this EMIS has been a material flow reference model for native starch production based on Material Flow Network methodology. Native starch is used in food, textile and paper industry. Starch production is one of the largest and most important agro-industrial branches in Thailand. However, the

production of native starch from Tapioka plants has negative impacts on the Thai environment. In particular this is obvious in water pollution caused by high organic exposure of sewage from the production process. Beyond that starch losses accumulate during the different steps of the production process. In order to investigate and improve ecologically relevant processes material and energy flows at the production of native starch are analyzed by means of material flow analysis. For example by balancing these flows it can be estimated how much and where exposed sewage at the production process of starch per ton is accumulated. Based on the results measures for improvement can be taken with positive effects on the environment. In Thailand around 85 native starch factories exist running similar production technologies. A reference model can provide a standardized Material Flow Network for these factories in order to reduce their effort in material flow analysis.

3.1 Components of Environmental Management Information System

The EMIS aims at the enlargement of the information basis for the management in the native starch factories and therefore at an improvement of the management decision process under environmental as well as economic criteria.

The MIS Software for Eco-Efficiency consists of different components including the data base application, the material flow reference production models as well as the Assistant programs. An essential component of the EMIS for eco-efficiency is a production data base application for the Native Starch sector. The main focus of the data base applications is short term data (daily to monthly and beyond) and monitoring of the production process on a timely basis. They provide (economic as well as ecological) Key Performance Indicators (KPIs) on a short term scale as well as short term reports supporting factories in day-to-day business. The data base applications have been introduced only recently in a number of native starch factories in different regions of Thailand as a data collection and data storage facility in a stepwise procedure. The users in the companies will also be able to produce some standard reports and outputs from the standard relational data base management system.

In contrast material flow analysis reference Models (MFA-models) based on Material Flow Networks have a long term focus on flexible time periods such as quarters, six months, or yearly periods and beyond. Thus another central component of the MIS for Eco-Efficiency is a suitable EMIS software product such as Umberto® supporting the development of material flow analysis models. This software component reads in company measurement data from the data base by defined interfaces and runs Material Flow Reference Models for the selected industrial sub sector of Native Starch production (as well as other sectors like Palm Oil). The reference model represents the principal processing steps of the production process of native starch. Since MFA modelling is very much a data-driven approach, the data base application of the EMIS plays an important role as central data source. In addition process knowledge by specialists in the project team has been incorporated into the modelling process.

Due to the typical similarities in the production processes in most companies of an industrial sub-sector the MFA model represents a branch-specific Reference Model used by all companies. Quantitative differences between the productions in the single companies are covered by varying parameter values derived from the respective company data base. Structural differences in the production processes, however, require model adaptations on submodel level, i.e. there might be different submodel versions in operation in single factories taking care of differences in technical production processes. The MFA model delivers long term KPI´s, material and energy balances for the whole production process or of sections of the process, can highlight critical flows and provides what-if-analyses comparing different options for improvements and cost reductions. Typical analysis results are for instance comparisons of company eco-efficiency with benchmark values defined for the industry under investigation.

IT-users in the companies such as production managers will be working with the runtime version of the EMIS model ("Umberto®-Runtime") and will be supported by so-called Assistants in order to facilitate the EMIS application. The user is prompted for certain data

items which can be typed in manually or accessed from the data base in a guided dialog. The Assistants will transfer the input data and parameters to the EMIS model for processing. Beyond data input the selection of available reports and analysis is also supported by the guided dialog. The Assistant program is an essential prerequisite for bringing MFA-models to non-technical users in Small and Medium Enterprises typical for agro-industrial sectors in Thailand. The EMIS architecture and the Assistant programs are discussed in more detail in Pimpisut et al. [2007].

3.2 Reference Model for Native Starch Production

The development of the reference model followed the typical steps of the model building cycle, see e.g. Page [2005]; the most relevant ones are addressed in this section. The modelling software Umberto® has been applied in the project in order to provide material flow models and transition modules in reusable form. Due to the use of Umberto® with its interactive graphical modelling approach the steps model design and model implementation are closely linked.

At the beginning of the modelling process the general description of the production process of Tapioka Starch has been defined as main goal of model development. Next to generalized empirical data from the production data base information from technical documentation and literature has been incorporated. Inconsistencies in the data had to be identified and cleared. Significant differences in the technical processes resulted in alternative material flow networks on the submodel level.

At the beginning of the modelling process all material and energy types in the reference model have been identified and installed in the material administration of the Umberto® software tool. Materials and energy forms (uniformly named materials in Umberto®) are structured into hierarchical material groups in order to allow a clearly arranged material management. Beyond that this structure can be found in the eco-balances later. Ecological, economic and technical characteristics can be defined as special material properties. For example starch content has been designated as technical material property for the material Fresh Tapioca Root. Ecologically relevant characteristics are sewage types coming up during the different stages of starch production. These are described by COD (chemical oxygen demand)- and BOD (biochemical oxygen demand)-values.

The modelling itself is started with the decomposition of starch production into different production sections, i.e. Material Preparation, Extraction and Drying Process, Utility, and Treatment System (see Fig.3 below). The reference model has been designed on two aggregation levels. On the higher aggregation level the four main sections of starch production are modelled in a rough form. System boarders of the investigated production system are specified by Input and by Output-places. For the linkage of the four sections or the transitions Connection-places (◎) are used. Connection places serve as a well defined transfer from the output of one transaction to the input of the following transaction, if these processes are directly linked. Material and energy flows are displayed by directed edges.

On the second aggregation level the four main sections are modelled as partial reference models with a higher degree of detail. The models are created as subnets of embedded transitions on the first layer (represented as sub models with double lined squares).

These subnets are stored as modules in a library and are retrievable anytime. A horizontal adaptability is reached by this modular architecture of the reference model. In this way different subnets are created for varying production steps in starch production. For example the use of a biogas system is optional in the treatment system. Therefore two alternative subnets are specified for this section.

In the partial reference model Material Preparation the process steps from acquisition to grinding and milling of Fresh Tapioca Root are modelled. The partial reference model Extraction and Drying Process comprise the production steps from Fruit Water Separation over Starch Refining until Cooling and Bagging. Here the separation of fruit water in starch production is optional. Therefore two different Material Flow Networks have been specified including Extraction with as well without fruit water decanter.

The partial reference model Treatment System is also modelled and specified in two different variants as mentioned above (i.e. Treatment with biogas plant and without biogas plant), because the Biogas system is only optional in Thai starch industry. Finally transitions for Water Treatment, Hot Air and Electricity Generation are specified in the partial reference model Utility (s. Fig. 2).

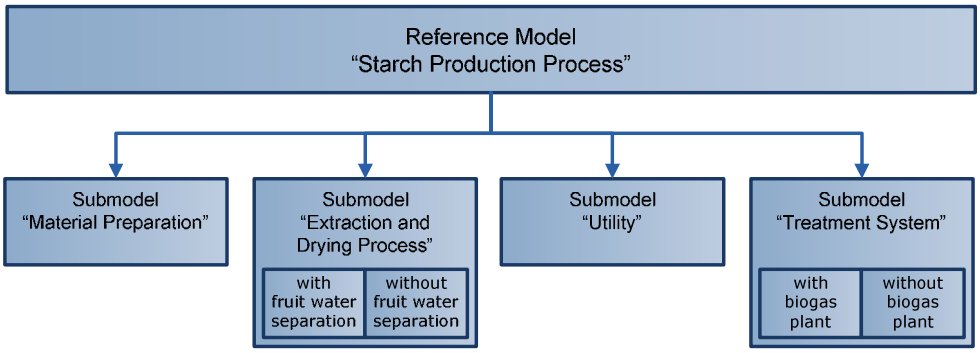


Figure 2: Structure of Reference Model „Starch Production Process" [Raspe, 2007]

The computation of the Material Flow Network for native starch production, based on data for transition parameters, input-output values and material properties, results in an input-output eco-balance. In the balance sheet input/output materials are sorted by material hierarchies defined in material management. The sums of input and output materials are indicated in their basic units of kg and kJ. The computational direction can be influenced by the specification of material flow quantities either as input or as output parameters. For example with the quantity definition of Fresh Fruit Root as input value the model is computing in a forward direction; i.e. starting from the input side all material and energy flows are computed depending on input values. Whereas with the option of backward computing of the network the output value of end product Native Starch is specified and the network is computed step by step towards the input side. Both computational directions have been implemented in the reference model by appropriate transition specifications. Another form of output with Material Flow Network models are so-called Sankey diagrams showing mass-proportional flows through the networks. In this context Sankey diagrams are applied to visualize starch losses along the production process steps of significance interest to the starch factories. In this way identification of significant losses at critical process steps can be highlighted (s. Fig. 3).

An indicator system (called Key Performance Indicators- KPIs) for eco-efficiency has been developed by the Thai project partner DIW in order to monitor and benchmark starch production in the participating Thai factories under economic as well as environmental aspects. The implementation in the reference model has been carried using the Umberto®-Validation System. The Key Performance Indicators have been specified by defining equations with mathematical expressions and applied to the eco-balance data produced by the reference model. These indicators can be grouped by production section.

In Fig. 4 shown in the next section we can see the set of KPIs used in the reference model for starch production, however indicating only fictitious values. Please note that these values are only of exemplary character, because real factory data have not been made available for publication.

4. CONCLUSIONS

The aim of this work has been the structured development of reference models for material flow analyses. An exemplary reference model has been assembled for Thai starch industry as part of an Environmental Management Information System project. The motivation is the simplification of material flow analyses by means of reference model building. Reference models facilitate model building in a special domain because they serve as templates.

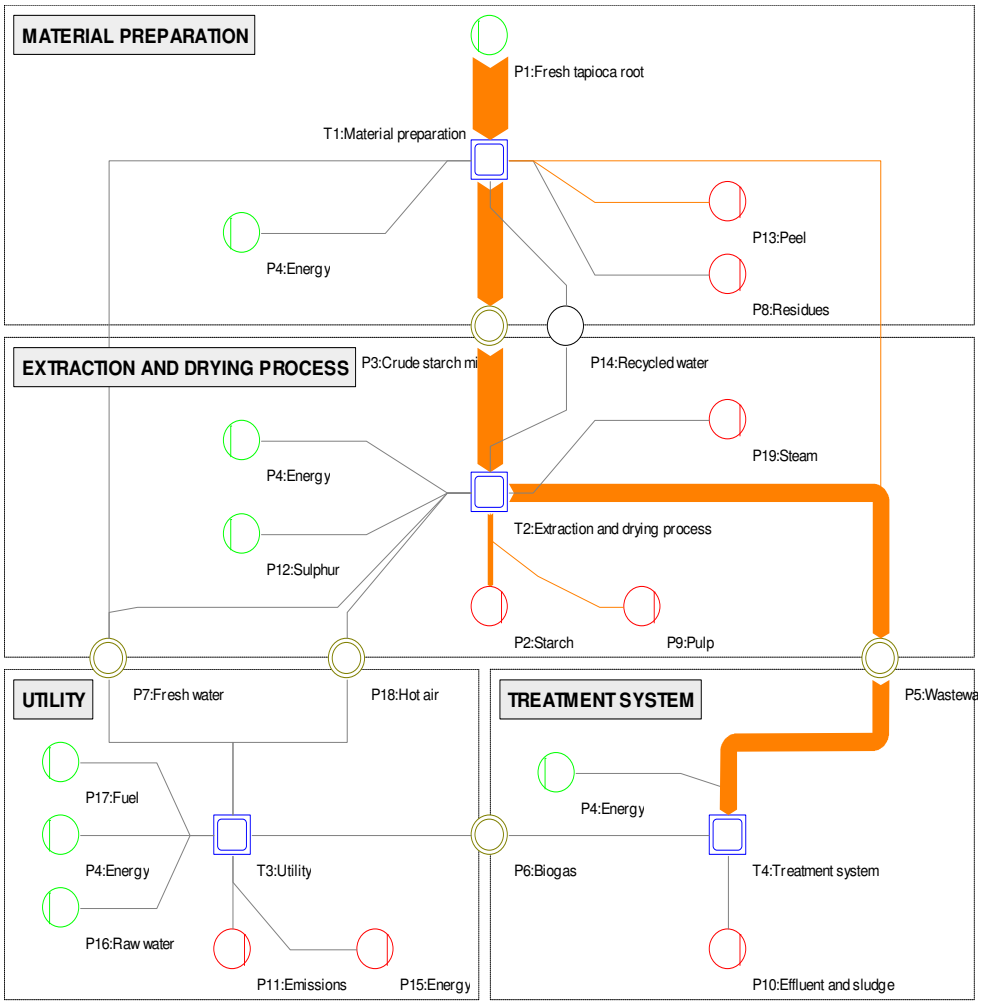
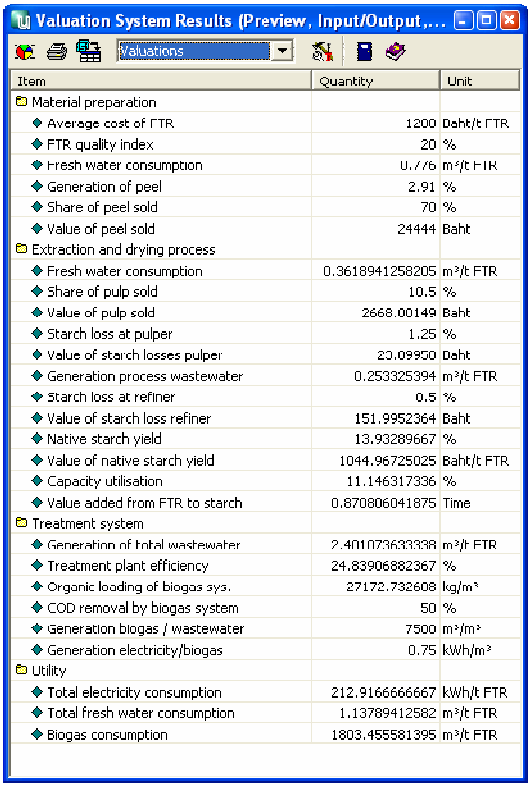


Figure 3: Material Flow Network of Reference Model „Starch Production "

The reference model has been constructed in a systematic fashion following the model building cycle from computer simulation as a guideline. The implementation has been carried out with the material flow analysis modelling tool Umberto® which allows graphical model construction. Based on the information provided in the project a reference model has been developed for Thai starch as well as palm oil industries. It represents the structures, the data and the Key Performance Indicators in a production domain in a general form which can be easily adapted to special production processes in different factories. The modelling approach of using individual subnets allows a further specification of the reference model ("horizontal adaptability"). With the possibility of storing different production variants as modules in a library (as offered by the software tool) a reference model is obtaining completeness for varying production facilities in an industrial branch such as starch production in the course of practical application step by step.

However, the approach of reference model building is of general interest, it does not restrict itself to agro-industrial domains such as starch or palm oil production. The approach of reference models in material flow analysis can be used for many other industrial branches and in other countries as well as long as the production processes are fairly standardized.



Item	Quantity	Unit
Material preparation		
Average cost of FTR	1200	Baht/t FTR
FTR quality index	20	%
Fresh water consumption	0.776	m³/t FTR
Generation of peel	2.91	%
Share of peel sold	70	%
Value of peel sold	24444	Baht
Extraction and drying process		
Fresh water consumption	0.3618941258205	m³/t FTR
Share of pulp sold	10.5	%
Value of pulp sold	2668.00149	Baht
Starch loss at pulper	1.25	%
Value of starch losses pulper	20.09950	Baht
Generation process wastewater	0.253325394	m³/t FTR
Starch loss at refiner	0.5	%
Value of starch loss refiner	151.9952364	Baht
Native starch yield	13.93289667	%
Value of native starch yield	1044.96725025	Baht/t FTR
Capacity utilisation	11.146317336	%
Value added from FTR to starch	0.87080641875	Time
Treatment system		
Generation of total wastewater	2.401073633338	m³/t FTR
Treatment plant efficiency	24.83906882367	%
Organic loading of biogas sys.	27172.732608	kg/m³
COD removal by biogas system	50	%
Generation biogas / wastewater	7500	m³/m³
Generation electricity/biogas	0.75	kWh/m³
Utility		
Total electricity consumption	212.9166666667	kWh/t FTR
Total fresh water consumption	1.13789412582	m³/t FTR
Biogas consumption	1803.455581395	m³/t FTR

Figure 4: Key Performance Indicators derived from Material Flow Network Model

REFERENCES

Department of Industrial Works (ed.): Management Information System (MIS) Guidelines for Eco-efficiency – Palm Oil Industry. Internal Report. Bangkok 2006.

A. Moeller.: Stoffstomnetze – Konzeption eines rechnergestützten ökologischen Rechnungswesens. Diplome thesis, Department of Informatics, Hamburg University, 1995.

A. Moeller: Grundlagen stoffstrombasierter Betrieblicher Umweltinformationssysteme. Bochum 2000.

A. Moeller, A. Rolf, B. Page, B., V. Wohlgemuth,: Foundations and Applications of Computer Based Material Flow Networks for Environmental Management. In: Rautenstrauch, C., Patig, S. (Eds.): Environmental Information Systems in Industry and Public Administration, Hershey, London 2001, S. 379-396.

B. Page, W. Kreutzer: The Java Simulation Handbook – Simulating Discrete Event Systems with UML and Java. Shaker Publ., Aachen, Germany 2005.

B. Page; V. Wohlgemuth: Linking Economic Optimisation and Simulation Models to Environmental Material Flow Networks for Ecoefficiency. In: L.M. Hilty, E.K. Seifert, R. Treibert (eds): Information Systems for Sustainable Development. Idea Group Publ., Hershey, PA USA 2005, pp. 94-108.

D. Pimpisut, B. Page, T. Spehs, V. Wohlgemuth: An Environmental Management Information System for Eco-Efficiency of Agro-Industries in Thailand based on Material Flow Networks. In: O. Hryniewicz, et.al. (Eds.): Proc. 21st Intern. Conf. Informatics for Environmental Protection. Warsaw, Poland, 12-14th Sept. 2007, Shaker Publ., Aachen 2007, pp. 219-227.

A. Perini, A. Susi: Developing a decision support system for integrated production in agriculture. Environmental Modelling & Software 19 (9), pp. 821-829.

M. Raspe: Entwicklung eines Referenzmodells für Stoffstromanalysen am Beispiel der thailändischen Stärke-Industrie, Master Thesis at the Department of Engineering II, University of Applied Sciences, FHTW Berlin, 2007

V. Wohlgemuth, B. Page, W. Kreutzer: Combining discrete Event Simulation and Material Flow Analysis in a Component-based Approach to industrial environmental Protection. Journal Environmental Modelling & Software 21 (2006), pp. 1607-1617.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

The Growing Importance of Open Service Platforms for the Design of Environmental Information Systems

Thomas Usländer, Fraunhofer IITB
thomas.uslaender@iitb.fraunhofer.de

Abstract: Environmental Information Systems (EIS) play a key role in our understanding of the past, current and future status of the environment. In the last 10-15 years, the design of EIS has undergone fundamental changes following both the requirements of the users and the capabilities of the underlying information technologies (IT). This paper first identifies three major trends that have determined these changes: 1) integration of several thematic domains to support interdisciplinary environmental tasks, 2) distribution of EIS to a wider spectrum of users, and 3) permanent functional enrichment with sophisticated functions such as environmental simulations or geo-processing capabilities. It then deduces from these trends the need and the growing importance of open service platforms based on international geospatial standards for the EIS design and its operation. In particular, the open geospatial service architecture of the European research project ORCHESTRA (Open Architecture and Spatial Data Infrastructure for Risk management) is presented in the context of the Web service architectures of the standardisation organisations W3C, OASIS and the Open Geospatial Consortium (OGC). The paper concludes with an overview about ongoing research topics such as the integration of sensor networks, the exploitation of semantic technologies and design methodologies for service-centric computing.

***Keywords**:* Environmental Information System; Open Geospatial Service Platform; Service-Oriented Architecture; Open Geospatial Consortium

1. INTRODUCTION

Environmental Information Systems (**EIS**) play a key role in our understanding of the past, current and future status of the environment. Usually based on large databases that are indirectly (offline) or directly (online) coupled to environmental sensors they allow the user to query and process environmental information and visualize it in thematic maps, diagrams and reports. More advanced functions cover, for instance, the estimation of future values of environmental parameters based on simulation or stochastic models as a basis for decision-support or early warning.

In the last 10-15 years, the design of EIS has undergone fundamental changes following both the requirements of the users and the capabilities of the underlying information technologies (**IT**). This paper first identifies the major trends that have determined these changes, deduces the growing importance of open service platforms for the EIS design and concludes with an overview about ongoing research topics.

On the one hand, the described trends reflect longstanding experiences in the design, deployment and operation of regional and national EIS gained in a German cooperative research network between stakeholders and IT providers [Mayer-Föll et al 2007]. On the other hand, the paper summarises the requirements, IT research topics and latest results of European research projects in the domain of environmental risk management aiming at establishing the basis for a so-called "Shared Information Space for the Environment in Europe (**SISE**)" as requested by the European Commission [Coene and Gasser 2007].

2. TRENDS IN ENVIRONMENTAL INFORMATION SYSTEMS

The following three major trends resulting from the demands of the stakeholders have determined the EIS design in the last years:

1. Domain Integration
2. Wider Distribution
3. Functional Enrichment

2.1 Domain Integration

Domain Integration responded to the demand of enabling the correlation of EIS information and services across various thematic domains, mainly driven by the needs to understand the complex inter-domain relationships in ecological systems. Traditionally, it has been essential for the success and the acceptance of EIS that they are tailored to the thematic domain (e.g. air, water, soil), i.e., that the end user easily recognises in the user interface the notions and names of his/her respective thematic world and the information objects he/she has to deal with. Moreover, the EIS should adhere to the organisational working rules that the end users are used to.

However, thematic information systems had more and more to integrate aspects of other environmental domains in order to handle multi-domain correlation effects. In order to enable a short learning curve, the integrated EIS should follow the same user interaction principles. From the IT perspective, these demands have initially (in the mid and end of the 1990s) resulted in the development of EIS programming frameworks based on mainstream but mostly platform-specific development environments of those days (e.g. Microsoft Active-X, Java/JavaBeans). Non-proprietary middleware technologies such as **CORBA** (Common Object Request broker Architecture) specified by the **OMG** (Object Management Group) could overcome some of the platform-dependencies, but in the end have only been partially successful as standard distributed EIS middleware.

One of the technical obstacles of CORBA has been the tight coupling of client components to remotely accessible objects which led to problems in the robustness of the whole distributed applications. Nevertheless, CORBA laid the basis for thinking in terms of "distributed architectures" based on the notion of "interfaces" specified in a platform-neutral interface definition language. As the middleware battle between Active-X/.NET and CORBA was not decided definitely in the IT market, "Web services" have emerged as the middleware of choice also for distributed Web-based EIS applications crossing thematic and organisational boundaries. The European Commission [2008] follows this trend in their ambition to conceive a "Shared Environmental Information System (**SEIS**)" in Europe. End-user institutions are expected to make investments to "render their existing systems interoperable and link them to an integrated system of systems".

2.2 Wider Distribution

The trend towards wider distribution responded to the demand to open up the EIS to a wider spectrum of users (from employees in environmental agencies, over politicians in ministries up to the citizen) as well as to design the functions of an EIS as callable units from other applications. This trend has been mainly driven by European and national directives such as the European Directive [2003/4/EC] on Public Access to Environmental Information or the Water Framework Directive (**WFD**) [Directive 2000]. At the technical level these directives resulted in the requirements to offer environmental information in a variety of formats and aggregation levels to a multitude of users.

An example is given WFD implementation requirements. The aim of the WFD is to ensure that all European waters, these being groundwater, surface or coastal waters, are protected according to a common standard. As the WFD requires a water management policy centered on natural river basin districts instead of administrative and political regions, the agencies have to co-ordinate their work, possibly across national borders. However, the WFD is not only a fundamental rethink of the EU water policy, its implementation is also a

challenge for the supporting information technology (IT) and, especially, for a WFD-specific information management [Usländer 2005].

Already in the WFD monitoring phase, there is a huge need for harmonisation and possibly standardisation to achieve an efficient implementation of the WFD within Europe. The Common Implementation Strategy established by the European Commission currently foresees three different WFD "products":

- Seamless and harmonised geometric data (e.g. topological consistency)
- Centralised WFD database (1st phase) with a data exchange based on ESRI shapefile format or the Geography Markup Language (**GML**).
- Federation of spatial WFD data servers (2nd phase) based on the standards and recommendations of the Open Geospatial Consortium (**OGC**) and aligned with other pan-European activities on spatial data integration such as **INSPIRE** (Infrastructure for Spatial Information in the European Community).

Today, we are in the transition between the 1st and the 2nd phase: Data is currently exchanged based on **XML** (Extended Markup Language) schemas (either agreed between the EU and the member states or within the member states) whereby investigations about the best architectural choice for the 2nd phase is the subject of European research projects and the implementation of the INSPIRE Directive [2007].

The IT answer to these demands has been essentially influenced and pushed forward by the growing acceptance of the World Wide Web as "computing platform" and not only as information medium. The initial driving forces of "Web Services" have been to better overcome enterprise firewalls and to provide an elegant platform-neutral solution in the middleware battle. However, the acceptance of "service-oriented architectures (**SOA**)" also for the design of EIS resulted from the encouraging perspective in deploying an EIS as a set of re-usable components (services) with well-defined interfaces. These services shall be callable from the Intranet/Internet, thus offering environmental information in an "open" manner to an increasing and heterogeneous user community.

2.3 Functional Enrichment

The trend towards functional enrichment followed the demand to make more sophisticated functions directly available within an EIS, such as environmental simulations or geo-processing capabilities. This helps to lower the purchase, development, user training and maintenance costs. Here, the SOA approach as described above has helped a lot as these functions may also just be loosely-coupled with the EIS as a remote Web service instead of being provided in stand-alone systems, e.g. a Geographic Information System (GIS). Standardised geospatial Web services (e.g. the Web Map Service) as specified by the OGC have captured the generic parts of such remote functions.

The emergence of these IT solutions has created a more challenging vision: an ″ideal″ IT support that would make information available on demand for the end users and would enable the service providers to offer high-quality services at considerably lower cost in a plug-and-play manner. As illustrated in figure 1, EIS applications of various types running in control or information centres, tailored to the analysis of environmental data or to its visualisation in maps or diagrams, have to be coupled with data of various types such as geospatial thematic data, documents or information about environmental phenomena observed by monitoring stations, cameras or satellites – a vision formulated by Denzer [2005] of a functionally rich but generic platform as a need to effectively build environmental decision support systems.

An essential element of such an ideal IT support is an "open geospatial service platform" which provides seamless access to resources (information, services and applications) across organisational, technical, cultural and political borders, thus overcoming real-world heterogeneity and assuring a sustainable investment for the support of future, yet unknown requirements. "Open" hereby means that service specifications are published and made

freely available to interested vendors and users with a view to widespread adoption. Furthermore, an open service platform makes use of existing standards (e.g. ISO and OGC) where appropriate and otherwise contributes to the evolution of relevant new standards.

3. OPEN GEOSPATIAL SERVICE PLATFORM

3.1 The ORCHESTRA Architecture

Based on a systematic analysis of system requirements, the European research project **ORCHESTRA** (Open Architecture and Spatial Data Infrastructure for Risk management) has specified and implemented a reference model and a series of architecture services that provide the generic and platform-neutral functional grounding of such open geospatial service platforms [Klopfer and Kannellopoulos (eds.) 2008]. This Reference Model for the ORCHESTRA Architecture (**RM-OA**) has been accepted as a best-practices document for architectural design by the OGC [Usländer 2007a]. Figure 1 shows the abstract principle how a geospatial service platform may couple applications and data sources.

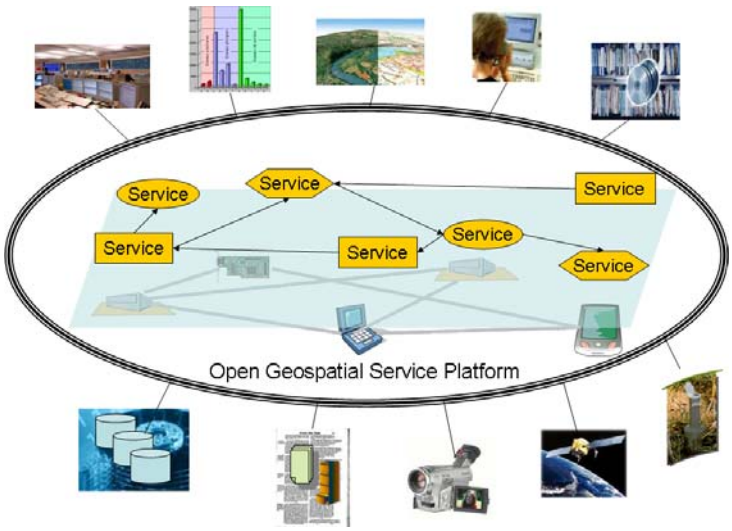


Figure 1. Open Geospatial Service Platform

The RM-OA is built upon two main pillars: a process model and a conceptual model. The ORCHESTRA process model applies an incremental, iterative approach for the analysis and design phases. Usually, a multi-step breakdown process across several abstraction layers is necessary to analyze user requirements and map them to the capabilities of a service platform. As illustrated in figure 2, ORCHESTRA distinguishes between an abstract service platform that is specified independently of a given middleware technology and a concrete service platform.

The abstract design phase leads to platform-neutral specifications following the rules defined in the RM-OA for the abstract service platform. They represent the functional, informational and non-functional requirements of the problem domain. The concrete design phase maps the abstract specifications to a chosen concrete service platform. Currently, this is the ORCHESTRA Web Services platform based on the Web Service Description Language (WSDL) bound to the SOAP protocol, and a GML profile as the current mainstream service platform technologies for geospatial applications. In the engineering phase the platform-specific components are organized into service networks taking into account the qualitative requirements and translating them into operational policies.

In practice, these individual phases are often interlinked and repeated in an iterative fashion. Sometimes the abstract design phase is not required. Furthermore, existing service

and OGC service standards for Web services make a pure top-down approach impractical. Thus, in practice, a middle-out design approach is often the appropriate method.

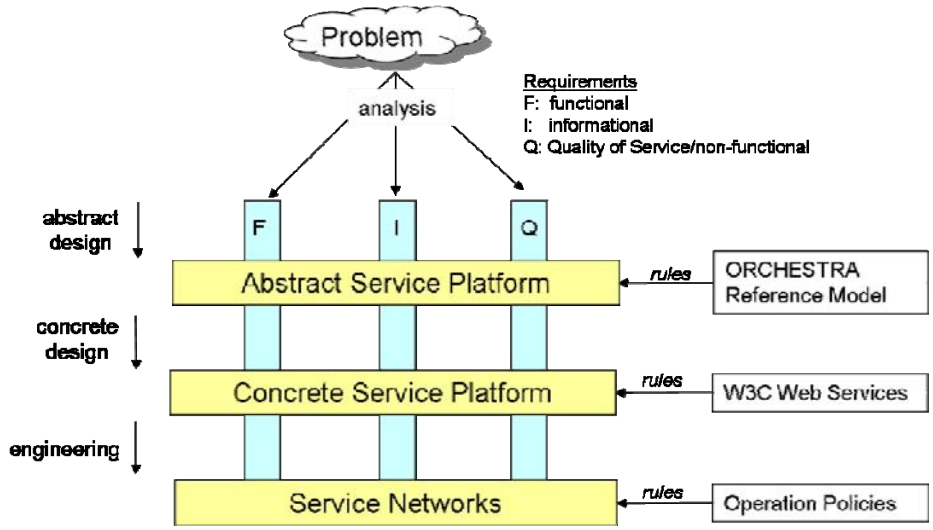


Figure 2. Abstract and Concrete Service Platforms

However, what is required is a clear structure for the documentation of the ideas and the results of the design phases. ORCHESTRA has adopted the ISO Reference Model for Open Distributed Processing [1998] for this task and has interpreted it for the design of geospatial service architectures. One major step forward in the RM-OA is a uniform meta-model for both the information and the service models that extends and complements the ISO 19109 General Feature Model [2005]. The key concept in the uniform meta-model is to put in the forefront the "interface type" as reusable specification unit. This allows the service designer to combine interfaces to new service types according to the functional requirements of the user. An example is a diagram rendering interface as an extension to a map service to enable visualisation of information in both map and diagram layers.

3.2 ORCHESTRA Architecture Services

Table 1 lists the major architectural interface and service types that have been specified in the ORCHESTRA Architecture on the abstract as well as on the concrete service platform level. This list of generic services has been systematically derived from system and user requirements in the thematic domain of risk management. However, their scope is intended to be independent of the application domain. Note that these services types consist themselves of a set of interface types that may also be attached to other service types.

4. ONGOING RESEARCH ACTIVITIES

There are ongoing research activities for service-centric distributed applications which are of high relevance to the future design of open service platforms for EIS. Some of them are highlighted in the following sections.

4.1 Semantic Web Services

There is ongoing research work in the field of semantic extensions of the Web (Semantic Web) which has already led to a series of basic recommendations of the World Wide Web Consortium (**W3C**) such as the Resource Description Framework (**RDF**) as a general method of modelling information as statements about resources in the form of subject-predicate-object expressions, and **OWL**, the W3C Web Ontology Language to define and instantiate ontologies.

Table 1. List of Major ORCHESTRA Architecture Services and Interfaces

Service and Interface Type Name	Application
Basic Interface Types	Enable a common architectural approach for all ORCHESTRA Services, e.g. for the self-description of service instances (capabilities), synchronous and asynchronous interactions, transactional support and access to knowledge basis.
Annotation Service	Generates specific meta-information from textual sources and relates it to elements of an ontology (e.g. concepts, properties, instances).
Authentication Service	Proves the genuineness of principals (i.e. the identity of a subject which may be a user or a software component) using a set of given credentials.
Authorisation Service	Provides an authorisation decision for a given context.
Catalogue Service	Ability to publish, query and retrieve descriptive information (meta-information) for resources of any type. Extends the OGC Catalogue Service by additional interfaces for catalogue cascade management and ontology-based query expansion.
Coordinate Operation Service	Changes coordinates on features from one coordinate reference system to another.
Document Access Service	Access to documents of any type (e.g. textual documents and images).
Feature Access Service	Selection, creation, update and deletion of feature instances and feature types available in a service network. Corresponds to the OGC Web Feature Service but is extensible by a schema mapping interface.
Map and Diagram Service	Enables geographic clients to interactively visualise geographic and statistical data in maps (such as the OGC Web Map Service) or diagrams.
Name Service	Encapsulates the implemented naming policy for service instances in a service network
Ontology Access Service	Supports the storage, retrieval, and deletion of ontologies as well as providing a high-level view on ontologies.
Service Chain Access Service	supports the creation of an executable service instance based on an explicit description of a service chain.
Service Monitoring Service	Provides an overview about service instances currently registered within service network incl. status and load.
User Management Service	Creates and maintains subjects (users or software components) including groups (of principals) as a special kind of subjects.

Work on semantic extension of Web Services (Semantic Web Services) is carried out in various research projects and is currently reflected in various competing submissions to the W3C, e.g. **WSMO** (Web Service Modelling Ontology) and **OWL-S** (OWL-based Web service ontology). It is too early to anticipate which of these technologies and frameworks will be dominating the market in the future as their practical relevance has only rarely been validated in real-world applications. However, the use of ontologies and related Semantic Web Services is one of the promising approaches to match between heterogeneous environmental data schemas and service types.

4.2 Integration of Sensor Networks

The European research project **SANY** (Sensors Anywhere) extends the ORCHESTRA architecture by the inclusion of sensors and sensor networks of various topologies [Havlik et al 2007]. The resulting Sensor Service Architecture is based on the OGC Observation and Measurement model [Cox 2007] and includes the service types of the OGC Sensor Web Enablement working group (see table 2). SANY organizes the service spectrum in a hierarchy of functional domains as illustrated in figure 3. Whereas the Sensor Domain is

dominated by proprietary solutions, a high percentage of the required functionality is covered by the application and tailoring of the generic ORCHESTRA and SANY service types.

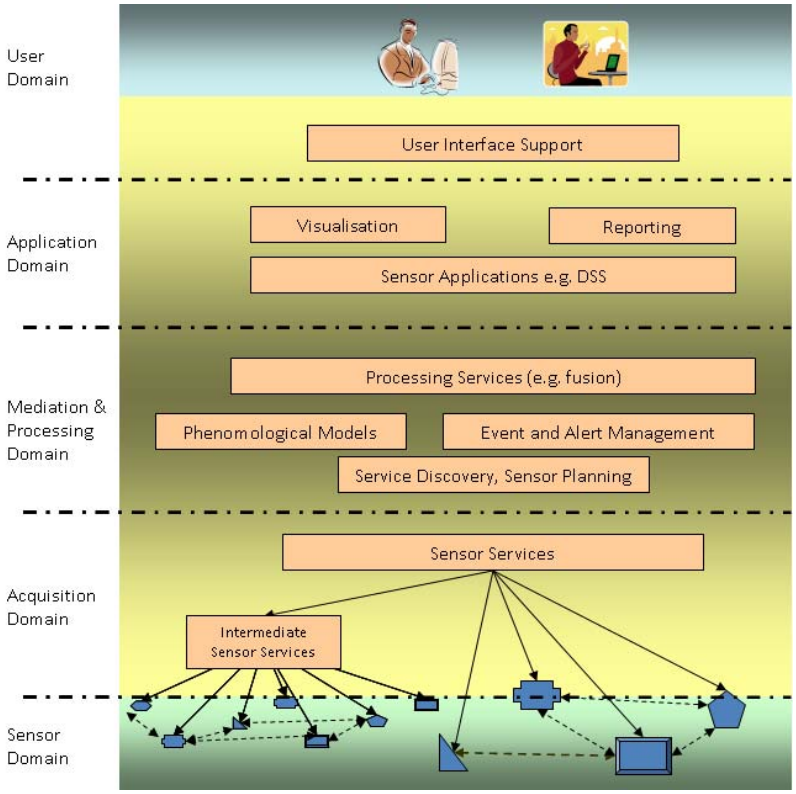


Figure 3. Functional Domains of the SANY Sensor Service Architecture

Table 2. List of Service Types of the OGC Sensor Web Enablement Group

Service and Interface Type Name	Application
Sensor Observation Service	Provides access to observations from sensors and sensor systems in a standard way that is consistent for all sensor systems including remote, in-situ, fixed and mobile sensors.
Sensor Alert Service	Provides a means to register for and to receive sensor alert messages.
Sensor Planning Service	Provides a standard interface to task any kind of sensor to retrieve collection assets.
Web Notification Service	Service by which a client may conduct asynchronous dialogues (message interchanges) with one or more other services.

4.2 Web Service Paradigms

Looking at the technical foundation for the open service platforms, there is an ongoing discussion about the basic Web service paradigm to be used. The OGC services have been designed before the W3C standards SOAP and WSDL have been accepted. Thus, they are still specified with an http/**KVP** (key-value pair) binding. Although the OGC Technical Committee has decided in June 2006 to provide additional WSDL/SOAP bindings for the OGC service interfaces, they have not yet reached the status of accepted OGC standards. Furthermore, the mass-market in the Geospatial Web tends towards another paradigm, the RESTful web services [Richardson and Ruby 2007]. RESTful web services aim at accessing and manipulating uniquely identified resources based on a uniform interface with

commonly agreed, well-defined semantics. It is usually discussed together with its realization in a Web environment i.e., based on the basic technologies of the World Wide Web.

Following the RM-OA, the SANY project understands a Resource-Oriented Architecture (ROA) as a platform-independent concept to be integrated into the service meta-model [Usländer 2007b]. At the level of the concrete service platform, there will be basic rules (e.g. to use http) and defined but conditional extensions (e.g. to use WSDL/SOAP or to provide RESTful Web services) depending on the needs of the environmental applications.

5. CONCLUSIONS AND RECOMMENDATIONS

Technologies for service platforms have continuously changed in the past and are expected to change in the future. Nevertheless, an "open" approach with functionally rich service platforms based on standards is indispensable when aiming at providing interoperable solutions. The ORCHESTRA and the SANY projects try to provide robust and sustainable architectures that may cope with technological changes. Semantic technologies will emerge but have to be integrated step-by-step into the existing open geospatial service platforms.

REFERENCES

Coene, Y. and R. Gasser, Joint Operability Workshop Report "Towards a single information space for Environment in Europe", Frascati, ftp://ftp.cordis.europa.eu/pub/ist/docs/environment/sise-workshop-report_en.pdf, 2007.

Cox, S, Observation and Measurements – Part 1: Observation Schema. OGC Document 07-022r1, OpenGIS Specification, 2007.

Denzer, R., Generic Integration of Environmental Decision Support Systems – State-of-the-Art, Elsevier Journal Environmental Modelling and Software 20, pp. 1217-1223, 2005.

Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for the Community action in the field of water policy, 2000.

Directive 2003/4/EC of the European Parliament and of the Council on public access to environmental information and repealing Council Directive 90/313/EEC, 2003.

Directive 2007/2/EC of the European Parliament and of the Council establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), 2007.

European Commission, Communication from the Commission to the Council, the European Parliament, The European Economic and Social Committee and the Committee of the Regions - Towards a Shared Environmental Information System (SEIS), COM (2008) 46 final, 2008.

Havlik, D., Schimak, G., Denzer, R. and Stevenot, B, SANY (Sensors Anywhere) Integrated Project. In: D. A. Swayne and J. Hrebicek, [Ed.], *Proceedings of the ISESS 2007*, 22-25 May 2007, Prague, Czech Republic.

ISO/IEC 10746 Reference Model for Open Distributed Processing, 1998.

ISO 19109 Geographic information -- Rules for application schema, 2005.

Klopfer, M. and I. Kannellopoulos (eds.), ORCHESTRA – An Open Service Architecture for Risk Management, http://www.eu-orchestra.org/docs/ORCHESTRA-Book.pdf, ISBN 978-3-00-024284-7, 2008.

Mayer-Föll, R. and W. Geiger (eds.), Wissenschaftliche Berichte Forschungszentrum Karlsruhe FZKA 7350, UIS Baden-Württemberg, F + E-Vorhaben KEWA Phase II 2006/2007, http://www2.lubw.baden-wuerttemberg.de/public/uis/kewa1, 2007.

Richardson, L. and S. Ruby, RESTful Web Services. O'Reilly Media, Inc. ISBN-10: 0-596-52926-0, 2007.

Usländer, T., 2005. Trends of environmental information systems in the context of the European Water Framework Directive. *Elsevier Journal Environmental Modelling & Software 20*, 1532-1542, 2005.

Usländer, T. (ed.), Reference Model for the ORCHESTRA Architecture Version 2. OGC Best Practices Document 07-097, http://portal.opengeospatial.org, 2007a.

Usländer, T., Integration of Resource-Oriented Architecture Concepts into the OGC Reference Model. OGC Document 07-156, http://portal.opengeospatial.org, 2007b.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Environmental Management Information Systems – Illustrations from Online Communication and Sustainability Reporting

R. Isenmann

Institute for Project Management and Innovation (IPMI), and Research Center for Sustainability Studies (artec), University of Bremen. Wilhelm-Herbst-Straße 12, D-28359 Bremen (Germany), (ralf.isenmann@innovation.uni-bremen.de)

Abstract: Environmental Management Information Systems (EMIS) are socio-technological systems used as business applications to gather, process, and provide environmental information, inside companies and in exchange with other actors in industry. They help to identify environmental impacts and support measures avoiding these impacts or reducing them. EMIS provide the necessary information support for decision making in companies. Hence, EMIS can be viewed as certain Information Systems (IS) usually implemented in companies as a part of their Environmental Management Systems (EMS). In order to give a tangible example with practical implications, the developments that EMIS have passed the last years are described along the field of "online communication and sustainability reporting" and illustrated by a case study. This area represents an emerging digital and fully ICT (information and communication technologies) supported approach within EMIS, using current internet technologies and services. It makes clear the array of capabilities of latest EMIS to be exploited for the improvement of advanced environmental and sustainability management, finally to the benefit for companies and their various stakeholders. The case study describes the concept and implementation of a software tool with shopping cart functionality providing sustainability reports à la carte so that stakeholders (i.e. users, readers) can create their own report on the fly, exactly meeting their detailed information needs and preferred media out from a single publishing database. This software tool which represents a module of a comprehensive EMIS is implemented as a web-based ICT application. Its performance goes beyond the leading-edge approach of O_2 who provides a personalized reporting feature on its website that could be regarded as best practice and pioneering effort in sustainability online reporting, so far.

***Keywords*:** Environmental Management Information System; Online Communication, Sustainability Reporting

1. INTRODUCTION

Environmental Management Information Systems (EMIS) are socio-technological systems used as business applications with the aim to gather, process, and provide environmental information, inside companies and in exchange with other actors in industry. Hence, EMIS can be viewed as certain Information Systems (IS) usually implemented in companies as a part of their Environmental Management Systems (EMS). Perhaps the most known and famous examples of EMS are the world-wide standard ISO 14000 series (ISO 2008) and the European Union's Eco-Management and Audit Scheme (EMAS) (European Commission 2008). These EMS are applicable for companies and other organizations to evaluate, report and improve their environmental performance. For example, EMAS has been available for participation by companies since 1995 and was originally restricted to companies in industrial sectors. Since 2001 it has been open to all economic sectors including public and private services. Despite some differences between ISO 14000 and EMAS in detail, the central objective is the same, i.e. to improve environmental performance on a continuous basis.

Since its launch more than two decades ago, the field that we call today as "Environmental Management Information Systems" (EMIS) has rapidly evolved (Donley et al. 1998). At first, Environmental Information Systems (EIS) were used primarily in the public sector, e.g. (i) to rise public awareness of the condition of public goods like the atmosphere or recreation areas, (ii) to support political decisions that have effects on nature by tools for modeling and simulation, (iii) and to help executing environmental policy e.g. by monitoring (Page and Hilty 1995). Several years ago, however, a growing market has come about for IS that support EMS. For sure, early pioneers in EMS and few sector-leaders in eco-sensitive industries have had in place some forms of EIS before. But at least since the late 1970s, EIS in business applications became a topic of broader interest in academia, industry, and administration, and – as a result – more widespread and finally an issue in its own right (Rautenstrauch and Patig 2001). The development path that EMIS have passed could be described as a movement away from standalone IS and single tools focused on certain environmental aspects like legal compliance or material input and energy consumption. These rather isolated ICT solutions mostly served basic functions in procurement and production, and hence they could provide solely limited managerial support.

Today, comprehensive EMIS are available that help to assess even a firm's integrated sustainability performance (i.e. environmental, social, economic, and mutual interrelations). These sophisticated forms of EMIS are based on flexible ICT architectures, often using current internet technologies and services. They are embedded in corporate-wide ICT networks combining business processes and organizational units. Further, interfaces facilitate information exchange with external stakeholders and enable continuous updates from data sources, inside and outside the companies. Such leading-edge EMIS offer a number of opportunities to support decision making, finally almost able to assess companies' activities in physical and monetary terms: (semi-)automated, in real time, online, and in a tailored fashion (Hilty et al. 2005). Currently, using EMIS is part of companies' daily affairs, and has entered the business mainstream. Due to technological progress in ICT applications and improvements in computer science, the question should be now to choose the right EMIS that meets current needs and future requirements, but no longer whether to use EMIS at all. The currently limiting factor may not be of technological nature based in ICT but rather organizational, rooted in corporate management and business strategy (Perl 2006).

2. ENVIRONMENTAL MANAGEMENT INFORMATION SYSTEMS

From an academic perspective, research on EMIS is located at the intersection between Environmental Informatics and Business Information Systems (fig. 1).

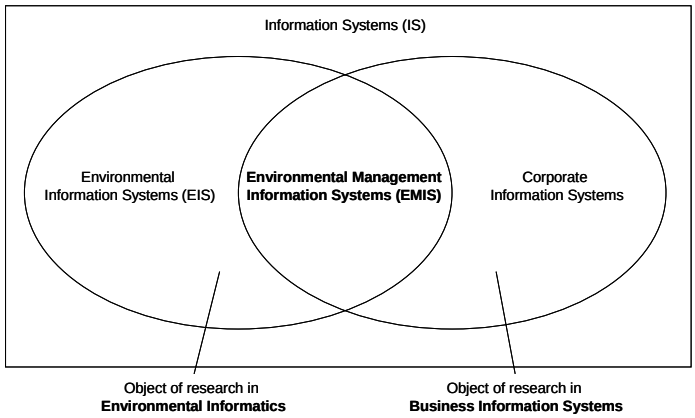


Figure 1: Classification of EMIS (Hilty and Rautenstrauch 1995, 295)

While the object of research in Business Information Systems is the design of IS primarily for companies serving industrial purposes, often with a commercial intention, the focus in Environmental Informatics is in the development and application of EIS in general, be it implemented in the public or in the private sector. From a systematic perspective, Environmental Informatics represents a goal-driven field of research, whereas the field of Business Information Systems looks like

an entity-based one with an industrial focus in business applications at firm's level. Regarding to these two fields of research and their specific intellectual heritage, it becomes clear that EMIS relate to both, Environmental Informatics as well as to Business Information Systems. On the one hand, EMIS are part of Environmental Informatics due to the environmental focus; on the other hand, EMIS also belong to Business Information Systems as EMIS are developed and designed for companies and implemented as business applications.

The central purpose of EMIS is to support EMS. EMS themselves prototypically fulfill a number of basic tasks, i.e.: analysis, plan, implementation and controlling, and check. These tasks are laying at the heart of any EMS, even the standardized ones like the European EMAS and the word-wide applicable EMS according to the ISO 14000 series. Any basic task in EMS corresponds to certain functions that EMIS may support (fig. 2).

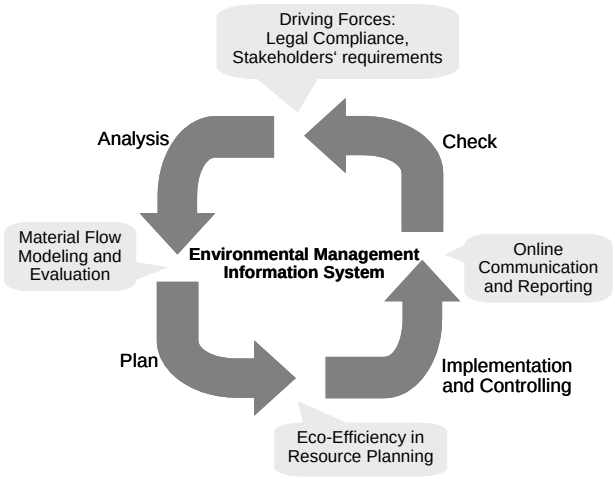


Figure 2: Process model of EMIS, illustration

Legal compliance and other stakeholders' requirements are the driving forces that push companies to improvements so that they may do more than smart "green-washing". First, EMIS help to comply with laws and regulations. They provide insights into complex legal frameworks, e.g. on the information retrieval networks for environmental regulations. Further, they assist decision makers in discovering business activities that are of relevance for environmental regulations, e.g. by emission monitoring. Second, EMIS also enable to get feedback from their target groups and to capture requirements from a larger audience. Such feedback loops are usually installed in the form of stakeholder relations management systems. Next, EMIS support material flow management, including material flow modeling and evaluation. For example, EMIS provide help when identifying the material and energy inputs, collecting the relevant data, and preparing an eco-balance. Based upon insights drawn from the eco-balance, EMIS further assist the impact assessment and balance valuation in order to assess environmental impacts due to material and energy flows. Once material flows are modeled and evaluated, EMIS help to improve eco-efficiency. The measures that could be taken range from product life cycles, production processes, to whole companies. These certain entities can be viewed as material flow systems and being optimized in terms of eco-efficiency. Therefore, a number of elaborated instruments offer methodological support, like Life Cycle Analysis (LCA) and Design for the Environment (DFE), just to mention a few. When the measures for improvement in EMS have been implemented and controlled according to eco-efficiency, then all activities and results need to be documented and communicated. Here EMIS provide an array of benefits to improve communication and reporting both, for companies and their stakeholders. For example, EMIS can overcome the restrictions of freestanding documents with limited contents, print media fixation, disclosure in a one size fits all fashion, and one-way-communication. Together, the basic tasks in EMS and corresponding EMIS functions are building a closed loop procedure, finally leading to a process model of continuous environmental improvements (fig. 2).

Ideally, the different EMIS functions are combined to a coherent system and fully integrated into companies' IS, while supporting all the tasks throughout an EMS (Rautenstrauch 1999). Despite progression companies have made in recent years and concepts of integration that have been proposed, however, EMIS are usually not incorporated elements, today (Isenmann and Rautenstrauch 2007). Particularly they are not part of Enterprise Resource Planning (ERP) systems. Such an integration has just partly taken place. For example, in my SAP® PLM (Product Lifecycle Management), there is a module on Environment, Health and Safety (EH&S). Its focus though is limited to the documentation of waste management following national laws for handling hazardous goods (SAP 2005). In order to give a tangible example with practical implications, the developments in EMIS are described along the field of "online communication and sustainability reporting" (Isenmann and Bey 2007; Isenmann et al. 2007) and illustrated by insights from a case study.

3. ILLUSTRATIONS FROM ONLINE COMMUNICATION AND SUSTAINABILITY REPORTING

Corporate sustainability reporting is a balanced reporting approach through which companies are communicating the three pillars of environmental, social, and economic performance and its mutual interrelations to various stakeholders. Hence, sustainability reporting can be viewed as the "front end" in the process model of EMIS (fig. 2), communicating activities and performance, inside and outside the companies. Further, it is a driver that stimulates companies to continuous environmental improvements when passing the circle again. At present, the vast majority of advanced environmental and sustainability reports are uniformed documents, usually prepared as one size fits all reports, still. This is even the case for online available reports.

4 CASE STUDY: SUSTAINABILITY REPORTING À LA CARTE – BEYOND BEST PRACTICE BY O_2

The case study describes the concept and implementation of a software tool with shopping cart functionality providing sustainability reports à la carte. Stakeholders (i.e. users, readers) can create their own tailored report in an interactive manner on the fly, while exactly meeting their detailed information needs and preferred media. This software tool is a module of a comprehensive EMIS. It is implemented as a web-based ICT application.

4.1 Challenges

The provision of reports on various media, and tailoring reports according to users' needs and preferences, while exactly meeting numerous standards, guidelines, and recommendations, are two major difficulties with which companies are struggling at present.

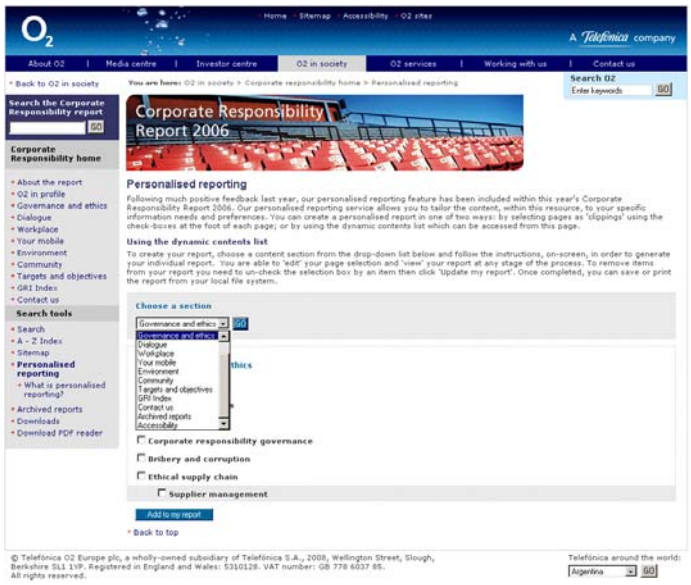


Figure 3: O_2's reporting feature (2008)

Today, an orthodox disclosure practice which merely provides isolated documents and stand-alone printed reports is not sufficient any longer. A substantial amount of information, matters of communication style, and the provision of tailored communication vehicles on various media are required (Isenmann and Kim 2006). O_2, a global player in telecommunications, provides a personalized sustainability reporting feature on its website (fig. 3). This could be regarded as best practice and pioneering effort, so far: Users can tailor the content by selecting issues from a catalogue, and they can choose the preferred media: HTML to view the report or PDF to print and save the report.

4.2 Concept

In order to go beyond this approach in tailored online reporting, a software tool with an elaborated shopping cart functionality was developed (see fig. 4: moving from cell I to cell IV). The conceptual basis is a classification proposed by Lenz (2003): He arranged the different strategies to fulfill target group tailoring along two dimensions (fig. 4): (i) degree of user modeling, representing the overall model of a user or a certain user group, and corresponding to information needs identified and preferences stored in user profiles: stereotyping, individualization, and personalization; (ii) degree of system adaptation, representing the capabilities to adapt the ICT application to users' needs, and corresponding to the extent of a system's facilities for tailoring: adapted, adaptable and adaptive applications. Following a generic information management approach (Isenmann and Marx Gómez 2004), three further tasks had to be carried out: (i) modeling the information demand, i.e. conducting a stakeholder analysis and information requirement analysis, (ii) modeling the information supply, i.e. designing suitable XML-based documents (schema, style sheets, document), and (iii) cross matching demand and supply, i.e. designing the software tool.

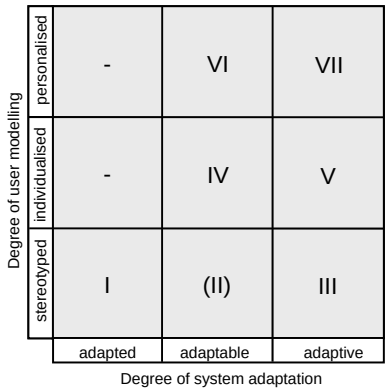


Figure 4: Classification of target group tailoring (Lenz 2003, 212)

4.3 Implementation

The software tool is implemented as a prototype. In order to showcase the performance and features of this tool compared to current best practice, the content of a digitalized sustainability report of Otto (2004), a German multi-channel retailer, was used. This report was already available as an XML document. At the heart of the software tool's ICT architecture lies Apache Cocoon (2008), a Java-based, modular-structured open source publishing framework, able to process XML schemas and to transform and format XML documents. It is thus suitable to perform single-source cross-media reporting. Sustainability reports à la carte are made possible, prepared by machine processing, and generated in an automated manner. The ICT architecture allows report contents to be stored, retrieved, edited, updated, controlled and then output cross-media in a variety of ways. The reason why Apache Cocoon has been employed lies, among other benefits, in its sophisticated application logic. The modular components could be arranged in a flexible way, serially grouped in pipelines where different reports are then dynamically created on the basis of an underlying XML-based schema (fig. 5).

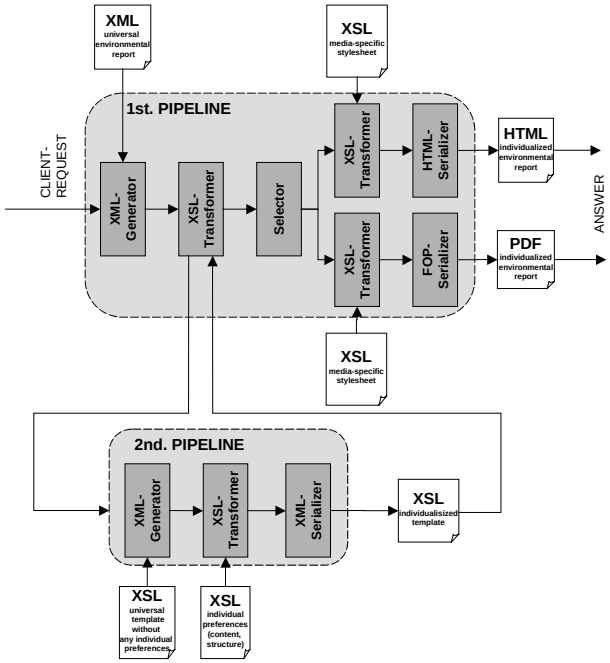


Figure 5: Apache Cocoon multiple pipelines to provide individualised reports (cell IV in fig. 4)

4.4 Sustainability reporting à la carte in practice

The procedure of using the software tool is like the following: First, users decide whether they may receive a predefined report corresponding to the information needs prototypical to a certain target group, or whether they may create their own report, exactly meeting their individual information needs (fig. 6). The software tool provides a catalogue of predefined report contents e.g. for investors, customers, employees, media representatives, suppliers, and non-governmental groups. Further, the catalogue also covers profiles exactly meeting the requirements of EMAS and of the GRI guideline (GRI 2006), i.e. the de-facto-standard in sustainability reporting. For example, some users such as local authorities are interested in how the company has met the EMAS requirements in order to aggregate environmental impacts for a certain area.

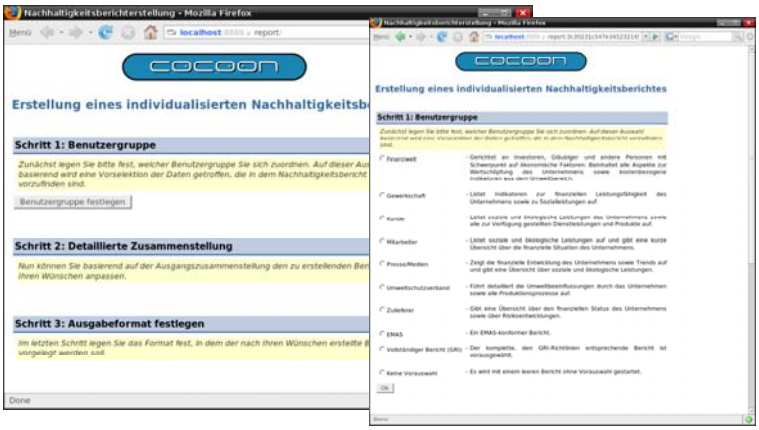


Figure 6: Software tool for sustainability reporting à la carte, illustration

Second, users can fine-tune in a flexible and easy to use manner what contents should finally be included in their shopping cart (fig. 7).

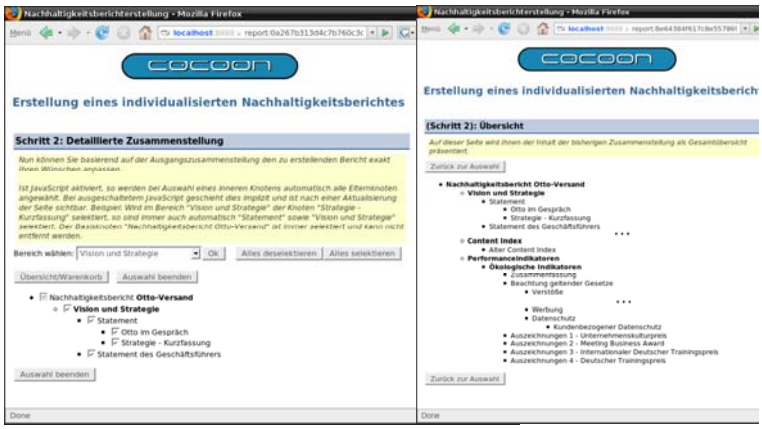


Figure 7: Fine-tuning the contents in the shopping cart, illustration

Third, users can choose different output formats (fig. 8): HTML to view reports in the web browser, PDF and Postscript to save and print reports, and – more important – XML to process data automatically by machinery. Among other aspects needed to link the micro-macro-perspective between single companies and their aggregation to a regional level (Seifert 2000; Rathje 2001), it is the progression in ICT that has now opened the window to better exchange environmental information among companies and within the industrial sector. No less important, the interoperability with public agencies could be improved.

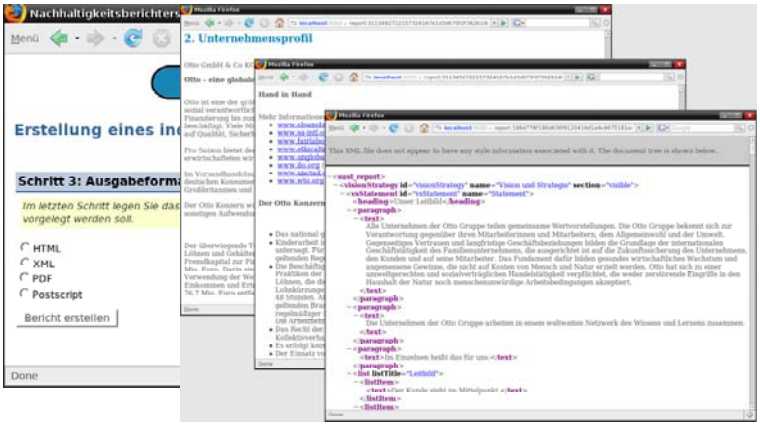


Figure 8: Selecting different output formats in the shopping cart, illustration

In contrast to the progression O_2 and other companies – even the award winning ones – have made the last years, however, current sustainability (online) reporting practice shows significant room for improvements, particularly in interactivity and target group tailoring. The concept and implementation of a software tool with shopping cart functionality shed light on how advanced environmental and sustainability reporting as an integral part of EMIS could be further developed, especially in terms of interactivity and target group tailoring.

5. CONCLUSIONS

The example that has been presented from the field of "online communication and sustainability reporting" is merely a part of a comprehensive EMIS. The whole EMIS includes a number of modules. For example, another module supports the continuous workflow throughout the EMIS process model (see fig. 2), ranging from (i) the extraction of environmental data out from ERP (enterprise resource planning) systems, (ii) the modeling of the material flow network, (iii) eco-balancing, (iv) impact assessment and (v) balance valuation performed by the software tool "Umberto", through (v) the production of various communication vehicles (Marx Gómez and Rautenstrauch 2001). In total, the whole EMIS allows report contents to be created (semi-)automatically (i.e. the quanti-

tative data), stored, retrieved, edited, updated, controlled and then output cross-media in a variety of ways (the single-source cross-media principle).

Of course, communication and reporting represent just one task that has to be managed within an EMIS. However, the example of the software tool presented here makes clear the array of capabilities of latest EMIS to be exploited for the improvement of advanced environmental and sustainability management, finally to the benefit both, for companies and their various stakeholders, even for public agencies. The case study may particularly highlight that leading-edge EMIS with a module for online communication and (sustainability) reporting could elevate the orthodox "one size fits all" disclosure practice of environmental reports to a sophisticated digital stage. Reports could be developed from being universal documents on print media to valuable communication vehicles which provide substantial and reliable information in a tailored fashion and are available on various media – due to an underlying single source cross-media principle. For example, following the EMAS reporting requirements, then the reports also may guarantee the reliability of the information provided (European Communities 2001).

Numerous stakeholder, be they private, public, or industrial, are no longer satisfied solely with environmental reports and other communication instruments on print media or mere electronic duplicates. Advanced environmental (and sustainability) reporting is becoming increasingly relevant both, for decision-making in industry and for public agencies. Responding to multiple inquiries that a variety of stakeholders are making to companies is time-consuming and costly (Axelrod 2000). Rather than endure these procedures, companies are recognising the value of having a readily available EMIS to provide the data which is needed.

Pioneering companies have started, or in the near future will start to implement, suitable applications. For example, Verie Sandborg, Baxter International's manager of EH&S requirements, regards a good environmental or sustainability report as an excellent source for responding to formalised requests for environmental or sustainability information (Axelrod 2000). Many of the questions asked are already answered in meaningful reports. Hence, it would be helpful to have proper EMIS with software tools in place supported through an efficient information management. Users could extract the information they need from a publishing database, and create an automatically generated tailored report themselves, i.e. users could generate their own "reports à la carte" simply by selecting keywords, clicking on preferences on a menu, or choosing a certain guideline – perhaps creating an environmental statement according to EMAS, or having a comprehensive sustainability report at one's fingertips. Especially an environmental statement (a report meeting the EMAS requirements) could be an excellent source and is therefore regarded as a core element of a new corporate performance evaluation system (Perrini and Tencati 2006) including an integrated approach such as sustainability reporting (GFEM and GFEA 2006):

- First, environmental statements provide a "true and fair view" (a reporting principle which is borrowed from financial reporting) as they guarantee the reliability of the information provided. Environmental statements are independently verified, which for sustainability reporting is still an open question.
- Second, environmental statements include integrated performance indicators such as eco-efficiency. Such quantitative data are crucial to uncover and highlight the mutual interrelations between environmental, social, and financial aspects. They are essential to make the integrated performance transparent, and help to conceptualise the "triple bottom line" (Elkington 1997), i.e. the core theme for corporate sustainability.

For example, a German company in the pharmaceutical industry (Weleda 2003) and a German public utility and transportation service (Heidelberger Versorgungs- und Verkehrsbetriebe GmbH 2005) are adopting this strategy. In so doing, they have integrated a validated environmental statement into their sustainability report.

REFERENCES

Axelrod, R.A., Brave new words: the financial value of environmental communications, *Environmental Quality Management* 9(4), 1-11, 2000.

Donley, M., Donley, C.M., Donley, B.J., and Donley, J.W., *Environmental management information systems report*. Colonial Beach: 3rd ed. Donley Technology, 1998.

Elkington, J., *Cannibals with forks. The triple bottom line of 21st century business.* Oxford: Capstone, 1997.

European Commission, *EMAS The Eco-Management and Audit Scheme*. http://ec.europa. eu/environment/emas/index_en.htm, 2008-02-19.

German Federal Environmental Ministry (GFEM), German Federal Environmental Agency (GFEA), *Guidance document for voluntarily developing an EMAS environmental statement into a sustainability report.* Berlin: GFEM, 2006.

Global Reporting Initiative (GRI), *Sustainability reporting guidelines. Version 3.0.* Amsterdam: GRI.

Heidelberger Versorgungs- und Verkehrbetriebe, *Verantwortung im Wettbewerb. Nachhaltigkeitsbericht und Umwelterklärung*. Heidelberg: HVV (available only in German), 2005.

Hilty, L.M., and Rautenstrauch, C., Betriebliche Umweltinformatik. *Umweltinformatik. Informatikmethoden für Umweltschutz und Umweltforschung*. Page, B.; Hilty, L.M. (Eds.). München: 2nd ed. Oldenbourg, 295-312, 1995.

Hilty, L.M., Seifert, E.K., and Treibert, R. (Eds), *Information systems for sustainable development*. Hershey: Idea Group Publishing, 2005.

Isenmann, R., and Bey, C., Environmental reporting on the internet – from a technical tool to a strategic necessity, *Strategic sustainability: The State of the art in corporate environmental management systems*, Sroufe, R., Sarkis, J. (Eds.). Sheffield: Greenleaf, 123-143, 2007.

Isenmann, R., and Kim, K., Interactive sustainability reporting. Developing clear target group tailoring and stimulating stakeholder dialogue, *Sustainability accounting and reporting*. Schaltegger, S., Bennett, M., Burritt, R. (Eds.). Berlin: Springer, 533-555, 2006.

Isenmann, R., Bey, C., and Welter, M., Online reporting for sustainability issues, *Business Strategy and the Environment, 16(7)*, 487-501, 2007.

Isenmann, R., and Marx Gómez, J., How to provide customized environmental reports properly, *Environmental online communication*. Scharl, A. (Ed.). London: Springer, 173-182, 2004.

Isenmann, R., and Rautenstrauch, C., Horizontale und vertikale Integration Betrieblicher Umweltinformationssysteme (BUIS) in Betriebswirtschaftliche Anwendungsszenarien, *Umweltwirtschaftsforum, 15(2)*, 75-81, 2007.

International Organization for Standardization (ISO), *ISO 14000 essentials*. http://www.iso. org/iso/iso_14000_essentials, 2008-02-19.

Lenz, C., *Empfängerorientierte Unternehmenskommunikation. Einsatz der Internet-Technologie am Beispiel der Umweltberichterstattung.* Lohmar: Eul 2003.

Marx-Gómez, J., and Rautenstrauch, C., Material based eco-balancing and evaluation of a beer production process. *Environmental Management & Health* 12(2/3): 286-300. 2001

O_2, *Corporate responsibility report 2006*. http://www.o2.com/cr/resource2006/personalisedreporting.asp, access 2008-02-19.

Page, B., and Hilty, L.M. (Eds), *Umweltinformatik. Informatikmethoden für Umweltschutz und Umweltforschung.* München: 2nd ed. Oldenburg, 1995.

Perrini, F., and Tencati, A., Sustainability and stakeholder management: the need for new corporate performance evaluation and reporting systems. *Business Strategy and the Environment* 15, 296-308, 2006.

Perl, E., *Implementierung von Umweltinformationssystemen. Industrieller Umweltschutz und die Komunikation von Umweltinformationen in Unternehmen und Netzwerken.* Wiesbaden: Gabler, 2006.

RATHJE, B., *Der Mikro-Makro-Link in der Umweltberichterstattung. Möglichkeiten und Grenzen der Verknüpfung einzel- und gesamtwirtschaftlicher Umweltberichterstattungssysteme*. Dissertation Thesis Universität Kassel, 2001

Rautenstrauch, C., *Betriebliche Umweltinformationssysteme.* Berlin et al.: Springer 1999.

Rautenstrauch, C., and Patig, S. (Eds) *Environmental information Systems in Industry and administration*. Hershey: Idea Group Publishing, 2001.

SAP, *Integrated industrial hygiene and safety with SAP® Environment, Health & Safety*. Walldorf: SAP 2005.

Seifert, E.K., Mikro-Makro-Link (MML), *Informationssysteme für das Umweltmanagement. Das Referenzmodell ECO-Integral*, Krcmar, H. et al. München: Oldenbourg, 235-284, 2000.

Weleda, *Transparenz 3*. Schwäbisch-Gmünd: Weleda (available only in German), 2003.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Developments in urban environmental information perception and communication

K. D. Karatzas [a] and J. Lee[b]

[a] *Department of Mechanical Engineering, Informatics Applications and Systems Group, Aristotle University, 54124 Thessaloniki, Greece (kkara@eng.auth.gr)*

[b] *School of Arts, Culture and Environment, University of Edinburgh, U.K. (J.Lee@ed.ac.uk)*

Abstract: Urban environmental information (EI) is related to knowledge concerning the status, the pressures and the consequences of environmental quality parameters to human life and the urban surroundings. One of the most important aspects in the life cycle of EI is knowledge generation and flow, data management and information dissemination between ICT systems and environmental actors, players, or interested parties. In the present paper, urban air quality serves as the application domain, for which the following issues are discussed,

- Analysis of the way that environmental information is related to urban structures (roads, buildings, etc)
- Investigation of the spatial and temporal urban dimensions of the perception and interpretation of environmental pressures from the citizens' point of view.
- Suggested presentation and communicating methods of environmental information to optimise its use and effectiveness towards quality of life

The discussion is initiated from the concept of Ontologies, and provides a perspective on urban environmental information perception and communication for air pollution. The insights reported may be used as a basis for the design of information/management systems.

***Keywords**:* Ontologies, communication, perception, quality of life, environmental information.

1. INTRODUCTION

One of the most cited definitions of ontologies states that *'An ontology is an explicit specification of a conceptualization'* [Gruber, 1993]. Another well known definition is that of Studer et. al., 1998, according to which: *'An ontology is a formal and explicit specification of a shared conceptualization'*. By "projecting" these definitions towards contemporary life in urban agglomerations, one could state that an ontology is a formal, explicit, and equally communicative concerning its content, specification of a conceptualisation, and as such it may be applied to urban life and activities. The latter makes use of the human centric principle concerning the relationship of the being with the world surrounding it, as stated in Plato's *Kritias* in the section on the *cave metaphor*. Thus, the main goal of an ontology may be considered *to effectively structure the representation and communication of information and concepts in an explicit way*. This aspect is of importance when it comes to the interactions that the human being experiences with its urban environment, especially taking into account that more than 60% of the European population lives within cities: there is a dense network of interactions between the individual and urban facilities and services, like transportation, energy, telecommunications, social and financial services. Citizens make use of such services in order to fulfil their needs and by doing so they generate a clear impact on the sustainability of the environment they live in. The introduction of ICT has resulted in the enrichment of available (to citizens) physical facilities and services with digital ones, thus creating the

concept of digital urban infra-structures, existing in "parallel" with the physical ones; Digital services for citizens are represented by local (city-based) content such as community events, nightlife, localized yellow pages, entertainment, visitor's guide, and e-commerce [ISP Glossary, 2008], and are provided via Internet-based information portals, personalized WAP, SMS and voice services, and electronic information boards. As these services are developing more and more in parallel with developments in ICT, they may support new (advanced) ways of managing quality of life on a personal level, with the aid of information related to the quality of the urban environment. The present paper addresses these issues for the environmental domain of air quality, as the latter represents one of the most profound, well recognised and widespread environmental pressures of contemporary Europe.

2. URBAN AIR QUALITY

As already proposed by Dewdney, 1998 "the urban surface is a continuous screen, or a series of overlapping screens, onto which representations are projected." Karatzas, 2003 and 2007, already suggested that urban air quality management and information systems may support decision makers towards environmental management for a sustainable society, and may also "trigger" the creation of new, user-friendly, human-centric environmental information services that advance the improvement of the general quality of life in the city. These services, when applied on an operational basis, can help in the "formulation" of a new, citizen level perception on the environmental semantics of their urban environment and on the influence of their behaviour to the sustainability of their everyday living: Perceived residential quality was found to depend on physical neighbourhood attributes, such as noise and air pollution, on psycho-social characteristics, like safety or crowding, and on the availability and quality of neighbourhood facilities [Van Poll, 1997]. The introduction of human-centric, environmental information services is thus expected to add a positive semantic (in equivalence to physical) attribute to the urban environment, as it will increase the people's feeling for access to services, facilities and amenities, that plays an important role in the perception of quality of life [Sénécal, 2002]. Air quality may play a leading role in becoming the pilot for flexible, adaptable, quality of life electronic information services, incorporating the recent advances in the research results of atmospheric sciences and health-risk assessment, the forthcoming developments of spatial services included in the INSPIRE directive (http://inspire.jrc.it/), and the mandates of EU for e-government and e-participation.

3. PEOPLE'S PERCEPTION AND AWARENESS OF AQ

Environmental awareness is being addressed with an increasing intensity in recent years. Thus, as an example, the German Federal Ministry for Environment, Nature Conservation, and Nuclear Safety and the German Federal Environmental Agency are investigating the environmental awareness and behaviour of German citizens, issuing study results every 2 years [Umweltstudie, 2006].Yet, studies focusing on AQ are not so common and systematic, but they have provided sufficient evidence for drawing a basic sketch concerning what people believe/have in mind about air pollution and related notions. Thus, according to Bickerstaff and Walker, 2003, studies of human response to air pollution first emerged in the 1950s in the USA, where social survey techniques were used. Environmental concerns were expressed for a national level, yet people failed to realise that pollution may be a serious threat to their local community [Murch, 1971]. Public opinion surveys were carried out by the State of California Department of Public Health in the late 1950s and measured psychological dimensions of air pollution. In the UK, early work on the perception of air pollution focused on the operation of Smoke Control Areas under the Clean Air Act of 1956.

In a paper of Skamp et al [2004] on student's ideas and attitudes about air pollution, the authors provide a review of similar work and present their own research results for Australia, for students from ages 6-20 years old. It is interesting to note that according to this study, the majority of students believe that air pollution is natural, and is attributed to animals. The majority of students identified the negative effects of air pollution for the environment and human health, asthma being the disease directly associated with the

problem. Another interesting finding was related to what could be done for improving air pollution, where the majority of students required more education and suggested personal action (environmental friendly behaviour). Thus, the issue of providing AQ information designed to trigger environmentally friendly behaviour, while educating the recipient on the subject, is emerging as important.

In another study, Elliott et. al., 1999, reported that an important parameter on human perception of air pollution is the role of sensate factors as primary predictors of concern, i.e., experience with black soot, bother by air-pollution-related odours, and reported impacts on health and daily life, bringing to the surface the issue of personal experience. Nevertheless, according to Bickerstaff and Walker, 2003 (and references therein), recent qualitative studies show a very low level of awareness and use of air quality information services - in particular those provided by central government. These studies have also revealed a consistent, if muted, scepticism of the motivations of government (or industry) in interpreting and presenting air quality data. When it comes to the relationship between information provided and public understanding, it is mentioned that "…the net result is that the information provided bears little relationship to most people's encounters with, or cultural understanding of, air pollution, and in consequence is either dismissed as an irrelevance or at worst taken as (further) evidence of disingenuous government motivations in the provision of information. The failure of recent changes in the format of air quality information …. to resolve these barriers to use, provides strong evidence of more deep-rooted problems in the communication process, relating to the construction of air pollution problems and assumptions about human behaviour embedded within scientific and policy communities".

A key issue is the understanding that the air as perceived, and actual pollution (the latter based on measurements), are not always associated, underlining an information gap (of both temporal and spatial nature), concerning air pollution levels and air quality perception. Brody et. al [2004], report that while there are research results indicating no correlation between perception and high air pollution levels, there are others demonstrating that awareness of air pollution is scientifically correlated with actual measurements, and that there is a distance decay of concern spreading outward from urban centres towards rural areas. They also report on the linkage between geographic locations and forming perceptions of air quality, yet, such perceptions are also influenced by social, cultural and behavioural parameters. The work of Brody et. al. [2004], for the Texas area also suggests that the role of media providing AQ information is of paramount importance for the way that citizens understand air quality and respond to governmental policies that seek to improve local and regional air quality conditions. As they put it ".. in this respect, more accessible and far-reaching communication channels should be established to provide the broader public with the best available information on air quality conditions". To this end, the conclusion reached by Higgs, 2005 for the dissemination of environment related information concerning siting of waste facilities may be of use: "participative approaches that use IT-based methods, based on combined geographical information systems (GIS) and multi-criteria evaluation techniques that could involve the public in the decision-making process, have the potential to build consensus and reduce disputes". Moreover, Lindley and Crabbe [2004], suggest that IT tools, especially GIS based, should be used for the spatial representation of AQ related information, while also additional visualization methods, including 3D, should be applied, thus verifying the findings reported by Karatzas et. al [2005], concerning the APNEE-TU project briefly summarised as follows|:

- Use of graphs, colours and smileys was found to communicate the content and the "threatening" level of warnings related to urban air quality episodes in various EU countries, in a much more effective way in comparison to the "traditional" usage of scientific terminology and numbers.
- Location-based services, provided with the aid of either personalised (mobile-phone based), or web based information channels, were preferred by citizens more than services that did not include any location oriented information.
- The visualisation pluralism achieved with the aid of on-line 3D animations of pollution was considered to more effectively communicate the air quality problem to citizens.

When it comes to the sociological dimension of the problem, Rydin [1998] notes that the language of air quality policy has traditionally not been placed in the context of moral responsibility to reduce harm to others, but in the context of individual responsibility for self-protection. Petts, 2005, also states that "Altruism gains importance in the context of equity arguments that those most affected by poor air quality are those least able to make a difference and the least contributors, to the problem", and poses an interesting question that we should have in mind when designing air quality information services: " Is it likely that information stressing the individual and social benefits of action will support choice and responsibility or is some further institutional insentivisation required". Petts also mentions that "numerous studies confirm that the public is worried about air pollution. Negative perceptions are shaped by proximity to industry, measurements, and personal experiences. However, there is little evidence of beliefs in strong causal links between diseases and air pollution; rather there is an understanding that chronic illnesses, such as allergies, asthma, and bronchitis, can be affected by air pollution". This finding suggests that AQ information should certainly address those suffering from respiratory diseases, yet it should also underline the consequences of long term exposure to high concentration values.

In the context of urban air pollution, it is argued that there is no benefit to the individual in taking action through the choice of a less polluting but more inconvenient form of transport, other than the potential satisfaction of altruism [Rydin, 1998]. On the other hand, local authorities may also support the opinion that from information supplied, the public can make informed decisions about their lifestyle and how they effect the environment [Jenkins, 2006].

4. AIR QUALITY AND URBAN STRUCTURES

From already published research results, it was evident that urban structure typology plays an important role in the development of everyday economic activities [Joosten and van Nes, 2005], while infrastructures and service availability are related to the perception of environmental quality [Majumder et. al., 2007] Thus it may be suggested that it influences the way that people perceive the quality of the environment they live in. It is therefore necessary to identify aspects of the urban typology related to such a perception. In order to do so, and as we lack field research and related results, and as the scientific literature barely exists in this field, we make use of subjective intuition and scientific judgement, a methodological basis that spans its limitations as far as the absence of actual data may allow.

One basic typology is the distinction between inside and outside spaces, the public versus the private and the physical relationship between those spaces. One of the typological elements that we identified was that of a threshold or a buffer zone (doorways, entrances, etc). This element is directly related to the way that people communicate with their environment and with each other, and its existence was identified to prepare the individual towards entering the proper communication mode and physical interaction mode. As a consequence of that, such buffer zones act as "smootheners" of environmental information towards the individual. This finding also led us to suggest that open spaces (parks, squares, open markets, etc) are actually an ensemble of communication functions that are clustered together on the basis of their resemblance and scope (marketing goods, communicating with people of similar interests, exchange of information concerning social matters, taking part in major social events, etc.). Another urban typology identified was that of transportation infrastructure. The physical proximity to such infrastructures (evaluated with the aid of visual contact or noise-air pollution sensing, should be considered to be the key element the way that people interpret the quality of the environment they live in [Elliott et. al., 1999, Petts, 2005]

We therefore seek ways in which these typological concepts may be developed towards ontologies that will help to structure the communication of environmental information more effectively.

5. COMMUNICATION OF AQ INFORMATION

Communication is related to: presentation; multimodality; representation. These are key parameters to consider in any communicational context. We need to consider the interplay of these factors with the intended typological framework. Here we focus on the use of mobile devices – phones, which are the ICT paradigm that provides a new communication language with and within the digital world. This emphasis has as a prerequisite that the society of users is not influenced by problems of digital divide, which is at least at a minimum in EU countries. When focusing on mobile devices, the following significant factors may be identified: (1) Geographically locatable, localised devices and services. (2) Small size; mobility. (3) Extremely wide availability and user base (ubiquity). (4) Audio interface that may be especially natural to use. (5) Restricted visual display area.

Communication using these devices can and should trade on the above factors, which have many important implications. Information needs to be packaged efficiently (5); extended explanations are to be avoided. Hence abbreviations are important, but these are also inherently ambiguous. Disambiguation can be helped by providing complementary information on additional channels, e.g. sound (4), but comes more from context and background knowledge. A major part of context may be physical location (1) – devices can refer to this directly and e.g. indicate when a change of location (2) has been detected. Indexical features of language (e.g. “here”, “this”) take advantage of such contextual cues so can be usefully exploited. Background knowledge is often derived from shared assumptions about the meanings of terms, symbols, indicators of various kinds. Shared assumptions come from (loosely speaking) a community of users with a commonality of background. This background may be shared cultural history among the users (from education, mass media, etc.), which is an increasingly reasonable assumption (3); or may be narrower shared practices that have evolved through joint activity or attention to some task, or have been acquired through explicit training.

So, one way to look at the challenge of these devices is to ask how best to exploit the features (1-4) to overcome the display problem (5). But alternatively, one can see whether the affordances of the display along with the other factors can be combined to provide an unusually fertile opportunity for developing new communicational possibilities, such as novel symbol systems.

The difference between icons and symbols can be characterised in information theoretical terms by looking at how information is distributed between the sign itself and the user’s knowledge of how the sign was used in the past [Shannon, 1948]. As a community uses signs in communication, this distribution tends to change in the direction of exploiting the increasingly shared (background) knowledge of the community about the use of the signs; and hence signs become more symbolic[cf. Garrod et al., 2007]. This process can be exploited to create new “languages” (perhaps graphical languages, animated languages, sonic languages, etc.) that offer us new opportunities to communicate effectively and with acceptance (or “buy in”) from appropriate user groups. The ways in which these languages can emerge depend in complex ways on the nature of the devices combined with the interests, abilities and perceptions of the users. A typical example of such a development is the now familiar “text language” that has emerged among the younger generation of users of SMS messaging, which is of course restricted to the only visual messaging modality available on early mobile phones.

So when it comes to air quality information, we need to define our user groups carefully, to determine communities of communicators who will form the basis for this kind of development. We then need to provide them with suitable communicational tools. Communication in this case needs to be a two-way process: it is not enough to push information out to the users, we have to facilitate their reacting to it in some way, passing it on to others, responding with their opinions, etc. These imperatives reflect the views of Brody et al [2004] and Higgs [2005], as quoted above. They are also in some respects the basis of the much-vaunted “Web 2.0” paradigm in internet terms. Similar thinking applies in the mobile context. Users need the flexibility to define their own ways of working, as they have done with SMS. Newer devices allow for languages which can be on the one hand much richer, and on the other hand potentially more specialised to particular content domains.

To see how this might interact with the notion of typology introduced above, let's consider again the transitional boundary that exists between (a) the inside (of buildings, spaces, enclosures, even cities) and (b) the outside (the environment), the private and the public. We could call this the "threshold" – an emotive idea in architecture, and perhaps elsewhere. How do people perceive environmental threshold conditions, and how do they communicate about them? What are the roles of these conditions as barriers to movement and progress, as comforting and protective shields against various kinds of threats, or simply as transitions between contrasting situations? Such questions would be highly relevant to investigating how information should be presented, and how people would decide to present it. Identifying the crossing of (something that would be perceived as) a threshold could be an appropriate trigger for delivery of information; understanding how the threshold is perceived should certainly condition the presentation. But also, an important aspect of the private/public threshold is that the private can refer to the individual (as perhaps we initially expect), but also to the group or community. In this sense, people may perceive a physically identifiable boundary as a threshold and in a related way may perceive communicating with members of a different group as crossing a threshold. Often these notions coincide, where groups (e.g. citizens of a particular district) tend to identify with a spatial area, but they may not (as e.g. the young and the middle-aged, as distinct groups, commonly inhabit the same household). And indeed, it may be that we can use these overlaps to aid communication across thresholds. In the "buffer zones", people meet and communicate differently as compared with the inside or the outside; there is a certain ambiguity about which zone they are in, and an opportunity to allow permeation between languages. We can perhaps use these zones not only to facilitate information exchange (between groups), but also to introduce or insinuate new communicational ideas, "seeding" the communities on either side to develop their own interpretations, usages and forms of language, but nonetheless to share a certain core of understanding.

6. THE USE OF ONTOLOGIES

It follows from the above that different groups within the wider community will develop their own conceptualisations of phenomena, based on their perceptions, which may differ markedly, and also on their inevitably divergent group communication histories. Therefore, there will be differences when it comes to considering formal characterisations of these conceptualisations, which is to say that we will need different ontologies. To some extent, ontologies can be expected to characterise groups. In the buffer zones, we will observe ontological shifts.

This should come as no surprise. Ontologies in practice are always somewhat mutable, subject to change and evolution, and bound to communicating groups that are themselves somewhat fluid [cf. Lee and McMeel, 2006]. It means that from the point of view of describing a domain such as air quality, we should not seek to find some single ontology that will do this job for all comers. Rather we must recognise that conceptualisations of air quality will differ between communities; and we (whoever we are – scientists, etc.) are just another community. It is our job to communicate with diverse groups. We need therefore to approach them through suitable buffer zones, and "seed" these with our own concepts of the material in our domain. They will then, inevitably, take off in their own directions and become to some degree incommensurable with an ontology that would have described our original view.

This does not, however, mean that ontologies are useless. We need them to import some formal structure to the communication systems that we develop, and then indeed we can use them to characterise the very shifts and changes in those systems that demand changes to our ontology. They are hence a critical tool both for research and for development. Moreover, the differences in ontologies do not necessarily result in difficulties in communication. ICT provide methods that allow for the "collaboration" between different ontologies, including ontology markup languages. Thus, the user of a quality of life information service may combine different concepts, thesauri, and metadata, by bringing together, on the fly, different ontologies, so that an information service may be personalised and materialised on demand.

7. CONCLUSIONS AND RECOMMENDATIONS

In the present paper the issue of environmental information perception, analysis and communication is addressed for the domain of urban air quality. The aim is to better understand how the physical interaction of people with the environment they live in is related to the way that they understand environmental pressures. This is a research work in progress, for which the following conclusions are drawn and recommendations are made:

The human perception of atmospheric quality is closely linked to the interaction of the individual with the physical space and the available interfaces between digital services and people. The communication of AQ information should take into account a number of factors that may be identified, and should also be based on a two way communication channel, that would allow people to interact and develop common perceptions and understandings on the basis of common interests. The use of multimedia communication means may enhance communication effectiveness and advance the usefulness of such quality of life information services. On this basis, the following may serve as recommendations for future research in this area.

1. Investigation of the spatial and temporal dimensions of human perception in relation to the actual situation concerning environmental pressures in urban spaces;
2. Assessment of the communication media and modes that address the spatial and temporal scale of both environmental quality and urban life (graphs, noise, use of vibrating mobile phones as alarm notifications, etc.);
3. Analysis of aspects of geometry, shape, size and the resulting concept of proximity, inclusion, exclusion concerning quality of life environmental information services;
4. Direct and intensive participatory design, with specific user groups, of mobile interfaces that combine these media and concepts flexibly for innovative communication.

ACKNOWLEDGEMENTS

This work is based on the scientific collaboration that the authors have developed in the frame of COST action C21: www.towntology.net, and is also related to work performed by the first author in the frame of COST action ES0602, www.chemicalweather.eu.

REFERENCES

Bickerstaff K. and Walker G. (2003), The place(s) of matter: matter out of place - public understandings of air Pollution, Progress in Human Geography 27,1 pp. 45-67

Brody S.D., Peck B.M. and Highfield W.E. (2004), Examining Localized Patterns of Air Quality Perception in Texas: A Spatial and Statistical Analysis, Risk Analysis, Volume 24, Number 6, pp. 1561-1574(14).

Dewdney, C. (1998) Last Flesh: Life in the Transhuman Era. Harper Collins, Toronto.

Elliott S.J., Cole D.C., Krueger P., Voorberg N. and Wakefield S. (1999), The Power of Perception: Health Risk Attributed to Air Pollution in an Urban Industrial Neighbourhood, Risk Analysis, Vol. 19, No. 4, pp. 621-634.

Higgs G. (2005), Integrating multi-criteria techniques with geographical information systems in waste facility location to enhance public participation. Waste Manage Res 2006: 24: 105–117

Garrod, S., Fay, N., Lee, J., Oberlander, J. and McLeod, T. (2007) Foundations of representation: Where might graphical symbol systems come from? Cognitive Science: A Multidisciplinary Journal, Vol. 31, No. 6, pp 961-987.

Gruber T.R. (1993), A translation approach to portable ontologies. Knowledge Acquisition, 5(2):199-220 (http://tomgruber.org/writing/ontolingua-kaj-1993.pdf

ISP Glossary (2008), *Digital City*, http://isp.webopedia.com/TERM/D/Digital_City.html (last visited April 4th, 2008)

Jenkins N. (2006), Local Air Quality Management - Air Quality Action Plan for Rushmore Borough Council, http://www.rushmoor.gov.uk/media/adobepdf/0/3/AirQualityActionPlanMay06_1.pdf

Joosten V. and van Nes A. (2005), How block typology influences the natural movement economic process – micro spatial conditions on the dispersal of shops and cafes in Berlin, Space Syntax Symposium 2005, http://www.spacesyntax.tudelft.nl/media/Posters/225.pdf

Karatzas, K. (2003) Environmnetal Informatics: Concepts and Definitions, Proceedings of "The Information Society and Enlargement of the European Union"-17th International Cenference Informatics for Environmental Protection, A. Gnauck and R. Heinrich eds., Part 1: Concepts and Methods, pp. 146-151, ISBN 3-89518-440-3.

Karatzas K., Endregard G., and Fløisand I. (2005). Citizen-oriented environmental information services: usage and impact modelling, Proceedings of "Informatics for Environmental Protection- Networking Environmental Information"-19th International EnviroInfo Conference, Brno, Czech Republic, pp. 872-878 http://www.apnee.org/documents/papers/APNEE-TU-Karatzas_Endregard_Floisand_EnviroInfo2005.pdf

Karatzas K. (2007), State-of-the-art in the dissemination of AQ information to the general public, Proceedings of the 21st International Conference on Informatics for Environmental Protection - EnviroInfo2007, (Hryniewicz O., Studziński J. and Romaniuk M., eds.), Vol. 2., pp. 41-47, Shaker Verlag, Aachen, 2007, ISBN 978-3-8322-6397-3 (conference date and location: Warsaw, Polant, Sept. 12-14, 2007

Lee, John and McMeel, Dermott (2007). 'Pre-ontology' considerations for communication in construction. In J Teller, J Lee and C Roussey (eds), Ontologies for Urban Development, (Computational Intelligence series vol. 61,) Springer Verlag, pp169-179.

Lindley S.J. and Crabbe H. (2004), What lies beneath?—issues in the representation of air quality management data for public consumption, Science of the Total Environment, Vol. 334– 335, pp. 307–325

Majumder A.K., Hossain E., Islam N. and Sarwar I. (2007), Urban environmental quality mapping: a perception study on Chittagong metropolitan city, Kathmandu Univ. Journal of Science, Engineering and Technology, Vol.I, No.IV, August, 2007 (http://www.ku.edu.np/kuset/fourth_issue/orginal/6%20A.Kamruzzaman-Edited.pdf)

Murch (1971), Public Concern for Environmental Pollution, The Public Opinion Quarterly, Vol. 35, No. 1. (Spring, 1971), pp. 100-106.

Petts J. (2005), Health, responsibility, and choice: contrasting negotiations of air pollution and immunisation information, Environment and Planning A, Vol. 37, pp. 791 -804

Rydin Y. (1998), Managing urban air quality': language and rational choice in metropolitan governance,' Environment and Planning A, 30, 1429 – 1443

Sénécal G., (2002): Urban Spaces and Quality of Life: Moving Beyond Normative approaches, Horizons 5, No. 1, pp. 20-22.

Shannon C.E. (1948), A Mathematical Theory of Communication", Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, July, October, 1948

Skamp K., Boyes E. and Stanisstreet M. (2004), Students' Ideas and Attitudes about Air Quality, Research in Science Education 34: 313–342, 2004.

Studer R., Benjamins R. and Fensel D. (1998), Knowledge Engineering: Principles and methods. Data and Knowledge Engineering, 25: 161-197.

Umweltstudie, 2006, Environmental Awareness in Germany 2006, http://www.umweltbewusstsein.de/ub/englisch/2006/index.html, last visited April 4th, 2008.

Van Poll, R. (1997). The perceived quality of the urban residential environment. A multi-at-tribute evaluation. Ph.D.-thesis, University of Groningen, The Netherlands.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Environmental Information Systems of Public Administration – Status and Trends

W. Geiger[a] **and W. Pillmann**[b]

[a] *Forschungszentrum Karlsruhe, Institute for Applied Computer Science, P.O.B. 3640, 76021 Karlsruhe, Germany (werner.geiger@iai.fzk.de)*

[b] *ISEP International Society for Environmental Protection, Bechardgasse 24/12, 1030 Vienna, Austria (pillmann@isep.at)*

Abstract: As a basis of political decisions and administrative measures for the long-term conservation of the natural resources of life, public administration acquires, processes, and supplies a variety of information on the environment. The environmental information systems used for this purpose are subject to constant further development. The present paper will focus on requirements made on environmental information systems in Europe, which essentially result from directives of the European Union, and outline some basic technical trends in the field of public environmental information systems. Furthermore, such systems will be classified roughly depending on the providers, target groups, intentions, etc. By way of example, selected public environmental information systems in Germany shall be described.

Keywords: Environmental information systems (EIS); EIS for the public (PEIS); Public access to environmental information; Requirements on EIS; Categories of EIS.

1. INTRODUCTION

Environmental information systems (EIS) of the public administration are major instruments for environmental monitoring, environmental reporting as well as for administration and public relations work. Requirements made on information and communication systems of public administration in the environmental area are constantly growing. New or increased requirements on the systems result from increasing information obligations among administrative bodies and towards the public as well as from new hazards affecting the environment and man via the environment.

The present contribution will focus on the status and trends of environmental information systems of public administration. As this subject is a very complex one, the contribution does not claim to be complete, of course. A good overview of the activities in this field is given by the proceedings of the EnviroInfo conferences, among others by the EnviroInfo 2006 [Tochtermann and Scharl, 2006] and the EnviroInfo 2007 [Hryniewicz et al., 2007]. The present article will highlight some basic trends with particular attention being paid to the development in Europe and especially Germany. Requirements on EIS of public administration, R&D cooperation within the administration, major technical trends, categories of EIS, and some exemplary application systems will be dealt with in more detail.

2. REQUIREMENTS, R&D COOPERATION

The tasks to be fulfilled by the environmental administration are subject to constant change resulting from changing legal requirements and general conditions as well as new ICT technologies.

In the past, emphasis was put on supplying information internally for supporting the administration in its planning and execution duties and on supplying information to persons responsible for decision-making. Meanwhile, information of the public and public participation have gained importance. The legal basis in the European Union comprises the EU Directive on Public Access to Environmental Information [EU, 2003] and the corresponding national laws in the individual EU member states (Environmental Information Acts).

Additional requirements on environmental information systems of public administration resulted from new or more strongly perceived hazards for the environment and man via the environment. Examples are public health risks by fine dust pollution, biodiversity loss, and climate changes with their effects on man and the environment. As regards natural disasters (e.g. floods, droughts, forest fires), requirements are increasing in connection with early warning systems and catastrophe management.

Furthermore, reporting obligations of subordinate to superordinate national or European authorities have been extended, e.g. by the European Union Water Framework Directive and the EU Directive on Environmental Noise. Administrations are also obliged to comply with agreements on an enhanced national and international data exchange, e.g. for nuclear reactor monitoring.

Due to budget constraints in many EU member states, however, availability of staff and financial resources mostly has not kept up with the increasing requirements on environmental information systems. For this reason, efficient development of such systems has gained importance. R&D cooperation was aimed at efficiently using the resources available and jointly investigating and developing new approaches to supplying environmental information to the administration and public. Examples of such cooperation agreements and projects in Germany are KoopUIS (Cooperation on the conception and development of software for EIS) and PortalU (http://www.portalu.de/) on the federal and state level as well as KEWA (Cooperative development of economical applications for the environment, transportation, and adjacent areas) that is pursued together with the municipal level [Mayer-Föll et al., 2007]. In addition, framework concepts for environmental information systems are developed, including guidelines and recommendations for the development of the individual information systems, e.g. RK UIS (Framework Conception EIS) in Baden-Württemberg [Mayer-Föll and Kaufhold, 2006].

3. MAJOR TECHNICAL TRENDS

Major trends with respect to the technology and infrastructure of environmental information systems in public administration are the enhanced use of internet technologies (up to Web 2.0 technologies) as well as of the internet itself as an infrastructural basis, of knowledge-based methods and techniques (e.g. ontologies) as well as of GIS technologies and geo-infrastructures for the networking, integration, and multiple use of services and information.

Broad use of the internet was and is significantly driven by the already mentioned obligation to supply environmental information to the public (Environmental Information Acts). As the information usually is supplied in a decentralised manner by the respective data-collecting authority, it was required to provide a general overview of the environmental information offered and to establish efficient means to search for the information needed. For this purpose, meta information systems and environmental web portals have been developed (cf. Chapter 5).

To enhance their usability, service concepts up to service-oriented architectures are increasingly applied in the development of web-based EIS, and the corresponding web services are made generally available by the systems. To support potential users of such web services (systems developers) in the search for appropriate services, central web service registries with descriptions of the services are implemented, e.g. UIS-UDDI [Paoli et al., 2007].

On the European as well as on the federal and state levels, frameworks are developed for the efficient generation of environmental information systems. In this connection, the EU project *ORCHESTRA* (Open Architecture and Spatial Data Infrastructure for Risk Management, http://www.eu-orchestra.org/) has to be mentioned. Within this project, a variety of basic services based on OGC standards and the respective architecture are specified and implemented for applications in the field of environmental risk management.

In addition, initiatives were launched for the set-up of IT infrastructures in the environmental and related fields, in particular geo data infrastructures. Examples are the *Infrastructure for Spatial Information in the European Community (INSPIRE)* [EU, 2007] and similar initiatives to establish spatial data infrastructures on the national level (e.g. *GDI-DE* in Germany, http://www.gdi-de.org/).

4. APPLICATION CATEGORIES

Depending on the providers, target groups, purposes, etc., various types of environmental information systems of public administration may be distinguished (see also [Geiger et al., 2006]):

- Depending on the administration level responsible for the EIS or offering the information:
  local-area EIS, municipal EIS, regional EIS (state EIS), national EIS, and EIS of international organisations. Systems of the lower administrative levels are dominated by information on individual environmental objects (e.g. tree catasters) and local information, such as local traffic information. EIS of higher administrative levels predominantly supply general information on the respective state/nation and statistical data, and they primarily address supraregional and global challenges like climate change.

  An important trend with respect to this aspect consists in enhanced cooperation of authorities on the same administration level and beyond administration levels. This trend is due to technical needs (e.g. cross-national high-quality flood forecast or supraregional supply of air quality data), and by cooperative development of systems, economic efficiency can be increased (cf. Chapter 2).

- Depending on the target group of administration/public:
  systems that support the administration itself in the planning and execution of environmental duties, monitoring of the environment, risk/accident management, and in supplying information to decision-makers, in contrast to environmental information systems for the public (PEIS). The latter comprise systems that address the broad public (interested citizens) and systems that rather supply information to the expert public (e.g. contracted engineering offices).

  The significance of EIS for the public is increasing, in particular due to the extended right of free access to environmental information (cf. Chapter 2). Of the EIS for the administration, those information systems are gaining relevance, which serve to comply with reporting obligations to the European Commission.

- Depending on the intention of supplying expert information or promoting environmental awareness:
  systems that were initially designed for professionals and just grant the public access to the expert information available in the environmental administration versus systems designed specifically for the general public to promote its environmental awareness. For systems to promote environmental awareness the expert information has to be processed in view of the broad public, both in terms of contents and presentation.

  In the past, most of the environmental information systems were initially designed for professionals. Meanwhile, systems designed specifically for the general public are gaining importance (cf. Chapter 2). They are aimed at changing the environmental awareness and, hence, achieving a more environmentally compatible behaviour.

- Depending on the contents:
  systems that predominantly serve as portals to other information systems like state environmental portals or meta information systems versus systems that predominantly hold their own contents.

  The relevance of environmental portals of the administration is increasing, because they considerably support the search for quality-assured environmental information and offer the chance of a wider use of the contents generated by the authorities with high expenditures. As far as systems without own contents are concerned, use of and usability by other systems (networking) is gaining importance.

- Depending on use on desktop/in the field:
  systems designed for use on desktop computers (incl. laptops) with broad-band internet access in contrast to systems designed for use on mobile computers in the field, i.e. with a small display and – so far – no or low-performance internet access.

  The former category has been the dominating one up to now at least. In the environmental area, context-based mobile applications, however, promise to be of particular use. When low-cost mobile devices with higher capabilities will be available, such as high-performance successors of today's mobile phones with good internet access and GPS, a strong increase in mobile applications has to be expected.

5. EXAMPLES OF APPLICATION SYSTEMS

In the following sections, four selected examples of environmental information systems of the public administration in Germany shall be presented in more detail: the German Environmental Information Portal PortalU® and environmental information portals on the state level, the Theme Park Environment for the broad public, and the mobile guide to nature MobiNaG for areas to experience nature.

5.1 Environmental Information Portals

The probably most extensive environmental information system in Germany is the German Environmental Information Portal PortalU, http://www.portalu.de [Vögele et al., 2007]. PortalU is being developed within the framework of a cooperation between the Federal Republic of Germany and federal states (cf. Chapter 2). It shall give the public access to the environmental information of the administration. Consequently, PortalU is an instrument to implement the Environmental Information Act.

The portal offers central access to more than a million of web pages and database entries of public institutions and organisations (Figure 1). PortalU uses the InGrid® software [Klenke et al., 2006] that is based on a strongly modular, distributed architecture. A central component is the integrated search engine that offers advanced search functions relating to the environmental topic, spatial reference, time reference, etc. using the Semantic Network Service (SNS) of the Federal Environmental Agency [Rüther and Bandholtz, 2005]. Since May 2006, this system has replaced the preceding system German Environmental Information Network *gein*®.

Approaches to environmental information portals have also been developed on the federal state level. The state environmental portals try to make accessible as comprehensively as possible the administrative environmental information of the respective state, including the municipal level, via centrally provided search and navigation functions. So far, comprehensive state environmental portals have only been offered in Germany by the states of Baden-Württemberg (http://umwelt.baden-wuerttemberg.de/servlet/is/811/) and Saxony-Anhalt (http://umwelt-st.iai.fzk.de/servlet/is/811/). These states have bundled their resources in a joint project (meanwhile, also the state of Thuringia has joined this cooperation) in order to implement state environmental portals on a joint technical basis [Schlachter et al., 2007]. The state environmental portals offer access among others via environmental topics, a full-text search, a list of providers as well as a list of sites with

measurement values and other current values. The classification of the contents of the state portals (environmental topics etc.) follows to the classification in the national portal PortalU, and the state portals supply meta data of their information sources – without the information sources of the municipal level – to the national portal.

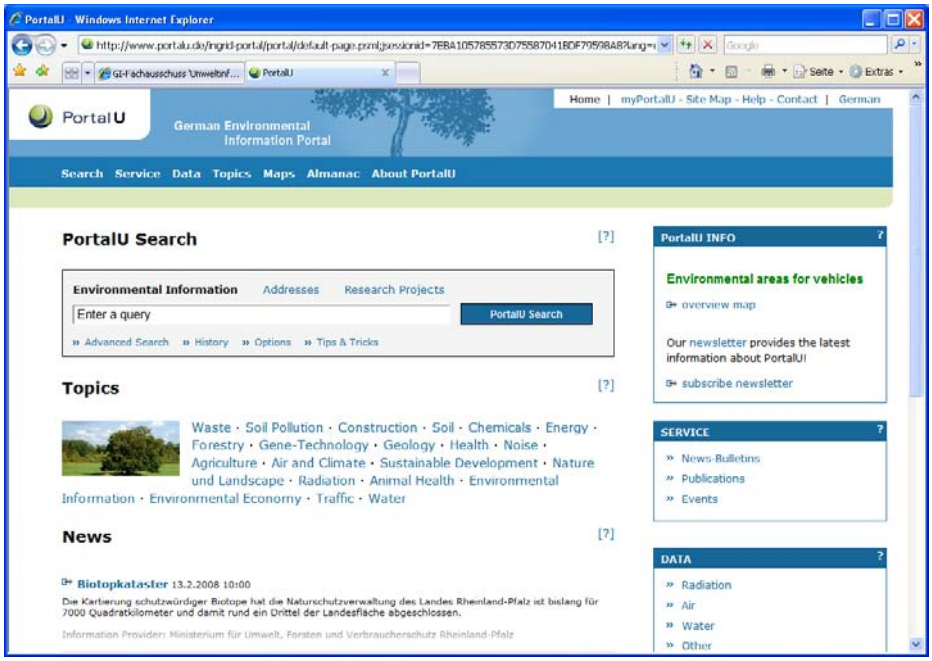


Figure 1. Homepage of the German Environmental Information Portal PortalU

In terms of the categories of EIS described in Chapter 4, PortalU is a national EIS, while the state portals are regional EIS. Target groups are the broad public as well as experts. The systems provide access to expert information and represent portals without own contents. The systems are conceived for access via desktop computers.

5.2 Theme Park Environment

In contrast to the above environmental portals, the Theme Park Environment of the state of Baden-Württemberg (http://www.themenpark-umwelt.de) is not intended to give the public access to the expert environmental information already available at the administration, but to supply information on regional environmental objects and topics to the ecologically interested public in an attractive manner and, thus, to enhance interest in the environment and environmental awareness.

The system largely uses multi-media information like maps, photos, audio files and also life videos. Development of this system focused on the implementation of flexible information management based on ontologies, integration of various external data sources in the system (EIS map service, Google Maps, Wikipedia, Environmental Portal BW) as well as on the supply of a powerful author component. With this author component, experts from various fields may decentrally acquire the information elements and flexibly integrate them in presentations together with contents from external sources [Düpmeier and Geiger, 2006].

In terms of the above EIS categories, the Theme Park Environment is a regional (state) environmental information system for the public. It is aimed at promoting environmental awareness, contains own contents, and is conceived for desktop computers.

5.3 Mobile Guide to Nature MobiNaG

The mobile guide to nature MobiNaG goes a step further, its intention being to promote the environmental awareness of the population [Ruchter, 2007]. The project is aimed at developing a mobile information system that motivates the user to directly experience nature and supports him in doing this.

The mobile guide provides environmental and nature information on PDAs. The core context-based services of the system include navigational assistance, such as orientation and routing along specific tours, as well as information on natural phenomena along the tour (see Figure 2). The information at each station is tailored to the target group and adapted to the location and to the season of the year. In cooperation with the nature protection centre in Karlsruhe-Rappenwört (Germany), a prototype was developed and evaluated with visitors of this centre. Evaluation has shown that in particular young people and young adults prefer this mobile guide to printed brochures due to its "coolness" and their enthusiasm for computers. Based on the experience gained from this evaluation, a first operational version of the mobile guide to nature was generated [Ruchter et al., 2006].



Figure 2. Screenshots of the MobiNaG prototype with the two basic services of navigation (left) and tour information (right) [Ruchter et al., 2006]

In terms of the EIS categories, the mobile guide to nature MobiNaG is a local-area EIS for the public that is aimed at promoting environmental awareness. It has own contents and was developed for use in the field.

REFERENCES

Düpmeier, C., and Geiger, W., Theme Park Environment as an example of environmental information systems for the public. *Environmental Modelling & Software*, 21, 1528-35, 2006.

EU, Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on Public Access to Environmental Information and Repealing Council Directive 90/313/EEC. Official Journal of the European Union, L 41/26, 2003.

EU, Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Official Journal of the European Union, L108, 2007. http://eur-lex.europa.eu/LexUriServ/site/en/oj/2007/l_108/l_10820070425en00010014.pdf.

Geiger, W., Düpmeier, C., Ruchter, M., Schlachter, T., Weidemann, R., Mayer-Föll, R., Ebel, R., and Linnenbach, M., Experience gained from the use of environmental information systems for the public in the state of Baden-Württemberg. Voinov, A. [Ed.], Summit on Environmental Modelling and Software: Proc. of the iEMSs 3rd Biennial Meeting, Burlington, Vt., July 9-13, 2006. Burlington, Vt.: Environmental Modelling and Software Society, CD-ROM, ISBN 1-4243-0852-6, 2006. http://www.iemss.org/iemss2006/papers/s5/84_Geiger_1.pdf.

Hryniewicz, O., Studzinski, J., and Romaniuk, M. (Eds.), EnviroInfo 2007 - Environmental Informatics and Systems Research; Proc. of the 21st Internat. Conf. Informatics for Environmental Protection, Warsaw, Poland, September 12-14, 2007. Aachen: Shaker Verlag, ISBN 978-3-8322-6397-3, 2007.

Klenke, M., Kruse, F., Lehmann, H., Riegel, T., and Vögele, T., InGrid 1.0 – the nuts and bolts of PortalU. In: Tochtermann, K., and Scharl, A., EnviroInfo 2006 – Managing Environmental Knowledge; Proc. of the 20th Internat. Conf. Informatics for Environmental Protection, Graz, Austria, September 6-8, 2006, pp. 437-440. Aachen: Shaker Verlag, ISBN-13:978-3-8322-5321-9, 2006.

Mayer-Föll, R., and Kaufhold, G. (Eds.), Umweltinformationssystem Baden-Württemberg, RK UIS 06 - Rahmenkonzeption 2006. Universitätsverlag Ulm GmbH, ISBN 3-89559-261-7, 2006.

Mayer-Föll, R., Keitel, A., and Geiger, W. (Eds.), F+E-Vorhaben KEWA – Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt und Verkehr in neuen Verwaltungsstrukturen, Phase II 2006/07. Forschungszentrum Karlsruhe, Wissenschaftliche Berichte, FZKA 7350, 2007. http://uis.baden-wuerttemberg.de/servlet/is/27457/.

Paoli, H., Holtmann, C., and Ebel, R., UIS-UDDI – Entwicklung eines Web Service-Verzeichnisses für das UIS Baden-Württemberg. In: Mayer-Föll, R., Keitel, A., and Geiger, W. (Eds.), F+E-Vorhaben KEWA – Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt und Verkehr in neuen Verwaltungsstrukturen, Phase II 2006/07. Forschungszentrum Karlsruhe, Wissenschaftliche Berichte, FZKA 7350, pp. 21-30, 2007.

Pillmann, W., Geiger, W., and Voigt, K., Survey of environmental informatics in Europe. *Environmental Modelling & Software*, 21, 1519-27, 2006.

Ruchter, M., et al., MobiNaf. Ergebnisse der Evaluation und Studie zu weiteren Einsatzmöglichkeiten mobiler Naturführer. In: Mayer-Föll, R., Keitel, A., and Geiger, W. (Eds.), F+E-Vorhaben KEWA – Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt und Verkehr in neuen Verwaltungsstrukturen, Phase I 2005/06. Forschungszentrum Karlsruhe, Wissenschaftliche Berichte, FZKA 7250, pp. 171-176, 2006.

Ruchter, M., A new concept for mobile environmental education. Dissertation, Universität Karlsruhe (TH). Universitätsverlag Karlsruhe, ISBN 978-3-86644-182-8, 2007.

Rüther, M., and Bandholtz, T., Applications of Environmental Semantics. Networking Environmental Information. In: Hrebicek, J., Racek, J. [Eds.], Proc. of the 19th Int. Conf. Informatics for Environmental Protection, EnviroInfo Brno 2005, Brno, Czech Republic, Sept. 7-9 2005, pp. 44-48. Brno : Masaryk University, ISBN 80-210-3780-6, 2005.

Tochtermann, K., and Scharl, A. (Eds.), EnviroInfo 2006 – Managing Environmental Knowledge; Proc. of the 20th Internat. Conf. Informatics for Environmental Protection, Graz, Austria, September 6-8, 2006. Aachen: Shaker Verlag, ISBN-13:978-3-8322-5321-9, 2006.

Schlachter, T.; Geiger, W.; Weidemann, R.; Ebel, R.; Tauber, M.; Mayer-Föll, R.; Sawade, A.; Bachmann, V.; and Köther, B., Accessing administrative environmental information. Tatnall, A. [Eds.], Encyclopedia of Portal Technologies and Applications, Vol. 1 pp. 20 – 25. Hershey, Pa. Information Science Reference, ISBN 978-1-59140-989-2, 2007.

Vögele, T., Klenke, M., Kruse, F., Lehmann, H., and Giffei, C., PotalU – a tool to support the implementation of the European Environmental Information Directive. In: Hryniewicz, O., Studzinski, J., and Romaniuk, M. [Eds.], EnviroInfo 2007 – Environmental Informatics and Systems Research; Proc of the 21st Internat. Conf. Informatics for Environmental Protection, Warsaw, Poland, September 12-14, 2007, Vol 2, pp. 153-155. Aachen: Shaker Verlag, ISBN-978-3-8322-6397-3, 2007.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Interaction design for rural agricultural sensor networks

K. Walker[a], A. Kabashi[b], J. Abdelnour-Nocera[c], K. Ngugi[d], J. Underwood[a], J. Elmirghani[b] and M. Prodanovic[e]

[a] London Knowledge Lab, Institute of Education, University of London. 23-29 Emerald Street, London, UK (k.walker/j.underwood/r.luckin@ioe.ac.uk)
[b] School of Electronic & Electrical Engineering, University of Leeds. LS2 9JT, Leeds, UK (A.H.Kabashi/J.M.H.Elmirghani@leeds.ac.uk)
[c] Institute for Information Technology, Thames Valley University. St. Marys Road – Ealing, London, UK (Jose.Abdelnour-Nocera@tvu.ac.uk)
[d] Department of Agriculture, University of Nairobi. Chiromo Campus, Nairobi, KE (eckngugi@africaonline.co.ke)
[e] Department of Electronic & Electrical Engineering, Imperial College. Exhibition Road, London, UK (milan.prodanovic@imperial.ac.uk)

Abstract: We describe the ongoing design of a sensor network for small family farms in rural Kenya. The sensor network is just one part of an 'ecology of resources' in which handheld devices are used to bridge the sensor network and a computer-based access point. We describe the two villages where the system is deployed and the user requirements collected. We then describe the architecture of the sensor network and detail how it fits in with the larger integrated system. We then detail our approach to interface and interaction design, and conclude by describing the next steps in the project.

Keywords: interaction design; sensor networks; agriculture; mobile.

1. INTRODUCTION

In rural parts of developing countries, many people rely upon farming to provide food, yet lack valuable agricultural information about soil conditions, the weather forecast, pest and plant diseases, efficient irrigation methods, and crops they intend to grow. They also often lack basic literacy skills, and have little or no knowledge about Information and Communications Technologies (ICT). Agricultural information systems are common among large commercial farms in the developed world, but small, local implementations for the developing world require sensitive and participatory design practices, since existing technologies, interaction design methodologies, and usability testing techniques are all developed by and for the developed world.

One technology is nearly pervasive however: the mobile phone. For example in Kenya, the number of mobile phone users has gone from 300,000 in 2000 to more than 9 million in 2007; this is due in part to the fact that at any given time, two-thirds of the country's land lines are not working [Economist, 2007]. Mobile technologies can provide a two-way flow of information as rural farmers can not only receive data about markets, methods and pests, but also collect data on their crops which can be shared with scientists or shared databases.

Wireless sensor networks have also been proposed to help meet the information needs of the rural poor. For example according to Panchard et al [2006], information on the temporal and spatial variability of environmental parameters and their impact on farmers' crops, soil, and other aspects of farming can play a major role in formulating the farmers' strategy.

2. THE VESEL PROJECT

VeSeL (Village e-Science for Life) is a research project, part of the Bridging the Global Digital Divide network sponsored by the Engineering and Physical Sciences Research Council (EPSRC) in the UK. It runs for three years from September 2006. The aim of the project is to enable rural communities in Sub-Saharan Africa to use advanced digital technology to improve their agricultural practices and literacy levels, with particular emphasis upon educational barriers. This is a demand-driven participatory iterative which involves end users (farming communities) from the start, and develops and refines technological solutions based on their needs and with their ongoing participation in the design process.

In contrast to previous initiatives in this area, we view sensor networks as just one element in an 'ecology of resources' decision support system for resource-poor farmers. This approach is not focused solely on information delivery, rather a two-way exchange. In addition to access to communications and infrastructure (via the Internet and mobile phone network), users can access and act upon data from their environment via the sensor network, and also send data both manually and automatically over the same infrastructure. This approach emerged from our initial consultations with the communities, as they expressed the desire not just for information but to share their data with other communities and the wider world.

3. SITES SURVEY

3.1 Kiangwachi

Kiangwachi is located in the central highland area, close to Mount Kenya (Africa's second highest peak. The land is generally lush, green and fertile, and rainfall is regular, although some degradation has been caused by deforestation. There are large well-watered fields of maize, beans, wheat and vegetables and the soil is rich, dark red-brown and moist. Water is available locally, although petrol pumps are used to irrigate the crops.

The group of farmers we met were newly formed and located on the slopes of Mount Kenya in the central area. The villagers belong to the Kikuyu ethnic group but many speak English fluently. Most have between two and five acres of land on which they grow mixed crops of bananas, maize, French beans and baby corn. They engage in various farming activities to minimise the risk of a falling market or low harvest, and are not afraid to learn and try new things to improve their practice at any cost. These farmers expressed the need to know when to plant, when to harvest, and how to obtain good prices for their produce.

Rainfall is fairly consistent in the area. A river provides sufficient water for all nearby farmers; it never dries up and its current is fairly strong. However, not all are making the best use of water. Most farmers use motor pumps to transport water to farms. The land needs to be irrigated two to three times a week during the dry season, and irrigating a plot of two acres requires up to six litres of fuel; this can cost around 1,500KES (about $21USD) per day. Poorer farmers do not collect rainwater for re-use, as in Figure 1.



Figure 1. Rainwater collection tank.

Few properties have electricity, but most members have mobile phones which can be charged at the local town or by neighbours with access to electricity. Some villagers have solar panels which supply small amounts of power. The nearest fixed access to the Internet is in the town 20 km away.

3.2 Kambu

Kambu, by contrast, is a small town in the south of Kenya, halfway between Nairobi and Mombassa. The area faces severe challenges including drought and environmental degradation, as seen in Figure 2. It is claimed that given enough rainfall, crops grow much faster in Kambu then in most other regions; however, drastic climate changes have included the delayed onset of rainfall, unpredictable weather patterns, and poor rainfall distribution. As a result the area around Kambu is one of the poorest parts of Kenya.

There is no sign of running water in the villages; tap water can be found only in town centres. In fact, water is so precious in this area that it costs more than fuel for motors. As a result, local farmers we visited do not irrigate or use other means to water their main crops, which include maize, millet, sweet potatoes, pumpkin, beans, pigeon peas, greens, arrow root, cassava and yam. Water needs to be pumped using expensive petrol, or stored in tanks to provide gravity feeds. The nearest spring water is 30 km away.



Figure 2. Typical dry field conditions in Kambu.

Villagers blame the lack of water for their condition and the decline of agriculture. Most do not use mixed farming which ensures food availability in all seasons. Consequently, two out of three families experience food shortages. Between November and June there is plenty of food; then there is nothing. Only those who plan well get through the year painlessly. People do not farm to sell but only to consume, so they are left with no food and no money to buy food in the dry season. Improved seeds for dry land, though they exist, are not affordable to even the most successful farmers here. There is a clear lack of knowledge of well-planned, safer farming.

Low literacy levels (about 50 percent) have exacerbated environmental degradation through destructive practices such as charcoal burning, sand harvesting, overstocking and overgrazing. High population growth has meant cutting down trees for settlement and charcoal burning. Thus this creates a destructive cycle as the land loses water and wild animals, and the soil erodes further.

The farmers have indigenous knowledge, but expressed the need for better farming knowledge. Since the land is dry, several Non-Governmental Organisations (NGOs) have shown interest and worked with the community, but these efforts have not always been successful. The farmers — particularly older women — continue to work the land, as this is seen as a source of independence and economic security. The group we met is well-organized and self-motivated. A system of national extension officers has supported local farmers, but this facility has been withdrawn recently.

3.3 Initial findings

Initial field research was carried out in 2007 in the both communities by researchers from the University of Nairobi and Thames Valley University who are versed to some extent in the local languages and cultures. The methods included evaluations of interfaces of existing mobile devices such as phones, cameras and iPods; a structured survey of technology usage and coverage; and ethnographic study of local agricultural and community practices among 36 adult farmers in Kambu, and 40 in Kiangwachi; in both communities the farmers are part of agricultural self-help groups. We discovered the following, commonalities:

- There is a lack of familiarity with digital technology such that most have no clear mental model of how many technologies work or how they could be useful.

- Agricultural knowledge sharing issues frequently come up. Since the communities practice mixed farming based on interest, it becomes difficult to find out everything to know about a crop before investing in it. Consequently, halfway through the process, farmers sometimes realize they do not have all the knowledge required.
- Literate and older professionals tend to record their farming activities (when planted, when to harvest, how much harvested, how much spent of particular crops, etc.). Others do not keep records and thus often cannot tell if they are operating at a profit or loss.
- Less than 10 percent of villagers comfortably read English. However, 50 percent in one village, and 90 percent in the other, indicate they have used, and like using, computers. On the other hand, less than 20 percent have heard of email or the Internet.
- Nearly all, however, have access to mobile phones. This use is almost exclusively for voice, not text. 85 percent are not aware of money transfer and other mobile services.
- Users indicate a preference in user interfaces for color, iconic legends, text, and representational icons, in that order. Numeric keypad interfaces fare poorly, implying a dislike for current mobile phone interfaces.

In their irrigation and farming practices, both groups practice low-risk strategies, but the less resource-poor area is more likely to experiment with new crops or new technologies. Given that water is precious in one area, and could be better utilised in the other, any technological intervention should seek to maximise the impact of irrigation while minimizing water intake. Monitoring of the environment and assessing the impact of temporal and spatial variability would seem to be a productive strategy. Both groups desire a two-way sharing of data and a simple means of recording their farming activities.

We originally set out to do participatory design directly with individual members of the communities. However, based on a Hofstede [2001] analysis, confirmed by our initial fieldwork and anecdotal evidence, we have learned that although the communities tend to be collective in nature, they have strong power structures in which organizations and institutions with formal boundaries and hierarchical structures are recognized and accepted. Therefore, the hardware is initially in the care of prominent individuals in the community, but it is important that it be viewed as collectively owned.

Thus it is more useful not to consider individual users but instead community-centered design. For example, mobile phone sharing is common. While technologically users are more comfortable with phones than computers, literacy patterns and usage suggests ruling out text-based communication such as SMS.

From the data collected, we developed the following use cases for a hybrid sensor network/mobile phone system:

- Calibration: A wireless sensor network must be calibrated when first deployed, and this should be simple and very clear, communicated by simple visual and/or aural means.
- Alert: When a parameter reaches a specified threshold, the system should notify the farmer, by similar means, locally or over the mobile phone network.
- Irrigation support: The system should give timely feedback to monitor different irrigation strategies as they are tested.
- Water requirements: If a computer model can be built for particular crops and local conditions, predictions could be made about water requirements and strategies.

4. SYSTEM ARCHITECTURE

From the data collected and use cases constructed, the following hybrid system has been specified. This will be deployed first at a test site in the UK, then in a demonstration plot in each village. In consultation with the communities, French beans will be grown in Kiangwachi, and in Kambu, watermelons will be grown along with Jatropha which has been selected by the community for its potential as a biofuel. Biofuels have become a controversial topic in recent months and have a complex relationship to sustainability and

climate change: High oil prices make biofuels such as Jatropha attractive alternatives, but high food prices mean that land which could be allocated to food production may be diverted to biofuels, as has happened in the U.S. thanks to heavy subsidies. The Kambu farmers are aware of the increased global demand for biofuels and therefore see Jatropha as a potential cash crop; however their interest at this time is merely experimental, since their first priority is growing adequate food crops in their harsh climate.

4.1 Sensor network topology

The network topology consists of collaborative sensor node designs that self-organise, exploiting small-world principles inspired by social and biological networks, to ensure that the distance between nodes remains small as the network scales. The topology also makes use of advances in routing and information theory, including an efficient network architecture that achieves full connectivity with minimum edges.

Due to the particularities of deploying wireless sensor networks for agriculture in such a resource-limited context, new research challenges must be addressed. For instance the sustainability of the network requires minimal hardware cost and, ultimately, autonomous operation, since no technical expertise will be available for maintenance. This, along with the fact that deployment areas are typically large, requires scarce deployment of nodes in a large field. This contrasts to the conventional 'dense' deployment of sensor nodes which is a typical assumption in sensor networks designs [Akyildiz et al, 2002]. Consequently the benefit of using node scheduling approaches for power conservation is highly limited. Furthermore, the radio range of the nodes should be large enough to guarantee connectivity. This, along with a consequent requirement for relatively higher sampling rates, puts a further burden on the power budget of the nodes.

To address this challenge, the inherent parameter-homogeneity of a typical agricultural field is taken into account. Using pre-defined homogeneous zones [Camilli et al, 2007; Konstantinos et al, 2007] a zone-based topology management algorithm has been developed to project the network topology on the field homogeneity map. This optimizes coverage of the scarce node deployment while avoiding unneeded redundancy. Additional inter-zone nodes are also considered to account for fault tolerance and reliability.

Another special design challenge is the reliability of the telecommunications infrastructure in such a remote community. Typically these areas are classified as unprofitable by telecom operators, hence limited capacity is allocated for them, especially for data communications (e.g., GPRS). Consequently the backbone link for the sensor network is highly unstable in terms of availability, bandwidth and quality. To address this challenge we are developing appropriate architectures and protocols, initially inspired by the Delay Tolerant Networking (DTN) concept [Ho and Fall, 2004]. These include a multi-sink architecture where the sensed data is aggregated in several nodes, as opposed to a single sink. These sink nodes are located along the edge of the field for automatic data logging when handheld devices pass by. This provides an alternative means of data collection.

Another ongoing effort is the development of a backbone/storage-aware routing protocol. The essential idea is that the sink(s) broadcast the backbone link status, along with the available memory capacity. Also the nodes broadcast their storage status. When the backbone link is not available and the storage capacity of the sink(s) is not appropriate, then the data will be routed within the nodes until circumstances change. The routing criteria would be balancing the overall 'storage space' of the network. To enable this, several storage management and asynchronous messaging approaches are considered.

Due to the power constraints in the field environment, new forms of power-aware routing of data are required in which the known locations of sensors aid routing so as to minimise the power needed. Routing heuristics are being developed to exploit minimum energy paths, and maximise the minimum power over all network nodes or avoid power drain at strategically positioned sensor nodes (e.g gateway/sink nodes). We define virtual node locations based on the power needed to reach them, rather than their geographic location, and derive appropriate routing algorithms. The goal is to achieve a successful balance between power consumption, acceptable field coverage and network connectivity while

employing scarce deployment of sensor nodes. We are exploring power-save modes of operation in the sensor nodes, and studying the impact of additional power consumption during wake-up sequences, leading to optimised power conservation algorithms.

4.2 Sensing subsystem

Soil moisture (volumetric water content) and soil temperature are sensed directly using the Decagon ECH2O probes, with an accuracy of ±%3 without calibration.

Weather data such as ambient temperature, rainfall and humidity information could be collected from local weather stations. However, according to the World Meteorological Organisation, the collection of weather data in Africa has gotten worse as many of the automatic weather stations it helped set up have fallen into disrepair. Therefore we utilise on-board temperature, humidity and ambient light sensors.

Inspired by U.C. Berkeley's Mica Motes, the sensors are interfaced to an Atmel Atmega128L microcontroller, with 512K bytes of flash memory and 4K bytes of RAM.

4.3 Communication Subsystem

For the sensor network, IEEE 802.15.4 (ZigBee) compliant RF transceivers are used for intra-network communications between sensor nodes. Bluetooth is also used by sink node(s) to broadcast data to any passing Bluetooth-enabled device. In this way, people (more precisely handheld Bluetooth-enabled devices) are treated as mobile data loggers in the network. Data is also automatically uploaded by sink/gateway node(s) via GPRS data service, since this is the only current option for network connectivity in the regions in which we are working. We are investigating embedded GPRS modules such as the Libelium GPRS/GPS module. Data could also be sent via SMS using an SMS gateway such as Clickatell which is used for mobile phone application development by the University of Nairobi. However, this is not a priority for us given end users' reluctance to use SMS due to literacy levels and cost.

GPRS is also used in the villages for Internet access, and the initial resource kit sent to the two villages included a laptop with GPRS modem and solar charging kit. The cost of GPRS data service is initially paid by UK-based project partners, and we are investigating sustainable solutions for providing ongoing service. We are also investigating WiMax technology for long-distance (up to 50km) communication [see Dekleva et al, 2007].

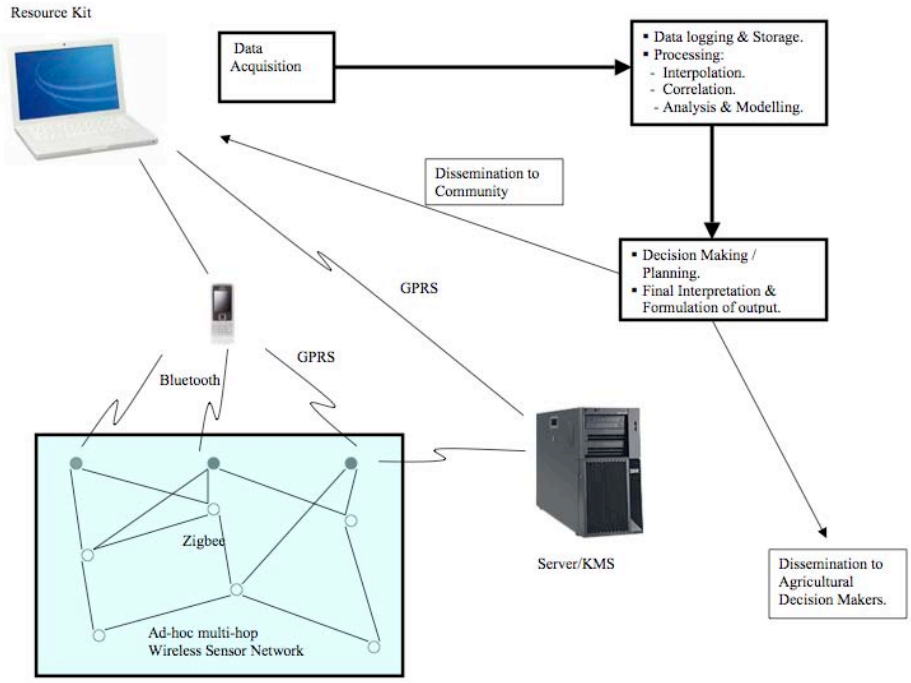


Figure 3. System architecture

4.4 Processing

Proper data processing is a critical element for the success of using technology in developing world applications, especially when considering a sensor-based system [Wang et al, 2006]. In order to build crop models, we need local weather data (including solar radiation, air temp, rainfall). Long-term records do not exist for the regions we are working in, so this data needs to be collected over time. We also need to record soil data, including depth, composition, density, carbon level, pH, aluminium, root abundance. Some crop-specific data can be gathered from existing crop modelling software such as APSIM [Keating et al, 2003]. However, packages such as these, having been created in and for the developing world, have usability issues in our context.

The sensor network can then help evaluate different water conservation measures delivered by the KMS. As mentioned previously, this system is a two-way exchange of information. A crop model which both informs and is informed by the sensor network can be complemented by an agricultural database for end users such as Infonet Biovision (www.infonet-biovision.org).

So for example, a user in one village can input a particular water management strategy to the KMS using a GPRS-enabled phone; this can be accessed and used by a user in another village. The sensor network there can automatically log the results of the intervention and feed them back into the KMS, making comparison possible between sites and over time.

Because users are treated as ‘mobile nodes’ in the system, user-generated data such as photos, text and audio interpretations can be uploaded and matched with sensor data. For example, a photo of the status of a particular crop can be matched with sensor data on that crop, providing a secondnary and more readily understandable ‘ground truth evaluation’ of the sensor data. As a later step, the feedback loop could be closed so that user-generated data documenting the results of an intervention can feed back to the sensor network to automatically generate water management strategies.

4.5 Data access subsystem

Data is collected, processed and served by the KMS which is built with the open-source Drupal content management system. Drupal Mobile is used to serve content to mobile devices. Usability is key, since as stated previously, many users have no ICT knowledge. Therefore real-time and archived data must be presented using simple interfaces. This development is a joint effort between the University of Nairobi which is undertaking mobile application and interface development, the University of Bradford which is working on server setup and administration, and the London Knowledge Lab which focuses on educational end use. The KMS is described in greater detail in Wirastuti, et al [2008].

The system is initially tested in a deployment at Kew Gardens, a large public botanic garden in London. This deployment will include interpretative data made available to the public both over the web and via Bluetooth broadcast locally in two test plots in which similar crops are grown.

4.6 Power subsystem

Individual nodes are battery-powered. A two-month typical battery life-time, for soil moisture sensing via sensor network, has been reported by Panchard et al [2006]. By optimizing the network design, as described previously, we intend to enable the batteries to cover, at least, the whole cropping season. Solar charging and power scavenging techniques are being investigated. A comparison between sites is planned, since for example UK has longer daylight hours during the summer, but with the sun at a more oblique angle.

5. CONCLUSION AND NEXT STEPS

We deploy technologies in a phased manner and with targeted activities designed to familiarize users with the technologies and the concepts underlying them. First, basic computing and network access is introduced; then, networking activities are carried out on mobile devices; then the sensor network adds the dimension of automatic data collection and ubiquitous computing. The basic computing and access resource kit, including laptop, GPRS modem and solar charging kit has been delivered to the communities in February 2008 and successfully received and tested. Mobile application development is being undertaken by the University of Nairobi and UK partners, and is expected to be deployed in the Spring. Meanwhile, the sensor network is to be deployed at Kew Gardens in March; after testing throughout the UK growing season it will be deployed in Kenya in the Autumn for the growing season there.

For ongoing sustainability we are working to keep costs down, and maintain agile development methods which are adaptable to change. Development is distributed and open source technologies are used to the extent possible. The communities continue to be involved in the design and evaluation at each stage.

ACKNOWLEDGEMENTS

We are grateful to the EPSRC for funding the VeSeL project under its Bridging the Global Digital Divide initiative (Grant number: EP/E007/198/1). We thank all our partners and contributors, and particularly the communities with which we are working.

REFERENCES

Akyildiz , I. F., W. Su , Y. Sankarasubramaniam , E. Cayirci, Wireless sensor networks: A survey, *Computer Networks: The International Journal of Computer and Telecommunications Networking* 38(4), 393-422, 2002.

Camilli, A., C.E. Cugnasca, A.M. Saraiva, A.R. Hirakawa and P.L.P Correa, From wireless sensors to field mapping: Anatomy of an application for precision agriculture, Computers and Electronics in Agriculture 58(1), 25-36, 2007.

Dekleva, S., J.P. Shim, U. Varshney and G. Knoerzer, Evolution and emerging issues in mobile wireless networks, *Communications of the ACM* 50(6), 38-43, 2007.

Economist, The. Kenya: Going up or down?, *Economist* (UK Edition) 9 June 2007. (http://www.economist.com/world/africa/displaystory.cfm?story_id=9304146)

Hartman, K., K. London, R. Bray and R. Faludi, Botanicalls: The plants have your number, http://www.botanicalls.com, 2006.

Ho, M. and K. Fall. Poster: Delay tolerant networking for sensor networks, 1st IEEE Conf. on Sensor and Ad Hoc Communications and Networks, 2004.

Hofstede, G., *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. Thousand Oaks CA: Sage Pubs, 2001.

Konstantinos, K. A. Xenakis, P. Kikiras and G. Stamoulis, Topology optimization in wireless sensor networks for precision agriculture applications, Proc. Int'l Conf. on Sensor Technologies and Applications, 526-530, 2007.

Panchard, J., S. Rao, T.V. Prabhakar, H.S. Jamadagni and J-P Hubaux, COMMON-Sense Net: Improved water management for resource-poor farmers via sensor networks, Int'l Conf. on Information and Communication Technologies and Development, 2006.

Wang, N., N. Zhang and M. Wang, Wireless sensors in agriculture and food industry: Recent developments and future perspective, *Comput. Electron. Agric*. 50, 1-14, 2006.

Wirastuti, D., R. Luckin, R.E. Sheriff, K. Walker, J. Underwood, and L. Dunckley, Development of a knowledge management system integrated with local communication channels and knowledge management initiatives for Kenyan rural farming, 5th Int'l Conf. on Wireless, Mobile and Ubiquitous Technologies in Education. Beijing, 2008.

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Graphical User Interface Design for Climate Impact Research Data Retrieval

M. Wrobel [a], **T. Nocke**[a], **M. Flechsig**[a] **and A. Glauer**[a]

[a]*Potsdam Institute for Climate Impact Research(PIK),*
Telegrafenberg, 14473 Potsdam, Germany
(wrobel/nocke/flechsig/glauer@pik-potsdam.de)

Abstract: Complex, heterogeneous environmental data sets are produced and stored in many scientific disciplines. For managing complex structured data, database and data warehouse technologies have been established, providing access via standardized interfaces. Nevertheless, there are some obstacles for researchers in applying the full functionality of sophisticated data management systems. Here, graphical user interfaces (GUI) are essential for bridging the gap between such systems and scientific users. However, designing / developing user interfaces for complex, environmental data sets is not trivial, facing user diversity and a variety of tasks. Consequently, for the example of climate impact research, this paper introduces an approach to support data retrieval. We propose tailored solutions for metadata navigation and filtering as well as for visualization and visualization design, and present three tools in combination with lessons learned for GUI design.

Keywords: Human Computer Interaction, Graphical User Interfaces, Data Retrieval, Climate Impact Research, Visualization

1 INTRODUCTION

Easy-to-use access to information from large heterogeneous environmental data sets is not straightforward. In particular in the multi-disciplinary field of climate impact research, providing appropriate tools for data retrieval is hampered by various obstacles, resulting from the heterogeneity of users as well as data and a dynamically developing research environment. Utilization of these data sets in an integrated manner requires the following two steps: (1) data integration and (2) provision of an appropriate design of human computer interaction (HCI) for data retrieval. In this paper, we focus on the second step. Therefore, we identify challenges that arise for graphical user interfaces (GUI) in climate impact research (sec. 2), outline related work (sec. 3), present our approach for data retrieval in our specific background (sec. 4) and resulting solutions (sec. 5). Finally, we discuss these solutions and the developed tools (sec. 6) and conclude with a summary and challenges for future work (sec. 7).

2 PROBLEM ANALYSIS

Designing intuitive access mechanisms to data in our context faces a variety of challenges, above all (i) data heterogeneity and complexity, (ii) the gap between required functionality and intuitive access to this functionality, (iii) user diversity and the variety of user tasks and (iv) the demand for smooth support of and integration into the users' ongoing research processes.

Data heterogeneity and complexity. First of all, data required for climate impact research refers to a variety of disciplines, has to be compiled from various sources, and can – due to the data complexity (e. g., different kinds of spatial grids, resolutions, formats) – not be managed using a single database scheme. Thus, comfortable integrated access is severely hampered.

Gap between required functionality and intuitive access to this functionality. Efficient access to data in our context may not be provided using a single standard software tool, but requires utilization and combination of functionality, partially provided by database management systems, geographical information systems, visualization systems and even programming languages. Since only very few users can be expected to be experts in applying all of these, there is a gap to be bridged between the required functionality, the intuitive access to it and the functionality provided by standard technologies and tools.

User diversity and the variety of user tasks. The HCI design in our context must be adaptable to heterogeneous users groups, including users with different skills and qualification grades (from students to senior scientists), with different retrieval goals (scientists, potentially also stakeholders and the general public) and from different disciplines (e. g. meteorology, hydrology, socio-economy). Moreover, these users are confronted with a variety of tasks, including gaining an overview of the available data sets, metadata and data navigation, filtering of subsets of interest, retrieval of derived / aggregated data and data exploration.

Demand for smooth support of and integration into the users' ongoing research processes. Identifying the required functionality needs input from many climate impact researchers. Their knowledge has to be incorporated into the design of data retrieval tools and appropriate interfaces without hampering scientists' ongoing, genuine research tasks. Here, a specific situation in our working context arises due to the fact that data sets required for specific research tasks are - for practical reasons - compiled and managed by individual scientists. This results in a locally available and quality-assured data landscape, which is still heterogeneous and steadily growing. Solutions for integrated access to these data resources have to achieve acceptance both from scientists that provide individually managed data, and those who need to access it. Here, well-designed functional graphical user interfaces are a driving argument to shift towards integrated data sources, competing with established work-flows based on proprietary data structures.

3 Related Work

Evolving information technologies are not easiliy adaptable to scientific requirements. For instance, relational databases that are the de-facto standard for data management in commercial applications, still have a low acceptance for everyday scientific data management processes (see Gray et al. [2005]). To reduce obstacles for managing heterogeneous data, material data integration based on data warehouses has been developed, facilitating flexible access to previously heterogeneous data (see e. g. Inmon [1996]). However, the resulting metadata and data landscape are still complex. Thus, concepts in human computer interaction concerning improved interface design (see e. g. Norman [1988], Shneiderman [1998]) are essential for reducing the gap between information technologies managing data and domain researchers.

In addition, visualization techniques have been established to filter, evaluate and explore data (see e. g. Bürger and Hauser [2007]). Besides a multitude of methods for spatial and temporal data representation (e. g. Aigner et al. [2008]), visual exploration techniques for multi-variate data (e. g. Bürger and Hauser [2007]) and metadata (e. g. Yang et al. [2005]) have been introduced.

4 Approach for data retrieval in climate impact research

Meeting the challenges outlined in section 2, our approach is to provide retrieval interfaces for integrated, data warehouse based data management structures, enabling intuitive access to metadata as well as to measured and simulated spatio-temporal environmental data sets. The approach is pragmatically-driven, balancing between user freedom and the task-related and computational restrictions. Our procedure is

1. **to develop the software iteratively in tight interplay with the users**: In our case, complete a priori specifications are usually not available or tend to be risky and failure prone. Instead, continuous and mutual learning of users and developers is required. Thus, we follow an incremental process model for software development and apply prototypes to involve users in this process.
2. **to identify required and attractive functionality and make it available intuitively:** we combine know-how from different information technologies and select/adapt suitable techniques for the specific requirements in our context (DBMS, GIS, HCI and visualization).
3. **to provide individual paths based on a modular interface design with recurrent interface elements:** Taking into account the given multitude of user abilities, preferences and tasks, we strive to support individual user approaches instead of forcing individuals to complete a rigid number of interaction steps. Additionally, due to the advantages of decomposing complex user tasks into simpler ones [cf. the OAI-Model, Shneiderman, 1998], we modularize our GUIs into subtask-specific elements that are easier to understand and handle. Moreover, recurrent usage of these elements – where appropriate – contributes to

both, increased consistency and reduced learning overhead.

4. **to choose interaction metaphors the users are familiar with:** although coming from different scientific disciplines, many users in our context are familiar with interaction paradigms such as form filling for data filtering or adjusting color maps for visualization. We build on such metaphors, enhance them if necessary and provide support for novice users.

As a result, we design portable, efficient, extensible object-oriented software (in Java and C++), providing easy-to-use and uniform access to climate-related data, supporting an accelerated, quality-ensured research process. Furthermore, as added value for data retrieval of simulation model output, we incorporate data analysis / exploration functionality.

5 Solutions for Climate Data Retrieval

In this section, we propose solutions for making the heterogeneous climate impact data sets easily accessible. We provide appropiate mechanisms to navigate through the available databased on metadata, and to filter and represent it. Therefore, we distinguish the following steps: (a) navigation through the data landscape, (b) filtering of data subsets, (c) selection of data representation technique, and (d) interactive visualization.

5.1 Navigation through the data landscape

To simplify access to the various available data sources we investigated specific GUI strategies providing overview and navigating through heterogeneous data sources. Here, besides the dialog-based, textual display of metadata, we use visual metaphors as a guide to improve orientation, recognition and comprehension of the provided data.

To access DBMS-integrated station measurement data compiled from world-wide sources (xDat tool), we built **a hierarchy of dialogs** that enables users to navigate through the different station databases and to review relevant metadata (e. g., type of database, content overview, see Fig. 1a,b). Color coding of the database types turned out to be a good orientation hint.

For a file based visual data exploration scenario (SimEnvVis tool), we developed a **tabbed-dialog based metadata representation**. Important types of metadata such as dependent variable ("results"), global, spatial and dataset-related metadata (cf. Nocke and Schumann [2002]) can be represented for several data subsets compactly (see Fig. 1i,j,k). Additionally, small icons are used to improve the comprehension of data/metadata. This includes grouping of dependent variables by subsets (color: subset, shape: the subset's dimensionality, Fig. 1k) and the illustration of subset specific metadata (grid and cell types, Fig. 1j).

Furthermore, to limit the amount of metadata to be displayed for the current task, we apply **task-dependent metadata presentation** for the retrieval of gridded climate data (PixDat tool, see Fig. 2). This greatly reduces the amount of data / metadata to be presented to the user, and thus significantly increases usability. Figure 2a represents the metadata required for displaying a single variable on a colored map, Figure 2b depicts the metadata relevant for visually comparing two variables (from one or two different data sets) and Figure 2d limits the represented metadata sets to data sets applicable for climate diagrams.

5.2 Filtering of data subsets

Filtering of data subsets to be retrieved is tightly linked with the navigation step described above. For instance, the data subset and variable selection in the file based data visualization tool SimEnvVis can be directly executed in the metadata view (see Fig. 1k).

However, for the retrieval of complex database managed data, both flexible and easy-to-use filtering is a severe challenge. To avoid user overload and to reduce complexity, we designed an extensible collection of configurable filter dialogs, each allowing the definition of specific query subtasks. A main filter dialog provides an overview of the filters available for the currently selected database, each filter's status (applied or not) and the currently selected filter criteria (Fig. 1c). Filters are presented on demand; users may freely choose their preferred filters in any order to define individual search criteria. This concept allows to provide GUI metaphors tailored for individual subquery definition tasks. The user may use interactive maps to define spatial restrictions intuitively (Fig. 1d) or hierarchical dialogs to select station types, depicted by both names and

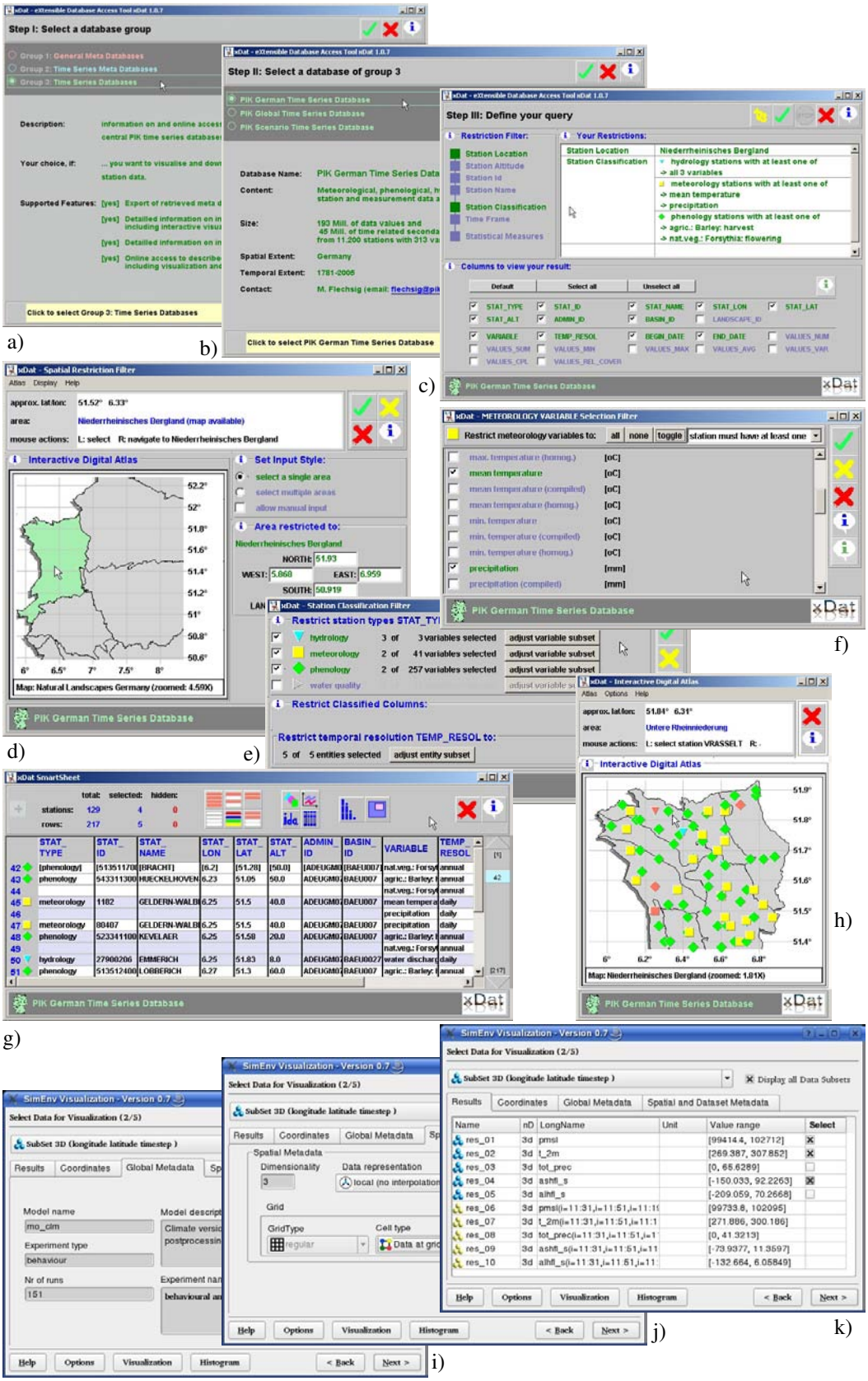


Figure 1: Metadata navigation (a,b,i-k) and data filtering (c-h) with the tools xDat (a-h) and SimEnvVis (i-k)

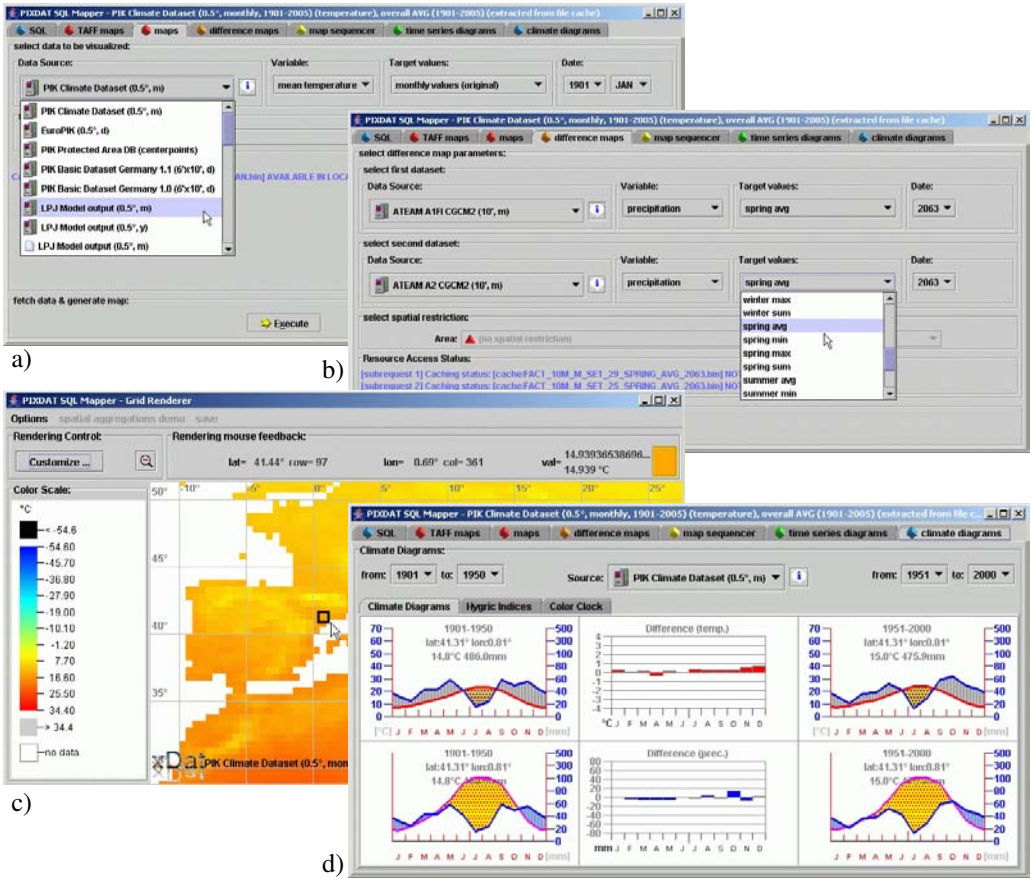


Figure 2: Task-based data filtering and interactive data visualization with the PixDat tool: a) single data set filtering; b) double data set filtering; c) map visualization; d) climate diagrams

icons, and according variables (Fig. 1e,f). To further reduce complexity, subsets defined using coarse filtering can be further filtered subsequently, e. g. using picking functionality on interlinked map visualization and tabular representation (Fig. 1g,h). Different structures / semantics of databases require flexible GUI adaptation. Thus, we provide configuration mechanisms assigning filters to specific databases and database attributes. For instance, picking an area on a map (see Fig. 1d) will transparently lead to different SQL statements depending on the stored geographical reference of the queried data (e. g. station locations or bounding boxes).

5.3 Selection of data representation

Considering the variety of visualization techniques available to represent the filtered data subsets, user support for selecting and parameterizing these techniques is benefical (e. g. Jiawei et al. [2004]). Thus – for the visual exploration tool SimEnvVis – a rule-based mechanism has been designed, which ranks techniques and supports users to launch them with appropriate starting parameters based on metadata, task specification, user preferences and available resources (Nocke [2007]; Nocke et al. [2007]). Figure 3 top-left illustrates a GUI list representation of automatically ranked visualizations available for the filtered data set, retrieving technique description and feedback about limitations of the technique for the current problem. Based on this information, the user can interactively (re-)select techniques.

5.4 Interactive Visualization of / on a data subset

Data retrieval strongly benefits from visualization techniques supporting different tasks. For intuitive retrieval of geo-spatial data it is essential to provide flexible, intuitive data selection and a first orientation about selected data using interactive maps (e. g. xDat and PixDat). Such maps provide a variety of interaction mechanisms (picking, rubber band selection, zooming), facilitating spatial filtering as well as intuitive comprehension of meta data / data in a geo-referenced manner (see Fig. 1d,h). In PixDat, where maps are used to represent measured or modelled data, we additionally provide a variety of color map definition mechanisms for filtering and data classification

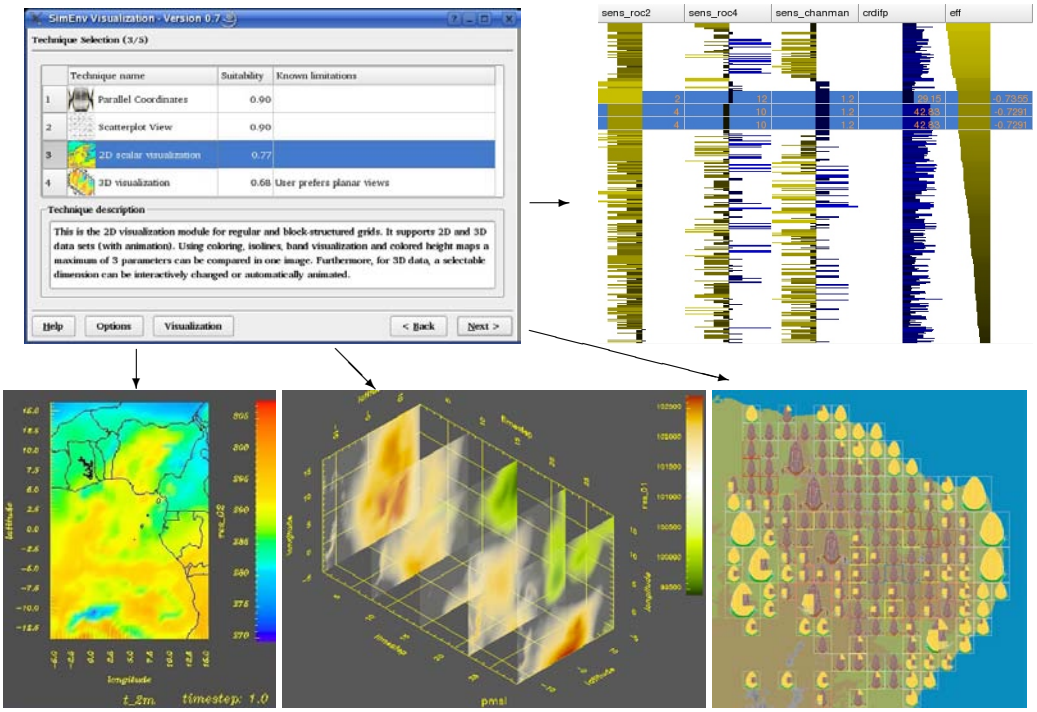


Figure 3: Semi-automatic visualization technique selection in SimEnvVis; top-left: list selection dialog; bottom-left: colored map; bottom-center: transparent slices tracking data extremes over space and time; bottom-right: icon based multi-variate visualization of maize cultivation conditions; top-right: graphical table

(see Fig. 2c) and a tailored visualization of aggregated annual climate conditions for a selected position (see Fig. 2d).

Otherwise, if the focus is on data analysis, a variety of interactive visualization methods are required, going beyond presentation oriented map-based techniques. Therefore, in SimEnvVis, visualization techniques for different data classes (1D, 2D, 3D, nD scalar data) have been integrated (see Nocke et al. [2007]). Besides providing standard techniques in this field (1D graph plots, 2D maps, ...), alternative visual metaphors for interactive multi-variate data exploration have been elaborated and integrated.

Due to the fact that scientific users appreciate direct feedback about the data values, we additionally provide tabular representations of metadata (see Fig. 1g) and data (see Fig. 3 top-right). Going beyond the presentation of textual values, we added a graphical mode, which maps the data values to bars inside a table cell. To explore multi-variate data with such a table, we provide cell shrinking, interactive re-sorting of rows and columns and displaying details for selected rows of interest. Thus, a large number of data records can be analyzed at a glance and be intercompared (see Fig. 3 top-right).

Furthermore, in the context of climate simulation data exploration, tailored spatio-temporal representations are required for different analysis tasks. Thus, besides the standard (animated) map visualization with color mapping (Fig. 3 bottom-left) and / or isolines, techniques comparing different time slices at a glance (Fig. 3 bottom-center) or comparing variables using metaphor-based multi-variate icons on a map (Fig. 3 bottom-center) have been successfully applied for interactive climate data analysis (see Nocke et al. [2007]).

6 Discussion

It is not trivial to decide whether certain sets of functionality should be integrated into one tool or rather be kept separate by designing many, less complex tools. Decisions should be made carefully, since the first approach leads to increased homogeneity as well as to increased application complexity and learning effort; the latter can lead to easier understandable, less complex applications, however, also to increased application heterogeneity. Avoiding both extremes, three applications integrating the steps outlined in the previous section have been designed and developed. These address various aspects/tasks for data retrieval for climate impact research and assign different interaction schemes for the outlined steps.

xDat (Wrobel [2004]) was designed to provide fast, integrated and direct access to a set of databases containing general and station specific metadata and a large variety of locally compiled time series, gathered from world wide stations and managed using data warehouse methods. The underlying interaction scheme we applied is abd (cf. sec. 5): The user starts choosing a metadatabase of interest (a) and is afterwards provided with flexible and intuitive filtering mechanisms for this database (b). Then, visualization (d), which is implicitly determined by the type of metadata to be retrieved, is used to represent and to perform further filtering on station metadata and to provide quicklooks for selected time series. The applied filter dialog concept (see sec. 5.2) is flexible enough to fulfil various retrieval tasks both for general and station specific metadata and sufficiently intuitive for novice users. The result representations provided – form based for general metadata, tabular interlinked with maps for station metadata – are adequate for typical result evaluation tasks occurring in this context.

PixDat, still under development, is the extension of xDat to access spatio-temporal data on structured grids with different resolutions, managed using a central database (Wrobel et al. [2005]). We apply the interaction scheme $cabd$: the user starts with selecting a task, which directly determines a visualization technique (c), resulting in a simplified navigation process (a) and subsequent subset filter criteria definition (b); finally, visualization (d) is used to present and evaluate the retrieved data. Due to the lower complexity of grid-based climate impact data sets compared to station data, here we apply compact, adaptable dialogs providing all relevant criteria at a glance, immediately adjusting to user selections. For instance, selecting a dataset influences the variables and temporal aggregations offered in the GUI. This immediate feedback supports orientation towards possible interaction options, allows for quick task definition and reduces interaction errors.

SimEnvVis is a visualization environment for the visual exploration of gridded, file-based climate data (see Nocke et al. [2007]). Using a wizard metaphor (interaction scheme $abcd$), it represents metadata of a given data source (a) and reduces the set of possible visualizations by interactive filtering (b) as well as by semi-automatically selecting available visualization methods and parameterizing them (c). Accordingly, a selected visualization technique can be started (d). At any time, users can short-cut this process by starting a visualization based on the current decisions, using standard values for later steps within the wizard. Thus, a flexible support mechanism generating entry points into complex visualization implementation environments (OpenDX, Ferret, OpenGL, Qt) has been developed, greatly simplifying the user's effort in generating images for a certain problem context. Furthermore, user profiling (e. g., distinguishing between novice and advanced users and incorporating preferences for certain visualization systems or metaphors) allows adjustment of the degree of information offered to the user and to individually steer the visualization design process.

For all these tools, our experience is that acceptance is increased if users do not perceive them as a competition to established expert tools they are familiar with. Rather, new applications should complement existing tools by combining a carefully chosen set of relevant 'best of different worlds'-functionality in both an easy to use and efficient manner. For instance, to support flexible query definition and a first intuitive result evaluation not requiring any GIS or SQL knowledge, xDat combines form filling with map interaction. Nevertheless, the user is free to export selected metadata and data directly to (further) explore it using his / her favorite expert tools. In a similar manner, SimEnvVis does not require any detailed knowledge about visualization networks or scripts, however, enables exporting them for later reuse by expert users.

To achieve location independent access to software functionality, *internet* technology has been established as de-facto-standard. Against this background, evolving *web* technologies play an important role for comfortable client deployment. GUI development in this context is still confronted with the well-known drawback between ease of deployment (thin clients) and rich functionality (rich clients). Thus, a retrieval system can benefit from supporting both rich and thin clients. For instance, PixDat can be accessed via a highly responsive Java Swing client (Fig. 2), while a subset of the functionality is also accessible via servlets.

7 CONCLUSION

We outlined our approach for interactive data retrieval, incorporating strategies for accessing heterogeneous databases intuitively as well as for evaluating and exploring extracted data. Therefore, we investigated GUI strategies to support metadata navigation and data filtering mechanisms ac-

companied by interactive visualization techniques. Thus, a variety of challenging tasks in the context of climate impact data retrieval have been essentially facilitated. In particular, we contribute a detailed problem analysis for our context, propose a tailored approach for these problems, describe its systematic application to the relevant interaction steps and discuss the resulting data retrieval and exploration tools.

However, there are still challenges for future work. First of all, this includes extending the proposed retrieval mechanisms, e. g. adapting them to highly heterogeneous socio-economic data. A further challenge is to provide a management mechanism for the rather complex filter and visualization parameterizations, improving later reuse and facilitating reproducibility of scientific results. Furthermore, it is necessary to investigate possible synergies of the GUI mechanisms introduced and of the related tools. Moreover, future work needs to investigate the opportunities arising from sophisticated metadata visualization (see e. g. Yang et al. [2005]) for the navigation and filtering of complex data sets.

ACKNOWLEDGMENTS

The authors wish to thank all the scientists at PIK that are testing and using the developed tools for their valuable support and hints. Furthermore, we thank Stefan Petri for LaTeX support.

REFERENCES

Aigner, W. et al. Visual methods for analyzing time-oriented data. *IEEE Transactions on Visualization and Computer Graphics*, 14(1), 2008.

Bürger, R. and H. Hauser. Visualization of multi-variate scientific data (state-of-the-art report). In *Proceedings Eurographics'07*, 2007.

Gray, J., D. Liu, M. Nieto-Santisteban, A. Szalay, D. DeWitt, and G. Heber. Scientific data management in the coming decade. Technical Report MSR-TR-2005-10, Microsoft Research, Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, January 2005.

Inmon, W. *Building the data warehouse*. John Wiley & Sons, Inc., New York, 2nd edition, 1996.

Jiawei, H., A. Bailey, and A. Sutcliffe. Visualization Design Reuse. In *Information Visualization (IV'04), London*, 2004.

Nocke, T. *Visuelles Data Mining und Visualisierungsdesign für die Klimaforschung (Visual Data Mining and Visualization Design for Climate Research), in German*. PhD thesis, University of Rostock, Faculty of Computer Science and Electrical Engineering, 2007.

Nocke, T., M. Flechsig, and U. Böhm. Visual exploration and evaluation of climate-related simulation data. In *Winter Simulation Conference (WSC'07), Washington D.C., USA*, 2007.

Nocke, T. and H. Schumann. Meta data for visual data mining. In *Proceedings Computer Graphics and Imaging (CGIM'02), Kaua'i, Hawaii, USA*, August 2002.

Norman, D. A. *The Psychology of Everyday Things*. Basic Books, 1988.

Shneiderman, B. *Designing the User Interface: Strategies for Effective Human-Computer-Interaction*. Addison-Wesley, Reading, 3rd edition, 1998.

Wrobel, M. *Multidimensionale, heterogene, visualisierbare Datenräume. Anforderungen, Entwurf und Implementierung einer adaptiven und interaktiven Schnittstelle für transdisziplinäre wissenschaftliche Daten im Kontext der Erdsystemanalyse (in German)*. PhD thesis, Freie Universität Berlin, 2004.

Wrobel, M., J. Glauer, M. Flechsig, and C. Rachimow. *xDat for regular spatial grids*. Status report, Potsdam Institute for Climate Impact Research, http://www.pik-potsdam.de/research/research-domains/transdisciplinary-concepts-and-methods/pixdat/xdat4rsg.pdf, 2005.

Yang, J., M. Ward, E. Rundensteiner, and W. Ribarsky. Value and relation display: A general framework for interactive exploration of high dimensional datasets. *IEEE Visualization Conference*, 2005.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

On the Design of an Immersive Environment for Visual Mining: Challenges and Results from an Early Study

E. Mozaffari, A. Seffah, A. Blakney and R. Naghshin
Environmental Software Technology Lab
Department of Computer Science and Software Engineering
Concordia University, Montreal, Canada

Abstract: Observational systems, with the complicity of Internet and information technology, have introduced new ways for collecting and disseminating environmental data. As a consequence, the analysis of increasingly large and complex environmental data sets has become a difficult and time consuming process which can hinder analytical comprehension and insight extraction for decision-making regarding humanity's environmental challenges. This paper introduces new concept of visual mining which presents the data in some visual form, allowing the 'human' to gain insight into the data, draw conclusions, and directly interact with the data. In addition, it discusses some principles toward VTab, a visualization tangible board to support visual mining as part of the whole environmental data lifecycle.

***Keywords**:* Visual Mining, Environmental data visualization, Immersive visualization.

1. INTRODUCTION

The visualization and analysis of large scientific data represents a very challenging task, especially in the earth, space and environmental sciences. The environmental data are complex, very large in volume, multidimensional, dynamic (time-varying in data and dimension) and consist of many parameters. It is important to take advantages of this flood of information, make sense of it and turn raw data into understandable information. Visualization provides an important means to handle the information for knowledge discovery processes. However, the visualization of data alone is not sufficient, human interaction with data should also be provided in a way that user can access and use the information.

Our goal is to move beyond "traditional" approaches to human-computer interaction to focus not only on how humans interact with machines, but on human access to and use of information. An important step toward meeting this goal is to assess existing technologies and develop new environmental data visualization methods and tools from a human-centered point of view that puts the emphasis on user needs. We believe that the most efficient way to achieve this is through a visual computing process that allows users to transform the raw data into useful knowledge and patterns via the use of visual representations and immersive interaction techniques to facilitate insight into that data.

This paper proposes some principles towards an immersive tangible visualization toolkit for the integration, interaction and visualization of large collections of environmental data. In addition to basic interaction and visualization tools, our goal is to provide data comprehension and visual mining features to support the whole environmental data engineering lifecycle. This includes: (1) the analysis of data, (2) complex phenomena and problems modeling, (3) design of solutions to overcome the problems or to understand the phenomena, and (4) decision

making in crisis situations. We intend to develop and validate new types of interactions that are appropriate to interact with different kind of environmental data.

2. FROM VISUALIZATION TO VISUAL MINING

Data mining has proven its capabilities to outperform traditional spatial analysis and statistical methods with better scalability, strong fault tolerance, and more importantly, the capability of inference/reasoning tasks (Miller et al., 2001). As a result, over the last decade, increasing efforts have been put into applying data mining techniques to handle the complexity of geo-spatial problems. Early studies (Koperski et al., 1996; Tung et al., 2000) show the capability of data mining to discover patterns in data. These efforts covered a range of spatial classification, association rule extraction, clustering analysis, outlier analysis and trend detection.
These efforts opened new possibilities for inductive spatial reasoning in the process of exploring both spatial and non-spatial attribute spaces. Geographers also realized potential advantages resulting from integrating data mining methods with geographic visualization tools to support human pattern-identifying abilities (Miller et al., 2001).
Visual data exploration aims at integrating the 'human' into the data exploration process by applying 'human' perceptual abilities to the large data sets available in modern computer systems. The basic idea of visual data exploration is to present the data in some new or more useful visual form, which allows the 'human' to get insight into the data, draw conclusions, and directly interact with the data. Visual data mining techniques have proven to be of high value in exploratory data analysis and they also have a high potential for exploring large databases (Keim, 2002).
Visual mining (Mozaffari et al. 2008) combines visualizations techniques and data mining. The goal of this area is usually to improve understanding of the data being presented. As 3D visualization techniques, such as Virtual Reality (VR), become more available and more commonly used, new and powerful visualization technologies, such as interactive and immersive visualization, are becoming an important part of the scientific process. Many tools and techniques are currently available to transform environmental data to its 3D representation and synthesize the data.
The integration and analysis of diverse data about various phenomena, in current Geographic Information Systems (GIS) and statistical tools, are usually represented as sampled or simulated data (e.g. on a regular or non-regular grid, or as point data). While GIS and existing tools are often used for integrating heterogeneous data, they still lack the capabilities to visually comprehend and link multivariate data while identifying relationships.

Building on data mining, we define visual mining as a process that:

- Aims at integrating the 'human' in the data exploration process with the goal to enhance their perceptual abilities to explore large data.
- Presents the data in some visual form, allowing the 'human' to gain insight into the data, draw conclusions, and directly interact with the data.
- Maps and transforms computer data into perceptual and visual representations that link factual (facts which can be reasonably assumed to contribute to the investigation, understanding, and solution of the problem) and singular (the available data) data while facilitating the overall comprehension of a large set of information.
- Lets the 'human' interact with and engineer data by manipulating visual representations, either as data objects or as objects representing relationships between data that are not necessarily visual themselves.
- Combines visualization techniques and the iterative cognitive process of data mining: (1) Filtering and analysis of data, (2) Modeling of complex phenomena and problems, (3) Design of solutions to overcome the problems or to understand the phenomena, and (4) Decision making.

As a cognitive information processing process, three major milestones characterize visual mining:

1) Gaining insight into singular data. Using different visualizations, the 'human' can build up an understanding of their data sets.
2) Relating factual and singular data. Besides visualization techniques, the 'human' exploits interaction techniques to directly interact and link the visualized data and navigate among links while relating data and facts.
3) Identifying patterns. Advanced data mining and visualizations techniques may be used for uncovering unknown patterns and relationships in large data sets.

Visual mining concept is similar to visual analytics (Keim et al., 2007) but it uses different visualization techniques simultaneously enhance users perceptual abilities to explore large data set and help them to relate actual and singular data in order to discover patterns in the large datasets. In addition, user can engineer data by manipulating visual representations, either as data objects or as objects representing relationships between data that are not necessarily visual themselves. In addition,

3. AN OVERVIEW OF THE PROPOSED INFRASTRUCTURE

Our objective is to provide an immersive platform for the integration, interaction and visualization of large collections of environmental data. A specific goal is the development and validation of a toolkit to support the visual mining process. The toolkit would offer interaction and visualization services to help users get insight into the data, relate factual and singular data, and discover patterns.

Figure 2 provides an overview of the proposed infrastructure. *Architecture of the toolkit* summarizes some of the services included in the toolkit. Table 1 contains a summery of descriptions of these services.

The *physical structure* portrays the proposed technological setup. The proposed visualization immersive tangible board (VTab) is built around a physical table that embeds a large interactive whiteboard and a second monitor to create a tangible workspace. Environmental data are represented by physical forms directly manipulable and perceivable by human sensors. Sensing based interaction mechanisms are used, allowing users to interact directly with the tangible forms on the large whiteboard. While exploring data, users will also be able to annotate data using, for example, real-time audio/video.

As it is shown in the *physical structure,* beside the interactive table, the environment consists of a ceiling-mounted video projector projecting computer generated world onto a white table around which the users sit. The users are equipped with an HMD system and electronic data gloves. They can be immersed into the virtual environment and they will be able to interacts with and engineer environmental data by manipulating visual representations. Sensing based interaction allows users to take advantage of their many senses and provides interactions that are closer to human cognitive processes. Also, this system includes a vertical display that shows information in more detail. It will be focused on the same position in the 3D model that user touched on the table.

We believe that VTab enhances comprehension and visual mining capabilities to support the whole environmental data engineering lifecycle. The *Visual mining process* in figure 2 shows visual mining milestones (described in section 2) which can be achieved using proposed approach. In order to validate proposed approach empirical user study should be conducted. The followings are steps that should be taken in order to perform the empirical studies:

1) Observe and describe a ***phenomenon*** which is an interaction between a human and a computer.
2) Formulate a ***hypothesis*** (research question) to explain it.
3) Use the hypothesis to ***predict*** (predictive model) or ***describe (***descriptive model***) other phenomena.***
4) Perform ***user study*** to test hypothesis.

Then the user study should be perfumed in a sequence shown in user study protocol (more details are given in section 4).

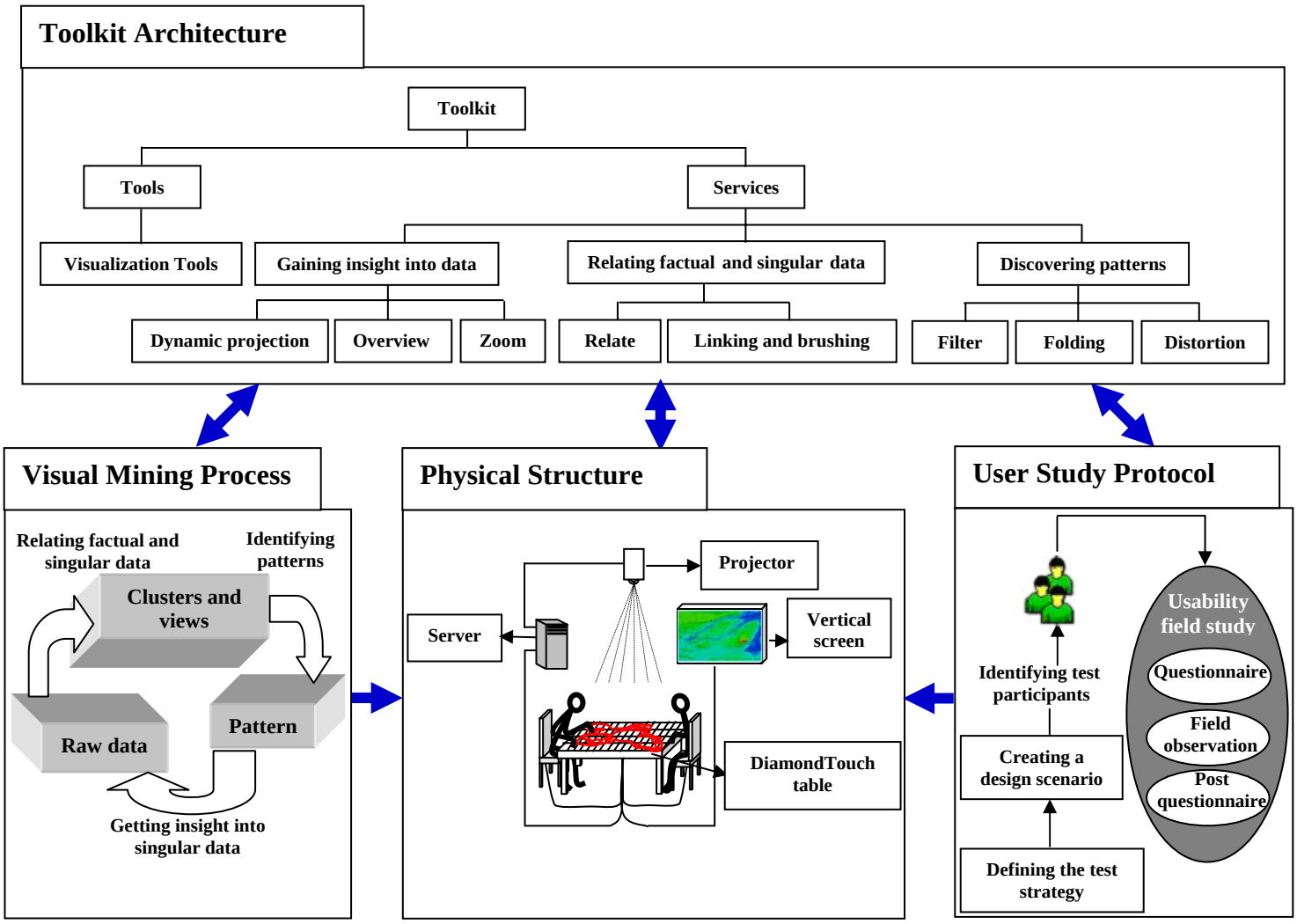


Figure 1. The proposed infrastructure

Table 1. Examples of services provided

Visual Mining Milestones	Tool	Description
Gaining insight into data	Zoom	Look at specific items of interest
	Overview	Gain an idea of the entire collection
	Dynamic projection	Dynamically change the projections
Relating factual and singular data	Relate	View relationship among items
	Linking and brushing	Multiple views dynamically linked, a subset of the data can be interactively selected and selected data portion will be highlighted in related views
Discovering patterns	Filter	Eliminate uninteresting items
	Folding	Pull a portion of the dataset: distant portions of the dataset can be positioned side by side for easier comparison
	Distortion	Show portion of data with a high level of detail while others shown in lower level of detail

4. RESULTS OF AN EARLY STUDY

For a better understanding of interaction techniques for visual mining, we conducted a field observation (Figure 2). Originally in this method the investigators observe the users in their natural environment while they are doing the related tasks (Nielsen 1993, Blomberg 1993, and Plowman 1995). In our study, we simulated the situation by creating a design solving problem and inviting users to participate in two three-hour tests sessions. After the field observation, we also asked users for their opinion about the tools they used. Observing the participants and analyzing the data gathered from the questionnaires and the interviews, we slightly modified our original physical setting.

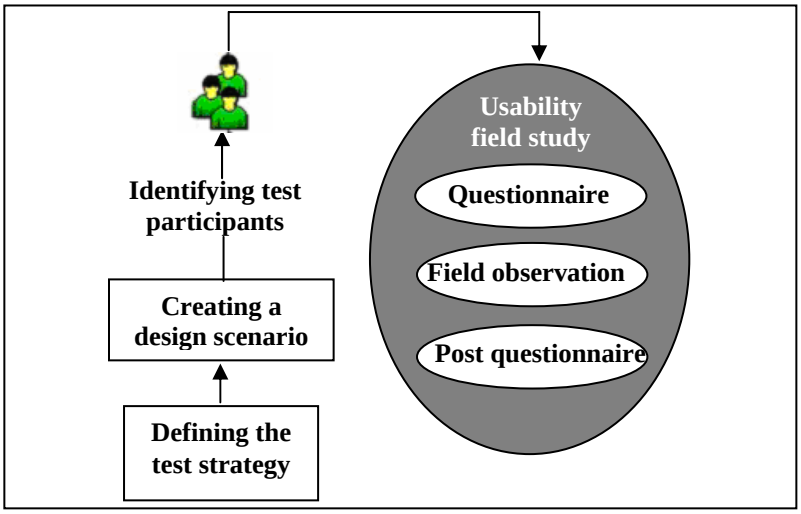


Figure 2. Protocol used in our field study

First, we used two computers (one user station one observer station) a digital camcorder, wireless microphone and the software Morae. Morae software was recording the participants'

interactions with the designers' station by recording the computer screen output, as well as mouse clicks and keystrokes. It was also recording the output from the digital camcorder and the microphone. It then combined everything into one editable data file. It also provided us with real-time remote viewing and annotation (across a computer network) of a recording in progress. We realized that one camera was not enough for recording all interactions. Therefore we modified our setting (Figure 3).

1. Digital camera capturing the user, using the computer which recorded via Morae
2. Overhead camera capturing the table from the top recorded via Adobe premier
3. Corner camera capturing the participants and the white board
4. Projector for projecting the computer screen on the white board
5. Design station consisting software design tools (the users' interaction with this station was recorded via Morae)
6. Design station consisting multimedia software
7. Observer station recording overhead camera and Morae
8. Observer station recoding corner camera
9. Microphone capturing participants' conversations
10. Observers
11. Participants
12. White Board
13. Dividing panel between monitor station and the design environment

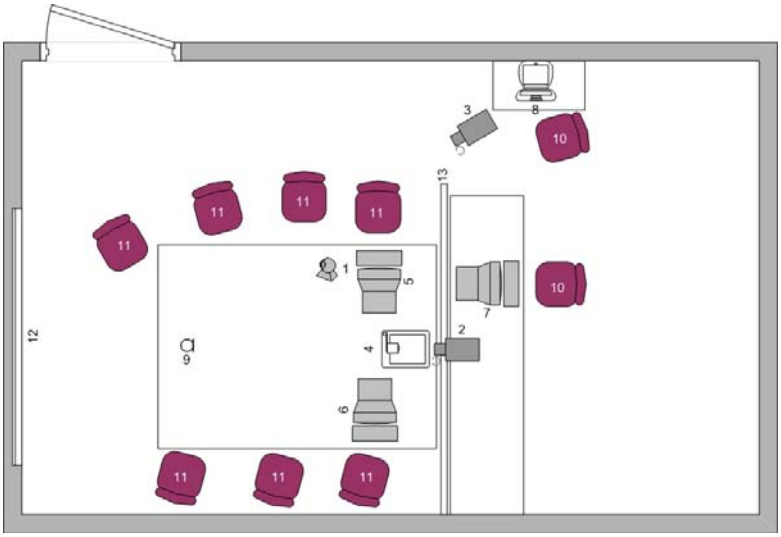


Figure 3. Study infrastructure

The tools that were provided for the designers in the first session included:

- Design stationeries (different pen, pencil, papers and markers),
- One computer including software such as Microsoft word, Microsoft Visio, and Photoshop.
- A white board with different color of markers for brainstorming,
- A projector was connected to the design station. The designers could project the monitor on the whiteboard.

Getting the feedback from the designers in the first session, we added more tools including:

- One computer as design station,
- Adobe Illustrators and Macromedia Flash for graphic designers,
- Rational Rose for software designers,
- Internet access.

The findings from this study are as fallows:

Software versus tangible tools. The most popular tools in both sessions were pen and papers and the whiteboard. Designers attempted several time to use computer software such as Photoshop for creating the interface and game characters or Flash for making a prototype. However, they found these software tools time consuming and impractical so they went back to pen and paper. One of the software designers, after producing his class diagram on the paper, asked monitors if this was enough or he had to redraw them with Rational Rose in order to make it more presentable. The only effective use of software was during the last hour of the first session when they were documenting their design decisions using Microsoft Word. They projected the computer screen on the whiteboard, and one designer was typing others' suggestions.

Whiteboard versus pen and paper. Both tools were used extensively in both design sessions and were rated as the most usable tools. As we noticed, pen and paper are used mostly for individual works or when a small group of designers work together, whereas the whiteboard is used during the design brainstorming. The large size of a white board allows a larger group of designers to present and explore the design ideas rapidly. However, a horizontal interface where the designers can sit and work for a longer period of time will be more preferable during the detail design.
Some of the requirements, we gathered from this field study are:

- The new environment should accommodate multidisciplinary activities using a large vertical interface as well as team activities using tabletop.
- Each user, representing a certain design discipline, would have their own horizontal table customized according to their design perspective.
- The tables would be connected to the hub which is connected to a vertical smart board, and the information should be shared among the disciplines by transferring any design artifact to the smart board. The information on the smart board could be modified during brainstorming by the whole team (Figure 4).
- The software used in this setting should help coordinate multiple points of view. Design experts from different disciplines bring different point of view, which should be converged in the final product. Early design tools must present and visualize design knowledge in a way that accommodates multiple design perspectives and goals.
- Using the existing CAD tools based on WIMP interaction for sketching, storyboarding, and prototyping is extremely expensive and time consuming. However, the disadvantage of making sketches using pen and paper is that these sketches are hard to modify and evaluate as the design evolves. The new setting should be based on tangible and sensor interaction. It will support natural interaction used for sketching and rapid prototyping, making the design iterations faster.

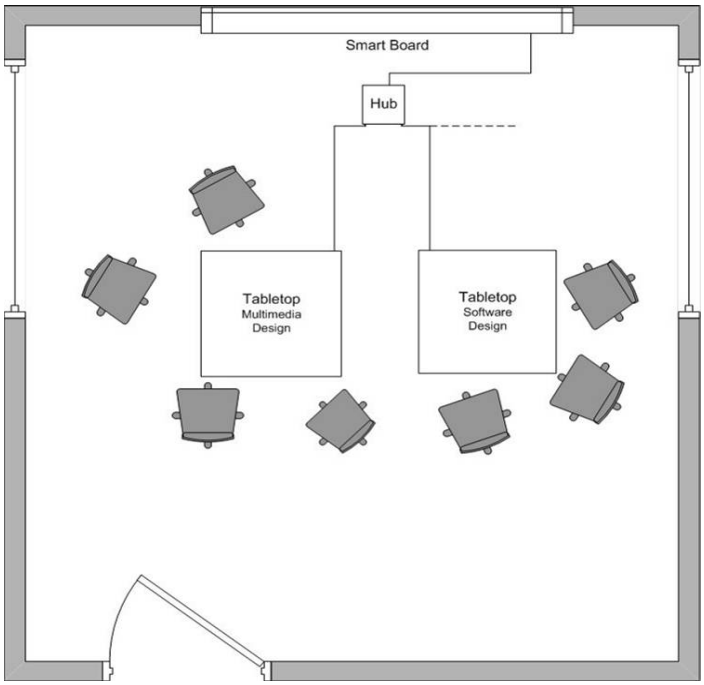


Figure 4. Layout

5. **CONCLUSION** and **CHALLENGES AWAITING FOR FURTHER INVESTIGATIONS**

This paper introduces an innovative visualization tangible board for enhancing comprehension and visual mining capabilities to support the whole environmental data engineering lifecycle. The proposed system includes visual paradigms to enable the human mind to process vast quantities of information, interaction paradigms to allow a higher order of interaction that is closer to the human cognitive processes, and physical devices to allow humans to take advantage of their sensory perception. In addition we described proposed framework. In the future, we intend to develop new types of interactive methods which are more appropriate for interactions with different kinds of environmental data, and later we will develop formal methods to assess the usability of these interaction techniques.
There are many research challenges that are associated with visual mining process. The following is a list of challenges in our scope of work:

1) Develop new types of interactions that are appropriate to interact with different kind of environmental data.
2) Perform empirical studies to test the effectiveness of these interactions. There are three themes:

- Raise the research questions: Formulate a hypothesis to explain the human computer interaction.
- Observe and measure: Gathering observation either manually (human observers) or automatically (computers, software, sensors, etc.). In addition the measures should be defined in a way that they fulfill the requirements.
- User studies: Perform user study to test hypothesis. The data should be analyzed to answer the research question.

6. REFRENCES

Keim, D. A., Schneidewind, J., 2007: Introduction to the Special Issue on Visual Analytics, *SIGKDD Explorations Special Issue on Visual Analytics*, 9, 2007.

Keim, D.A., Information visualization and visual data mining. *IEEE Transaction on Visualization and Computer Graphics*, 8, 1-8, 2002.

Koperski, K. and Han, J., Discovery of spatial association rules in geographic information databases. *In Advances in Spatial Databases, Proceedings of 4th Symposium, SSD'95*, Springer-Verlag, 47-66, 1996.

Miller, H. J. and Han, J., Geographic data mining and knowledge discovery. *Taylor and Francis*, London, 3–32, 2001.

Mozaffari, E. and Seffah, A., From visualization to visual mining: application to environmental data, *The First International Conference on Advances in Computer-Human Interaction*, 2008.

Nielsen, J., Usability engineering. London: Academic Press, 1993.

Plowman, L., Rogers, Y. and Ramage, M., What are workplace studies for?". *Proceedings of ECSCW 95*, 309-324, 1995.

Tung, A., Hou J. and Han, J., COE: clustering with obstacles entities, a preliminary study. *In Proceedings of 2000 Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'00)*, 3, 165-168, 2000.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Reclaiming Garbage: Interaction Design for Participation

Marijke Rijsberman
Interfacility (marijke@interfacility.com)

Abstract: Since the introduction of organized sanitation practices in the West, local governments and waste management business in the private sector have essentially adhered to an out-of-sight, out-of-mind approach to residential garbage, by which, for a modest fee, individuals are relieved from responsibility for their own waste and from awareness of its impact on the planet. Together with the systematic deferral of the environmental costs of current activities, this out-of-sight, out-of-mind approach has fostered individual behaviors that exert a great deal of pressure on the environment. Current environmental challenges and rising environmental concern now dictate a greater awareness of the impacts of individual choices. This paper proposes that interaction design can play a significant role in breaking through the garbage "taboo," helping to open up existing sources of data about waste in such a way as to stimulate individual participation in sustainable practices, such as waste diversion and source reduction programs. To demonstrate the relevance of interaction design to the challenge I use the example of the California Waste Stream Profiles, a database of information about municipal solid waste in California. The exercise delivers a basic set of interaction design principles to support individual sustainability that can be applied more broadly to sustainability at an individual level.

***Keywords**:* Waste; sustainability; participation; motivation; interaction design.

1. INTRODUCTION: DEFINING THE PROBLEM

Modern sanitation practices in the US were born in the waning years of the 19th century. Driven by a public health and nuisance abatement agenda and propelled by indescribably filth in the larger cities across the continent, these practices focused primarily on shielding individuals from their own garbage. Trash lay so thickly in the streets of cities like New York that it formed an impediment to effective transportation.[1] The mandate of the municipal collection services and private haulers of the time was primarily to make the garbage go away and secondarily to produce some income with which to defray expenses for such services as street-sweeping. Initially, almost all the garbage was reused—sorted and recycled, fed to pigs, boiled and rendered for fertilizer, burned to generate steam, and even used for fill to "improve" real estate values.

But this general approach of simply making away with residential garbage and not burdening the producers with the details is not relevant to the challenges of our time. We've come to think that real estate values are not improved by the addition of trash, in part because our garbage has become highly toxic with the addition of metals, plastics, and chemicals. Now garbage is simply buried and preserved in highly engineered environments that try to reduce the production of methane and leachate attendant upon biodegradation. In addition, the volume has swelled precipitously since World War II, while landfill space is becoming a scarce commodity.[2] Older landfills have been determined almost without exception unsafe in the last 20 or 30 years and have been closed at a much faster rate than

[1] Miller (2000) contains the most fulsome descriptions of waste in New York City, including one startling photograph taken late in the 19th century of a street where garbage lay knee-deep over the entire length and width of the street and sidewalks.
[2] Melosi (2004) refers to an estimate of remaining space of approximately 18 years, which means current landfill would fill up by 2022. More aggressive diversion goals in the past few years may have pushed that horizon a little further out.

new landfill is being created. Siting new landfill is extremely difficult, being fraught with environmental regulation and further hampered by popular aversion and NIMBY-ism.[3] In consequence, state governments are hard at work to reduce the amount of waste that ends up in landfill within their own state. Some of this political activity merely attempts to shift burdens, as certain states look to export as much of their garbage as possible, while states that still have sufficient landfill space are beginning to ban imports.[4] More permanent solutions are also pursued, mostly by setting significant waste reduction goals. One example is California's goal of 75% diversion from landfill by 2010. Local governments sometimes add aspirational goals of their own, such as San Francisco's zero waste goal by the year 2020. Meanwhile, it is worth noting that, despite increasing diversion rates, the average volume of garbage buried in landfill is still increasing year by year as the US population grows, while per capita the numbers have held more or less steady since 1992, as packaging material becomes ever more voluminous.[5]

Responsibility for achieving diversion goals is mostly in the hands of local governments, which look primarily to recycling to relieve pressure on dwindling landfill space. There are plenty of reasons to quarrel with the emphasis on recycling, as opposed to source reduction, but I will set aside such objections for the moment, to focus on the challenge of individual participation in existing programs. When it comes to recycling, limitations on waste diversion are not only posed by fluctuating markets for recovered materials, but also by residential and commercial participation in source separation. In many cases participation stays well below the ceiling imposed by markets, even under current conditions favoring the use of virgin materials.[6]

San Francisco's effort to collect compostibles separately from garbage that goes to landfill offers an instructive example of recycling limited by participation rather than market conditions. The program is voluntary and currently garners a 40% participation rate.[7] Not all residents who participate do so full-time.[8] NorCal, the waste management company that implements the San Francisco program, has no difficulty bringing the resulting compost to market at current participation rates. Without increases in participation, San Francisco is unlikely to reach the state-mandated diversion goal of 75% by 2010.[9] NorCal employs a team of people who go door to door attaching information to residents' garbage cans to alert them to the possibility of reducing their landfilled waste by separating out food waste. A modest financial incentive is offered, in the form of lower monthly rates for households that manage to drop a size in waste bins. NorCal does not believe it is possible to increase participation further without making it mandatory, as Seattle has opted to do. San Francisco recently announced an intention to make source separation mandatory and to impose fines for failures to comply.[10]

However, whether large-scale composting and other recycling programs are mandatory or not, the success of such programs largely depends on the willingness of the public to accept and pursue changes in behavior. For that reason, it is important to look beyond the

---

[3] Pellow (2002) traces the phenomenon to the 1980s, when "the environmental movement's mass mobilization precipitated a 'landfill crisis' by rendering the siting waste of facilities [sic] nearly impossible" (p. 54).

[4] Some of these efforts are clearly for present political gain and take little account of environmental imperatives or even common sense. For instance, North Carolina last year declared a moratorium on garbage imports even though exports have exceeded imports in that state for several years running.

[5] See Zero Waste America, Statistics, http://www.zerowasteamerica.org/Statistics.htm.

[6] Glass cannot always be profitably recycled, while the recycled paper markets swing significantly, so that increasing participation in source separation may not invariably result in increased diversion rates for these materials. However the recycle rate for aluminum, the most profitable recyclable material, lies near 50% nationally, while the market would easily absorb a 100% rate.

[7] Based on an interview with Bob Besso, Waste Diversion Program Director for NorCal, on January 31, 2008.

[8] In a recent qualitative study, several participants noted that they separate food scraps only about half the time, mostly for practical reasons (Rijsberman, 2008).

[9] Bob Besso, interview with the author, January 31, 2008.

[10] Decision of April 22, 2008.

motivational repertoire of (sometimes very modest) financial incentives and behavioral rules imposed from above. Proper alignment of financial incentives and convenient separation and collection systems is undoubtedly important in pushing participation rates, but neither are likely to drive the desired behavior by themselves.[11]

It follows, then, that the very success of the revolution in sanitation, by which residential and commercial waste essentially became invisible to those who produced it, now threatens the viability of waste reduction goals. It shields individuals from the knowledge of why their own efforts matter, knowledge that might help motivate their participation in environmental programs. Garbage is our most concrete—and, given current waste management practices, permanent—contribution to the environment. The question of what happens to it is not idle and can reasonably be expected to have an impact on individual behavior. Interestingly enough, individuals are asking this question online, submitting queries such as "what happens to our garbage" to the major search engines (Rijsberman, 2008).

2. EXISTING ONLINE TOOLS

How might interaction design be applied to this complex challenge of disclosing information about household waste in order to increase individual participation rates in environmental sustainability, such as San Francisco's food waste recycling program? The general goal would be to make available environmental information in such a way as to increase the (perceived) benefits and lower the (perceived) obstacles to the desired behavior. I propose that better disclosure of the fate of landfilled garbage, couched in highly concrete and individual terms—what happens to **your** garbage after you set it by the curb?—may contribute to increasing the perceived benefits of participation. That is, interaction design can help break through the garbage taboo by presenting information in a highly engaging fashion. At the same time, bringing to bear local communities through the use of social software applications offers a possibility of lowering the barriers to participation.

To demonstrate the possible deployment of such a two-pronged approach, I turn to an especially rich and public source of information about waste in California, the California Waste Stream Profiles maintained by the California Integrated Waste Management Board (CIWMB), the regulatory agency that oversees solid waste management in the state and holds first-line responsibility for implementing the state's waste reduction mandate. The California Waste Stream Profiles data is publicly accessible on the CIWMB website (www.ciwmb.ca.gov).[12] It contains a wealth of information about the volume, movement, and diversion of municipal solid waste at the town, county, and state level, as well as an exhaustive inventory of active and closed, legal and illegal landfills and other waste handlers. I chose this dataset for demonstration purposes because I have used it for years for my own research related to garbage and have found it invaluable (though often frustrating) for the information it contains—data about the fate of garbage in California that is directly relevant to individual consumers. However, the CIWMB does not appear to have realized that this is the case and makes the information available in a manner highly unlikely to encourage behavior change or even discovery.

The CIWMB does in fact address individuals on its website, because it clearly has an interest in their behavior. For example, some of the materials contained in a section called "Reduce Waste" are relevant to individual householders. Waste Prevention World (http://www.ciwmb.ca.gov/WPW/) is described as a "site about doing more with less, with information for individuals, businesses, and governments on a wide range of waste reduction topics." "Waste Reduction at Home" (http://www.ciwmb.ca.gov/WPW/Home/) offers

[11] Participants in the qualitative study referenced above cited a lack of convenience as a barrier (Rijsberman, 2008). Since no waste separation system can be as convenient as a pitching all trash into a single barrel, it follows that individuals require some motivation outside of convenience to adopt the desired behavior..

[12] The CIWMB is not involved with the work described in this paper. Attempt to speak to staff and management at CIWMB about the possibility of offering a consumer-friendly interface for the waste data have failed to elicit a response.

advice: buy in bulk, bring unwanted articles to re-sellers, avoid disposables, and so on. In principle, consumers can contribute to the Waste Prevention Information Exchange (http://www.ciwmb.ca.gov/WPIE/), although Web2.0 has not yet penetrated to the CIWMB and the contributor has to email information to a site manager for inclusion in the exchange.

The consumer-oriented resources have significant drawbacks. The information architecture is confusing to say the least, and all the information is extremely text-heavy. None of the standards of social marketing or instructional design seem to have been applied in the design. However, the most significant problem with the pages is that all the information focuses on what you can do to reduce waste, completely bypassing the issue of why you should bother in the first place. Although the CIWMB is sitting on a treasure trove of information about garbage and how it impacts the local environment, the materials that specifically address household waste don't offer the consumer any way in to that data. There is nothing about the numbers, the consequences to the California landscape, or the environmental challenges posed by landfills, in all their precariousness. The landfill and waste stream information is so separate from the advice and so hard to find, that no consumer without a special wish and a dogged determination to find it would ever run the waste data to ground. In other words, the CIWMB website carefully preserves the taboo on trash that's now been with us for more than 100 years. Luckily, the information to help us break through that taboo is on hand within the CIWMB California Waste Stream Profiles, even if it takes a great deal of perseverance to get to it.

Table 1: Source CIWMB's California Waste Stream Profiles.

Location (County)	Steps	Landfill (Town,County)
San Francisco	65	Altamont Landfill (Livermore, Alameda)
Antioch (Contra Costa)	47	Keller Canyon (Pittsburg, Contra Costa)
Bakersfield (Kern)	38	Bakersfield Sanitary Landfill (Bakersfield, Kern)
Red Bluff (Tehama)	14	Red Bluff Sanitary Lanfill (Red Bluff, Tehama)

I checked the clickpath for the city of San Francisco and several other communities, including a smaller town in the Bay Area, a large urban conglomeration in the Central Valley and a small agricultural town in Northern California. In rural areas, the garbage situation is relatively straightforward and the clickpath to the data is shortest (see Table 1). But it takes 38 steps to determine that the lion's share of Bakersfield garbage stays in Bakersfield, and the number rises further for communities in urban conglomerations.

Many of these steps require informed choices. If you don't know that Antioch is considered a jurisdiction, add a few steps. If you don't know that the destination of Antioch's garbage is hidden inside a Waste Flows tab which points to a different set of pages which provide a roundabout answer, add many more steps. And of course this only takes the consumer to a knowledge of where the garbage goes, which is not quite the same as knowing what happens to it.

3. GENERAL DESIGN PRINCIPLES

Of course, it makes little sense to lambaste the California Waste Stream Profiles for doing a poor job of something it didn't set out to do and has yet to discover the usefulness of. It will be more profitable to consider the basic design principles that could inform a consumer-oriented interface for the data contained in the Waste Stream Profiles. The first concern is to apply general best practices from interaction design, including a need to shorten the clickpath and to use terminology that aligns with lay usage. But if the objective is to motivate people to participate in waste reduction, then a variety of other design criteria will need to come into play that are fundamentally different from the challenges posed by an ecommerce application, a productivity tool, or a collaboration environment, to mention just a few standard examples. The existing literature on the relevance of interaction design to sustainability, insofar as I am familiar with it, approaches sustainability from the opposite direction, working towards sustainable products rather than towards sustainable individuals

making informed lifestyle choices.[13] As a starting point, I propose that, to motivate behavior changes based on environmental concerns, an individual needs

1. Knowledge about the problem at hand as well as the potential solutions
2. Engagement with that challenge through an emotional connection with the information
3. A viable sense of identity as part of the solution, primarily to be conveyed through a connection to the local community

3.1 Knowledge

In *Fostering Sustainable Behavior,* Dough McKenzie-Mohr points out that the most effective way to convey environmental information to motivate behavioral changes towards greater sustainability is a two-step process. The first step is revelation. This involves bringing home the impact of the problem, without pulling punches or mincing words. The second step is reassurance that concrete steps can be undertaken to solve the problem (McKenzie-Mohr, 1999, pp. 90-91).

In terms of this paper's chosen issue of household waste, the revelation is a challenge in framing the problem appropriately, and this means learning from things that have not worked in the past. The alarm about garbage was first raised by journalists in the 1980s, when the dearth of landfill in the Northeast precipitated a gradual transition to the exportation of garbage to other states west and south. The alarm died down again when the new patterns involving long-distance transportation were established. The eventual "solution" is in fact an environmental setback on three counts. Transportation of garbage across huge distances is associated with a significant increase in the use of fossil fuels and output of emissions for every ton of garbage landfilled.[14] Moreover, exports of garbage to poor and remote areas negatively affect the balance both of environmental justice and also of unspoiled natural assets. While recycling was introduced in many places following the so-called "garbage crisis," the total per capita production of garbage has nevertheless risen significantly. In other words, it is apparent that framing the problem as an issue of landfill ***space*** does not lead to behavioral changes that make a positive difference.

It is more appropriate to frame the problem in terms of individual responsibility for the production of astonishing amounts of garbage, which is toxic and permanent, and which under current conditions ends up in a real place, as a persistent embarrassment in the landscape. In the case of California, most municipal solid waste in fact ends up within state boundaries and often in the same county or a neighboring one, in places that are immediately recognizable and meaningful to individuals. The Waste Stream Profiles can make the difference here. Following the admonitions of McKenzie-Mohr, it is also important to add specific suggestions for making a difference—such as participation in food waste separation and becoming a more discriminating consumer with respect to such things as product packaging and durability.

3.2 Engagement

I use engagement here in a double-edged way, as a sense that certain problems have direct relevance to oneself, as well as "a feeling of belonging, an experience of investment and ownership in the local, regional, national, and/or international political communities to which citizens belong," as Wikipedia defines civic engagement.[15] What matters in terms of engagement is the concrete and specific in a local context that manifests itself as a community. This makes it important not to talk about a national or global challenge that is easily overwhelmingly difficult to comprehend and solve. While the larger environmental challenges are not irrelevant, the focus is appropriately on the local community and the individual's personal impact.

---

[13] See for example Fogg (2002) or the literature review on Debra Lilley's Design Behavior website http://www.lboro.ac.uk/research/susdesign/design-behaviour/what_is_it.htm.

[14] Benjamin Miller, currently at work on a study of transportation, shows that solid waste is trucked extremely long distances, often passing many active landfills along the way. Email communication with the author, October 24, 2007.

[15] http://en.wikipedia.org/wiki/Civic_engagement

The local community is also important in the sense that individuals don't usually make changes in patterns of consumption and handling of their own waste products in isolation. People often learn in community, as became clear in my qualitative study of participation patterns in San Francisco (Rijsberman, 2008). One participant explained that she separates food waste because "it is the right thing to do for the environment. Our preschool (a progressive co-op) started composting 4 years ago and so I was able to learn through that the vast amount of things that can be composted, and it has made a big difference in the amount of our trash." McKenzie-Mohr points out that social marketing has been shown to be more effective when mediated through direct contact: "Community-based social marketing draws heavily on research in social psychology which indicates that initiatives to promote behavior change are often most effective when they are carried out at the community level and involve direct contact with people" (McKenzie Mohr, 1999, pp. x-xi).

The principle of engagement, then, suggests that personal responsibility and local community both need to be addressed in the challenge of designing for environmental change. Space for collaborative problem-solving and community building will both help support the ultimate goal of designing for individual behavior change.

3.1 Identity

People who do participate in source separation programs invariably say they do it for environmental, not financial reasons, even if they sometimes mention the cost savings as a nice perk (Rijsberman, 2008). A significant sub-theme is formed by a reflection on identity, a thought about who the participant is and who he or she would wish to be. One man explained that "I don't want to have to think of myself as the kind of person who just leaves my trash to future generations." The perspective of future generations, as a lens on current behavior patterns that brings identity issues into focus, is mentioned with some frequency. "My kids are going to look back at us and wonder how we could do it," noted one participant, "just like we look back to our parents' generation and think, how could they just throw their trash in the water?" (Rijsberman, 2008).

Advertising of consumer products has been very successful driving consumers in the opposite direction, creating desirable identities based on ownership of those products. But keeping up with the Joneses is not just a mechanism that drives to ever-increasing consumption. It can also work in the opposite direction. In a different field, the obesity study that came out in the *New England Journal of Medicine* in 2007 also underscores the conclusion that individual behaviors develop in a social context (Christakis, 2007). Possibly there is something to learn from the techniques that have been employed, as the field of social marketing attempts to do. Perhaps demystification of these manufactured consumerist identities will also be helpful. But first and foremost, the challenge for interaction design is to engage the community in articulating and creating an alternative identity.

While it is clear that a significant number of people, at least in the vicinity of San Francisco, do not wish to be personally implicated in an environmentally oblivious and even environmentally predatory identity, the alternatives may not be especially appealing. One participant in the study referenced above remarked that when she first awakened to the seriousness of the environmental challenges facing the planet, she didn't want to be perceived as an environmentalist. That identity had been contaminated, in a similar vein as the identity of "feminist" was contaminated in the 1970s and 1980s, by an idea of a certain strident unreasonableness, beyond the pale and outside the discourse of one's own community. This suggests that part of the interaction design challenge is stewarding a process of discovering a viable alternative. If we don't want to be the people who preserve their leftovers in gigantic, unstable piles (sometimes in or near their own communities, but preferably in someone else's), so that they can indulge in unbridled consumption, then who do we want to be and how do we get there? Since we don't yet have good answers to that question, the best option is to facilitate the process by which the answers may be articulated.

3. SKETCHING AN ALTERNATIVE DESIGN

It is possible to imagine many different designs of a consumer-oriented interface for the California Waste Streem Profiles that answers the above design parameters. Preliminary sketches for an alternative interface for the Waste Stream Profiles data (available online focus first of all on a simple two-step search that shows the user where garbage produced in a particular community ends up.[16] Unlike the existing interface, the consumer-oriented interface sketch assumes that landfill is of interest to the general public, both as destination of the garbage individuals produce and as elements in the landscape they inhabit. Search result provide the same data that are available in the California Waste Stream Profiles, but it does so in a single step. The resources now offered by the CIWMB in an independent area of the CIWMB website are envisioned as displayed in the context of the search. Perhaps more importantly, the alternative interface sketch envisions the presence of video to explain to a lay audience how the principles of modern sanitary engineering are applied in landfill construction and what that means for the garbage we produce.

The alternative sketch also assumes that the information about facilities now made available by the CIWMB (information about operations, environmental studies, permits, etc.) is best provided in the form of a wiki, to invite engagement and participation in the process of producing the knowledge of what happens to garbage (see Figure 7 of the alternative interface sketch). The wiki offers users the opportunity to add a layer of oral history and visual memory, as a form of recuperation and reanimation of a near-invisible aspect of our society. Such a wiki seeks to encourage a stronger sense of place and a greater awareness of the environmental cost of individual behaviors. The intention behind the presentation of landfill-related information is primarily one of description, providing a kind of inventory of what exists and how people relate to it rather than an environmental analysis or a cautionary tale with an axe to grind.

Both the Landfill Wiki (Figure 7) and the Community Compass (Figures 8 and 9) are in service to the engagement and identity principles. The collaborative wiki approach to the landfill information makes it possible for users to take ownership of the pages and to add their own stories and photographs to the general, more impersonal information supplied by the CIWMB. The Community Compass component makes an effort to connect users to others in their own community who have an interest in environmental issues and in changing their own behaviors to reduce their environmental load. Social software is used, then, to enable people to better understand and negotiate their local geography and to connect to the local community. Figure 9 offers a general impression of how a map-based search interface might tie online interaction to a sense of community, by enabling users to find like-minded people in their own neighborhood.

A way to present oneself to the rest of the community in terms of environmental aspirations, competencies, habits, and questions, is also envisioned, to support a sense of identity as an environmental actor. One might imagine here some system by which environmentally-motivated interactions with the website (the posting of events, the participation in discussions, and so on) as well as data from the individual's shared in the individual profile could all contribute to the person's status in the community, which in turn can be visualized in the map. Discussion capabilities, meanwhile, would help support a collaborative community-based learning process.

3. SKETCHING AN ALTERNATIVE DESIGN

The design of such an alternative is crucially dependent on two other processes: on-going user research to support an iterative design cycle on the one hand and stewardship of the community aspects on the other. The proposed alternative interface sketch would need to be validated, adjusted, and further developed in an iterative design and testing cycle before being built, to determine its likely effectiveness as an illustration of the design parameters that would bridge between environmental information and the sustainability of individual

[16] Available at http://www.interfacility.com/experiencewriting/garbage.html), Figures 5 and 6. Comments are welcomed in Rijsberman, "What Happens to Our Garbage," at http://landfill.wordpress.com/2008/02/03/what-happens-to-our-garbage/.

behavior. Further research would appropriately focus on the validation of the proposed design principles as well, leading in the longer term to the development of best practices for interaction design in the service of individual sustainability.

REFERENCES

Christakis, A.N. and J.H. Fowler (2007), "The Spread of Obesity in a Large Social Network over 32 Years," *New England Journal of Medicine*, 357, July 26, 2007, pp. 370-79.

Fogg, B.J. (2002), *Persuasive Technology: Using Computers to Change What We Think and Do*, Morgan Kaufman.

Lilley, Debra, Design Behaviour: Reducing the Social and Environmental Impacts of Use through Design, http://www.lboro.ac.uk/research/susdesign/design-behaviour/index.htm.

McKenzie-Mohr, D., and W. Smith (1999), *Fostering Sustainable Behavior: An Introduction to Community-Based Social Marketing*, New Society Publishers.

Melosi, M. (2004), *Garbage in the Cities*, 2d ed, University of Pittsburgh Press.

Miller, B. (2000), *Fat of the Land: The Garbage of New York—The Last Two Hundred Years*, Four Walls Eight Windows.

Pellow, D. N. (2002), *Garbage Wars: The Struggle for Environmental Justice in Chicago*, Cambridge, MIT Press.

Rijsberman, M. (2008), "Motivations of Recycling Behaviors: A Qualitative Study," www.interfacility.com/experiencewriting/motivation.html).

Rijsberman, M., Reclaiming Garbage: A Design Sketch, http://www.interfacility.com/experiencewriting/garbage.html

Rijsberman, "What Happens to Our Garbage," http://landfill.wordpress.com/2008/02/03/what-happens-to-our-garbage/

Zero Waste America, "Statistics," http://www.zerowasteamerica.org/Statistics.htm.

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

An Analysis of Interface Quality for Decision Support

Timothy Maciag [a] **Daryl H. Hepting** [a]

[a]*Department of Computer Science, University of Regina, 3737 Wascana Parkway, Regina, CANADA (maciagt@cs.uregina.ca, dhh@cs.uregina.ca)*

Abstract: This paper examines the quality of user interfaces in an application to choose products. A high-quality interface should enable users to perform prescribed tasks quickly and without errors. The interface should also enable users to find products which match their personal values. A high-quality interface can especially empower those diverse users who might otherwise be overwhelmed by technology but who need information to ground their actions. By examining user performance in such regards, insight into the quality of interface may be acheived. For this purpose, we examined user performance while utilizing different interfaces for product selections. An analysis and discussion of the examination is presented, along with opportunities for future work.

Keywords: decision support; usability; user evaluation

1 INTRODUCTION

Over the last few decades, many authors have discussed usability and utility in terms of interface design [Grinstein et al., 2003]. If a tool is not usable but has utility, users may not utilize the system to its full potential. If it is usable without utility, people will not come back. In this regard, Rosson and Carroll [Rosson and Carroll, 2002] describe three levels of designing user interfaces with various degrees of usability and utility, these being *effective, comprehensible, and satisfying* with *satisfying* being the level that interface designers strive to achieve in their designs. A satisfying user interface could be one that encapsulates the notions of effectiveness, efficiency, and comprehensibility.

When considering design solutions for the development of user interfaces which provide satisfying user interactions, the diversity of end-user may play an essential role. As such, an interface should be designed to incorporate aids for both beginners and experts alike, as outlined by Raskin's [2000] *myth of the beginner-expert dichotomy*. In such regards, Fischer [Fischer, 2001] describes *high-functionality applications* (HFA); those which enable beginners to learn the system over time while allowing experts to harness the full potential of available resources.

Specifically we are interested in designing satisfying HFAs in the form of decision support tools. Decision support tools, which include software tools that assist end-users' in tasks involving decision formulations, seek to aid end-users by highlighting key information relevant to the task, providing a medium for comparison, aiding query formation, and by providing an interface between the user and system data [Phillips-Wren and Forgionne, 2006; Pu and Chen, 2005]. To paraphrase Richard Hamming [1962], "the purpose of decision support is to easily find the product that matches one's values, not to easily find some product which does not interest me." For this purpose, we consider designing *augmented* [1991] HFAs which provide a balance between the capabilities of the end-user and core system functionality. This paper presents work towards these goals.

The rest of the paper proceeds as follows. Section 2 describes the decision support tools used. Section 3 describes the methods used and the metrics collected during two evaluations for product decision support systems. Section 4 describes the analysis and interpretation of the study data. Section 5 provides conclusions and areas for future work.

2 TOOLS FOR EVALUATION

For our analysis, we conducted an in-depth usability analysis of two support tools, one developed by the United States Environmental Protection Agency (US-EPA) and the other developed by us called *cogito*. Both tools provided varying interface designs for searching a database of 29 environmentally preferable cleaning products distinguished by eight environmental attributes: *skin irritation*, *food chain exposure*, *air pollution potential*, *product contains fragrances*, *product contains dye*, *product uses recyclable packaging*, *product is a concentrate*, and *product minimizes exposure to concentrate*. We examined the quality of the user interface in terms of measuring user effectiveness and efficiency, hypothesizing that higher levels of each may indicate greater satisfaction of use.

2.1 US-EPA Tool

The US-EPA tool consisted of three searching tools: a single attribute ranking tool (SART), a multiple attribute ranking tool (MART), and a weighted attributed ranking tool (WART). SART enables users to filter products using one attribute only. Users choose a single attribute and are able to sort the attribute's values in terms of those that are most environmentally preferable to least (ascending order). MART, illustrated in Figure 1, enables users to filter products using two methods: by assigning priority of up to four of the attributes and/or by setting thresholds for specific attributes. WART enables users to filter cleaning products by assigning weights to specific attributes. Users assign attributes weights by choosing numerical values between 0 and 1000[1]. Higher weights indicate greater importance. Although the functionality of the search/filter procedure differed for each US-EPA EPP search tool, the results displayed by each tool used a tabular format.

2.2 Cogito

Cogito [Hepting, 2003], illustrated in Figure 2, is an implementation of an augmented HFA. End-users are provided the functionality to scan and filter items within an information space. The interface is comprised of rectangular cells displaying system objects based on end-user queries. *Cogito's* built-in querying functionality equips end-users with the basis to form queries by selecting system attributes and applying thresholds accordingly.

To evaluate the *cogito* system with respect to the US-EPA tools we created two *cogito* interface representations, a textual interface which displays products using textual data in the form of HTML, and a graphical interface which displays products using a nightingale rose illustration, see Figure 1. Each interface representation utilizes the functionality of the *cogito* system similarly; only the visual representation of the products differ.

3 EXPERIMENT DESIGN

For our analysis, we conducted two experiments. The goal of the first experiment was to evaluate the three decision support tools provided by the US-EPA in order to formulate which of the tools provided users with the most adequate support. The goal of the second experiment was to evaluate the *cogito* interfaces with respect to the best of the US-EPA tools.

[1]The range *0 to 1000* was the one originally provided by the US-EPA and was not changed for the purposes of the evaluation

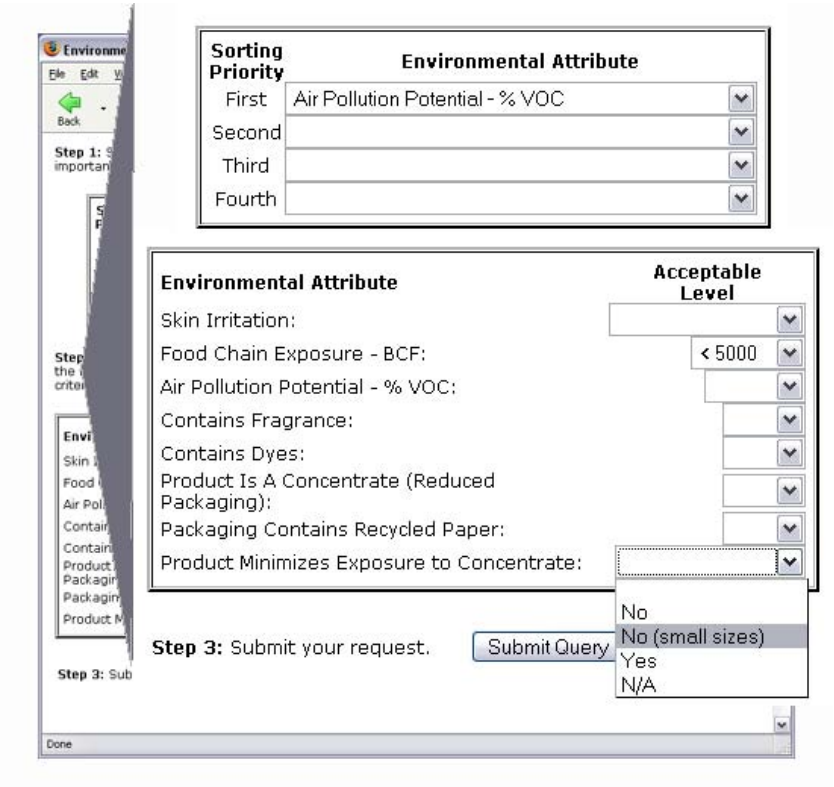


Figure 1: A screen capture of the US-EPA MART. The main interface is magnified. Users may filter cleaning products by setting priority of up to four attributes. Users may also chose acceptable attribute value levels.

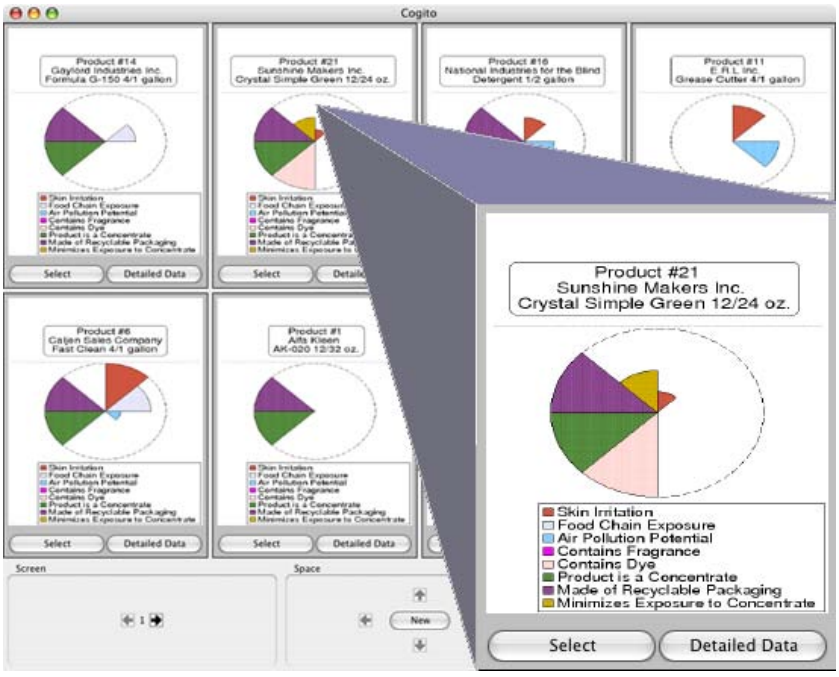


Figure 2: Screen capture of the graphical *cogito* interface. A cell on the interface is magnified.

In each experiment, 28 participants (11 female in the first experiment and 13 in the second) were recruited from the University of Regina Computer Science Participant Pool.[2]. In both experiments participants were asked to complete a pre-task questionnaire, answer a series of task questions, and complete a post-task questionnaire. The task questions were designed to evaluate the quality of the tool interfaces in terms of task complexity. People who use these tools are interested in finding products that match their values. The question of how well people can decide based on values is a more difficult task, because of the potential mismatch between user's model and the categories available. To some users, only one attribute might be important, while other users might need two, three, or more attributes to make their decisions. These preferences may also change over time. The authors created question sets to represent this variable complexity. There were a total of three sets, each with questions dealing with 1, 2, and 3 attributes.

To analyze user satisfaction, we recorded user response times and computed user task scores. Task scores were computed (Eqn 1) by determining the absolute and relative values of participant scores in relation to their responses. Eqn 2 was used to measure the efficiency of users. Based on the equation, participants who performed well on the interfaces in both task score and response time obtained a higher efficiency score. Task score results were squared to account for the occurrence of a high user task score matched with a *lengthy* response time and vice versa. It is also important to note that in order to discredit any bias in our final results, the interfaces were counterbalanced accordingly.

$$TaskScore = (1 - (\frac{\#CorrectAnswers - \#CorrectGiven}{\#CorrectAnswers})) \times \frac{\#CorrectGiven}{\#AnswersGiven} \quad (1)$$

$$Efficiency = \frac{TaskScore^{\,2}}{UserResponseTime} \quad (2)$$

4 Evaluation and Results

Results from both experiments are described in Table 1.[3] When observing results from the first experiment for tasks involving one attribute, users performed more adequately using SART whereas for tasks involving two and three, users performed more adequately using MART and WART. This also coincided with the participant preferences indicated in the results obtained in the post-task questionnaire analysis, where we asked participants to reveal their preferences in terms of interface. Based on their responses, for tasks involving one attribute SART was more preferred, whereas for tasks involving more than one attribute MART was more preferred. Based on these results, for our second evaluation incorporating the *cogito* tools, we developed a comparison structure of SART versus the textual *cogito* interface and MART versus the graphical *cogito* interface.

Observing the results in Table 1, it is clear that the users performed more adequately using the *cogito* interfaces. Figures 3 and 4 illustrate the user efficiency metric. Again, there is a clear indication that users performed more adequately using the *cogito* tools with respect to performing the prescribed tasks.

We analyzed results of our experiments using a 2×3 ANOVA (analysis of variance), comparing SART with the textual *cogito* interface and MART with the graphical *cogito* interface. F and p-values are reported to illustrate any significant patterns, interactions, and/or trends within the data.[4]

[2]The University of Regina Computer Science Participant Pool enables undergraduate students enrolled in computer science courses at the University of Regina to participate in evaluations similar to that described in this paper. Students receive up to 2% course credit as reward for their participation.
[3]Analysis conducted using SPSS.
[4]As F-values increase, p-values decrease. P-values less than 0.05 show, with 95% confidence, that there are significant differences between the estimated marginal mean values of the analysis data

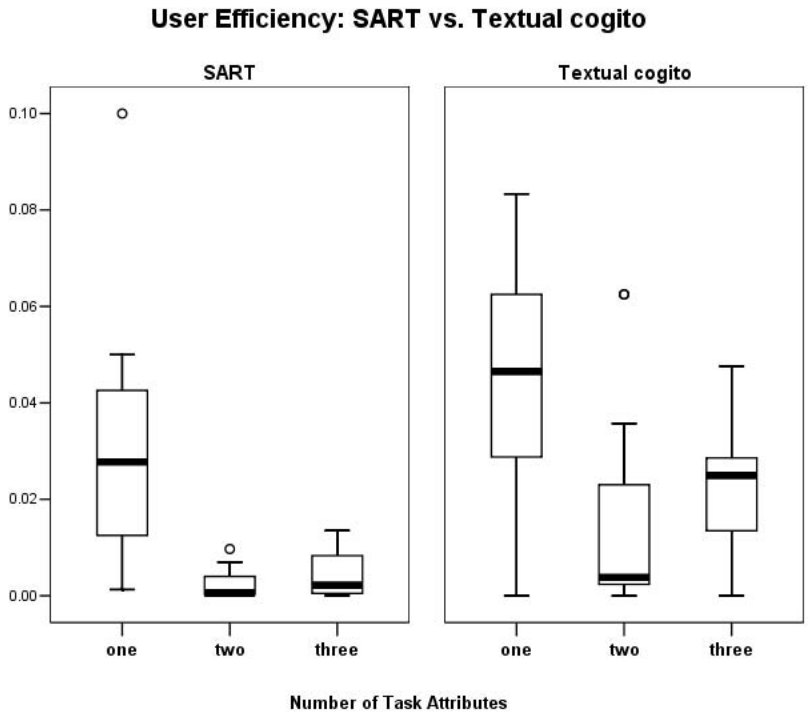


Figure 3: Illustration of user efficiency for the SART and textual *cogito* interface. The illustration provides a comparison of the data accumulated from both experiments. Here, both interfaces were assigned the same question set.

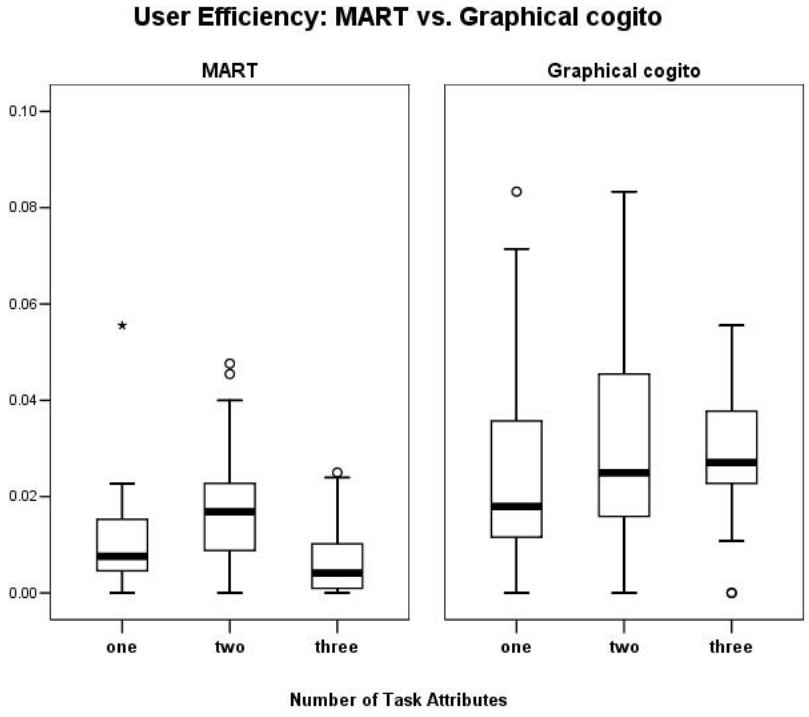


Figure 4: Illustration of user efficiency for the MART and graphical *cogito* interface. The illustration provides a comparison of the data accumulated from both experiments. Here, both interfaces were assigned the same question set.

Table 1: Experiment Results: Mean and Standard Deviations for User Response Time and Task Score for the US-EPA tools and the *cogito* interfaces (Colours are to indicate the same question set and are used to aid comparison)

	Time (sec)		Score (0-1)			Time (sec)		Score (0-1)	
Int.(Attr.)	$\bar{x}$	σ	$\bar{x}$	σ	Int.(Attr.)	$\bar{x}$	σ	$\bar{x}$	σ
First Experiment Results, US-EPA suite					Second Experiment Results, *cogito* Interfaces				
SART(1)	32.58	24.03	0.81	0.33	Text.(1)	24.00	11.58	0.96	0.19
MART(1)	83.18	50.65	0.74	0.35	Graph.(1)	50.89	33.66	0.86	0.30
WART(1)	61.68	36.80	0.61	0.49					
SART(2)	50.39	23.07	0.26	0.22	Text.(2)	34.04	29.66	0.51	0.38
MART(2)	58.36	35.83	0.79	0.42	Graph.(2)	35.11	15.04	0.79	0.42
WART(2)	60.21	29.38	0.48	0.42					
SART(3)	123.50	98.05	0.54	0.40	Text.(3)	42.12	15.85	0.88	0.28
MART(3)	103.57	81.64	0.60	0.39	Graph.(3)	37.93	16.32	0.92	0.40
WART(3)	68.07	44.27	0.65	0.48					

4.1 SART vs. Textual Cogito

Comparing the results from Table 1 for SART and the textual *cogito* interface, it appears the textual *cogito* interface provides users with better support. In terms of user response times for SART versus the textual *cogito* interface, there were significant interactions for both the task complexity (F = 35.13, $p \leq 0.000$), with linear relationships, and interface (F = 19.28, $p \leq 0.000$). In terms of user task scores for SART versus the textual *cogito* interface, there were also significant interactions for both task complexity (F = 68.29, $p \leq 0.000$), with non-linear relationships, and interface (F = 20.60, $p \leq 0.000$). Thus, results indicating the textual *cogito* interface being superior to SART were verified

Apparent from results in Table 1 and Figure 3 is a sharp drop in both SART and the textual *cogito* interface for the two-attribute question set. Upon further analysis, it became clear that users had to perform additional tasks in order to fully answer this question. By the definition provided by the GSA of *air pollution potential*, an air pollution value of *N/A* means that the product is exempt from materials and ingredients that have potential to harm the air. Thus, a value of *N/A* would be the same as having 0% air pollution potential. When observing the data, it was shown that 21/28 (75%) of participants who used SART did not include these products. As well, it was shown that 17/28 (61%) participants who used the textual *cogito* did not include these products. Here, we observe the crucial role between attribute terminology and end-user comprehension and how it reflects upon interface quality.

4.2 MART vs. Graphical Cogito

In terms of user response times for MART versus the graphical *cogito*, there were significant interactions for both the task complexity (F = 22.46, $p \leq 0.000$), with non-linear relationship, and interface (F 20.64, $p \leq 0.000$). However, in terms of user task score, there were no significant interactions in terms of task complexity (F = 0.01, p = 0.91) and only marginal significance in terms of the interfaces (F = 5.81, p = 0.02). Although this was observed, given the overall results described in Table 1 and the depiction of user efficiency in Figure 4, the graphical *cogito* interface provides the users with marginally better support.

Furthermore, with respect to participant preferences for interface design, the US-EPA versus *cogito* interface, the *cogito* interfaces were the most preferred consistent across all tasks, as described in [Maciag and Hepting, 2008].

5 CONCLUSION

This paper described how user performance could be used to measure the quality of user interfaces for decision support tools. Two experiments were conducted and results of each were matched and analyzed. The evaluation highlighted the need to provide users with a high quality interface for information seeking and searching tasks. The experiments provided indication that the *cogito* interfaces provided better support to participants. In terms of user performance, the *cogito* interfaces were shown to be successful. Some of the issues that remain include system terminology, e.g. attributes with values as *N/A*. Future work will include researching ways to provide personalized user interfaces for decision support tools. Some work in this direction has already been done [Maciag and Hepting, 2005; Maciag et al., 2005]. More adequately understanding user preferences will help personalize the user's search experience, highlighting those features, attributes, and values that the user prefers thus, improving their overall user experience.

6 ACKNOWLEDGMENTS

We wish to thank Don Sharpe at the University of Regina for his helpful comments on statistics. The authors also gratefully acknowledge the support of the Centre for Sustainable Communities at the University of Regina and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES

Fischer, G. User modeling in human-computer interaction. *User Modeling and User-Adapted Interaction*, 11:65–86, 2001.

Grinstein, G., A. Kobsa, C. Plaisant, B. Shneiderman, and J. T. Stasko. Which comes first, usability or utility. *In Proc. IEEE Visualization Conference*, pages 605–606, 2003.

Hamming, R. *Numerical Methods for Scientists and Engineers*. McGraw-Hill, 1962.

Hepting, D. H. Interactive evolution for systematic exploration of a parameter space. In Dagli, C. H., editor, *Proceedings of ANNIE 2003*, pages 125–131. American Society of Mechanical Engineers, 2003.

Kochhar, S., J. Marks, and M. Friedell. Interaction paradigms for human-computer cooperation in graphical-object modeling. *In Proc. GI*, pages 180–191, 1991.

Maciag, T. and D. H. Hepting. Analysis of user classifiers for personalization of environmental decision support system interfaces. *In Proc. ANNIE*, 2005.

Maciag, T. and D. H. Hepting. On the role of reflection and representation in environmental decision support systems. *In Proc. International Environmental Modelling and Software Society (iEMSs)*, 2008.

Maciag, T., D. H. Hepting, and D. Slezak. Personalizing user interfaces for environmental decision support systems. *In Proc. Rough Sets and Soft Computing in Intelligent Agent and Web Technology*, 2005.

Phillips-Wren, G. E. and G. A. Forgionne. Aided search strategy enabled by decision support. *Information Processing and Management*, 42:503–518, 2006.

Pu, P. and L. Chen. Integrating tradeoff support in product search tools for e-commerce sites. *In Proc. EC*, pages 269–278, 2005.

Raskin, J. *The Humane Interface: New Directions for Designing Interactive Systems*. Addison-Wesley, 2000.

Rosson, M. B. and J. M. Carroll. *Usability Engineering, Scenario-based Development of Human-Computer Interaction*. Morgan Kaufmann Publishers, 2002.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

On the Role of Reflection and Representation in Environmental Decision Support Systems

Timothy Maciag[a] and Daryl H. Hepting[a]

[a] Department of Computer Science, University of Regina, 3737 Wascana Parkway, Regina, Saskatchewan, Canada, S4S 0A2 {maciagt, dhh}@cs.uregina.ca

Abstract: Donald Norman describes how reflection enables us to critically analyse and review details, compare and contrast situational outcomes, and aid in our general decision-making abilities. Furthermore, he explains how through reflection, we increase our awareness - become smarter - which inevitably enables us to conduct more satisfying decision-making analyses. In a day and age where information is abundant, the activity of reflection may prove more difficult. This is particularly the case for evaluating alternatives for health and environmentally preferable product selection. Key in supporting consumers in such regard is the design of the user interface, one where the interactions provide satisfying user experiences through support for reflective activities supplemented by high quality representations. This paper will discuss the importance of reflection and representation in such regard by describing a framework for system design. A detailed description of the framework is provided along with a discussion describing qualitative results from a recent usability evaluation. Future work is also provided.

Keywords: Environmental decision support systems; Decision-Making analyses, Human-Computer Interaction; User-Centred interface design.

1. INTRODUCTION

Reflection is a mental activity we commonly apply in many of our *higher-level* (*non-experiential*) decision-making analyses [Norman, 1993]. As such, reflection could be considered a catalyst in facilitating our decision-making tasks as it provides the necessary cognitive activity which enables us to comparatively analyse criteria and concepts within a problem domain while effectively enabling us to conduct compensatory and non-compensatory decision-making analyses and evaluation(s) [Hoyrup, 2004; Blandford, 1991]. Reflection, and by extension – *reflective activity* [Courbasson, 2006], stimulates our decision-making processes, thereby enabling us to arrive at solutions based on our own experiences and understanding of the decision task [Hoyrup, 2004].

However, reflective activities can prove to be difficult given our own cognitive limitations [Norman, 1993]. Too much reflection and nothing would be accomplished whereas purely experiential activities may lead to poor decision solutions. Furthermore, left to our own devices, our ability to truly reflect and act upon a decision solution may be limited – as Norman [1993] and Fischer [2005] state: "*The power of the unaided individual mind is highly overrated.*" From this statement comes the realization that effective and satisfying decision-making may require the use of external decision aids. These aids could include low-level tools such as a piece of paper and a pencil but may also include higher-level tools such as those in the form of computer-aided support tools (decision support systems (DSS)). These higher-level tools may be preferred as they have potential to more effectively provide a more seamless interaction between the decision-maker and the data being analysed.

1.1 Usability in Decision Support Systems

When designing DSSs, a critical requirement in the design process is to define how best to develop an effective framework for user interaction. Norman [1993] suggests that we can achieve a high degree user satisfaction through the *power of representation.* Here, representation could refer to many aspects of design – including the user interface display, how system objects are represented on the interface, and the interface functionality, among others. The power of high quality representations cannot be understated. Consider the commonly misinterpreted proverb: "*A picture is worth a thousand words*,[1]" or its satirised version given by McCarthy "*1001 words is worth more than a picture*.[2]" – here we are provided with a deeper insight into what Norman is suggesting. In either case, regardless of their intended meanings, it becomes more clear that a representation in any form has the potential to greatly assist our decision-making abilities. Higher quality representations may ensure that decision-makers have more satisfying reflective experiences as they stimulate reflective thought – evoking a deeper exploration, which may empower the decision-maker to achieve more satisfying decision solutions [Norman, 2004]. However, designing quality representations is not a definite process – as what has meaning to some decision-makers may not have meaning to others. This task becomes even more difficult depending on the intended use of the DSS. For example, consider consumer-oriented DSSs – here, there may exist a variety of decision-maker – some with expert experience but also some who have limited understanding of the decision criteria. In this instance, more care is needed when designing the system representations.

Norman [1993] attempts to aid in this regard by describing what he believes constitutes a quality representation – being one that captures the critical aspects and decision criteria as it is viewed in the *represented world* (the "real-world") and correlating it with a depiction or illustration of the criteria as it would appear in a *representing world* (an abstraction of the "real-world" – e.g. a visual metaphor [Norman, 1988]). Here, the representation would only depict the necessary aspects of the represented world, while omitting all non-crucial aspects of its understanding. As described by Rosson and Carroll [2002], when designing such representations, it may useful to utilize the concepts of *realism* (realistic depictions) and *refinement* (more abstract depictions), as illustrated in Figure 1, which depicts two illustrations of the concept of recycling.

Rosson and Carroll [2002] indicate, in previous user studies, people have had high success in relating to realistic imagery such as that depicted on left-hand side in Figure 1. However, they add that in similar studies, people have also related well to refined imagery,



Figure 1. Two separate illustrations depicting the concept of recycling. The depiction on the left[3] illustrates the represented world (realistic) whereas the one on the right illustrates the representing world (refined)

[1] http://www2.cs.uregina.ca/~hepting/research/web/words/history.html (Accessed March 2008)

[2] http://www-formal.stanford.edu/jmc/sayings.html (Accessed March 2008)

[3] Image from: http://www.uoregon.edu/~recycle/housing_kitchens_text.htm (Accessed March 2008)

such as that depicted in right-hand side of Figure 1. In either case, they conclude that designers need to consider the *process of recognition* – meaning that in some instances, it may be that people take less time to cognitively process refined representations – as realistic imagery tends to be more complex in nature. As illustrated in Figure 1, it may be that more people will immediately understand the intended meaning of the refined image of recycling as opposed to the realistic illustration as it is unclear of the true activity being performed – *is she recycling? - or putting things in the trash? – or picking certain items out from each bin?*

We hypothesize that the role of reflection and representation is paramount in the design of any DSS. In this sense, successful frameworks for DSS design would enable decision-makers to effectively conduct reflective activities founded by their interactions with high-quality representations. Specifically in the research described in this paper, we are interested in studying these concepts in relation to environmental decision support systems (EDSSs). In the case of EDSSs, reflection and representation may play an even more critical aspect of system design given the unique nature of such systems, as will be discussed in the proceeding section.

1.2 Environmental Decision Support Systems

Swayne et al. [2000] define an EDSS as "*an information system containing at least one component whose purpose is to support human decision-making about an environmental issue.*" For the purposes of this paper, we focus on particular EDSSs for environmentally preferable purchasing. This type of EDSS is unique in that many of its potential users may not have specialized training in the decision domain (e.g. expertise levels may range from uniformed-less experienced user to the informed-expert). As such, information and the way it is represented is a critical factor in the underlying success of these kinds of EDSSs. In this regard, the usability of such systems is still an important determinant of their success [Frysinger, 2003], but the design's success may also be closely correlated with the decision-maker's perception and comprehension of the decision criteria and their ability to reflect and use the support tool(s) to arrive at satisfying decision solutions.

Given the abundance and complex nature of information relating to environmental and health related issues, the task of designing quality representations is even more arduous [Frysinger, 2005]. We hypothesize that EDSSs should enable users to: decipher quality information from the quantity, comprehend system representations, reflect upon obtained results, and formulate satisfying decision solutions [Hepting and Maciag, 2005]. We illustrate this hypothesis by describing a new framework for design and evaluating it with a previously developed EDSS designed by the United States Environmental Protection Agency (US-EPA). A usability evaluation was conducted comparing each EDSS framework and qualitative (and quantitative) results relating to user interaction were collected. For the purposes of this paper, we emphasise the qualitative results obtained from the evaluation as they may provide a deeper insight into what the user truly thinks about the system.

2. FRAMEWORKS FOR EVALUATION

Recently, the US-EPA developed an online EDSS for environmentally preferable purchasing of cleaning products using a database of 29 cleaning products distinguished between eight environmental and health related features – including:

- skin irritation, food chain exposure, air pollution potential (volatile organic compound percentage – VOC %), fragrance, dye, recyclable paper packaging, concentrated packaging, minimizing exposure to concentrated packaging

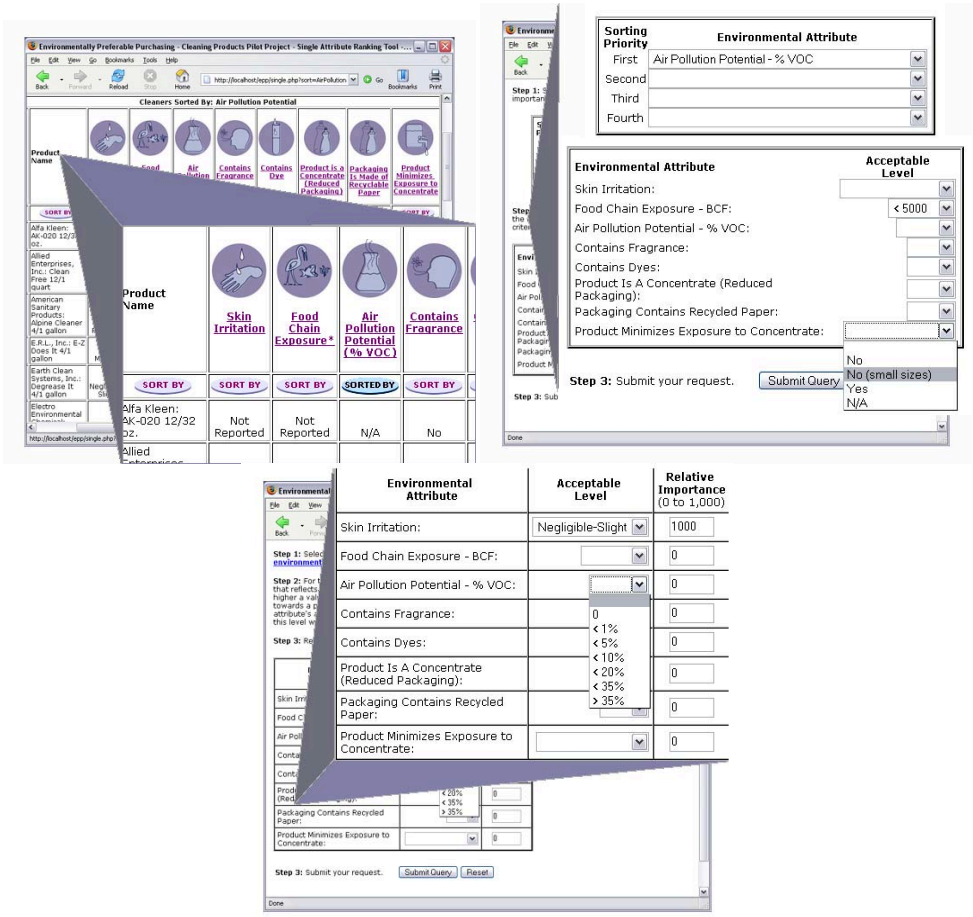


Figure 2. Single-feature ranking tool (upper-left), multiple-feature ranking tool (upper-right), and the weighted-feature ranking tool (bottom-centre). Each tool has a portion of its interface magnified for illustration purposes.

The US-EPA EDSS provided three different interface representations that enabled users to conduct reflective analyses. These were comprised of a single-feature ranking tool – enabling users to sort products by a single feature, a multiple-feature ranking tool – enabling users to sort products using up to four features with defined levels of priority, and a weighted-feature ranking tool – enabling users to sort products using up to all eight features with weighted importance values ranging from 0-unimportant, to 1000-most important. All three tools are illustrated in Figure 2. In all three instances, search results were presented to users in a tabular display, similar to that depicted in the illustration of the single-feature ranking tool in Figure 2.

In setting up our examination, we incorporated the product data and feature representations from the US-EPA EDSS into our own framework for design, which was based on a system originally developed by Hepting [2002] called *cogito*. *Cogito* differs from the US-EPA EDSS in terms of its primary interface representation as well as the way system objects are represented and displayed to the user (core system functionality also differs but that is beyond the scope of this paper). Instead of representing products in a tabular display, *cogito* uses a cell-type representation, comprised of up to eight cells per *cogito* page (up to how many pages are required). The system objects were represented in two separate ways, one being a textual (html-based) representation whereas the other, a graphical (nightingale rose-based) representation. Users were able to sort products by selecting which features and feature values (using in or all possible combinations) they were interested in reflecting upon. The *cogito*-based interface representations are presented in Figure 3.

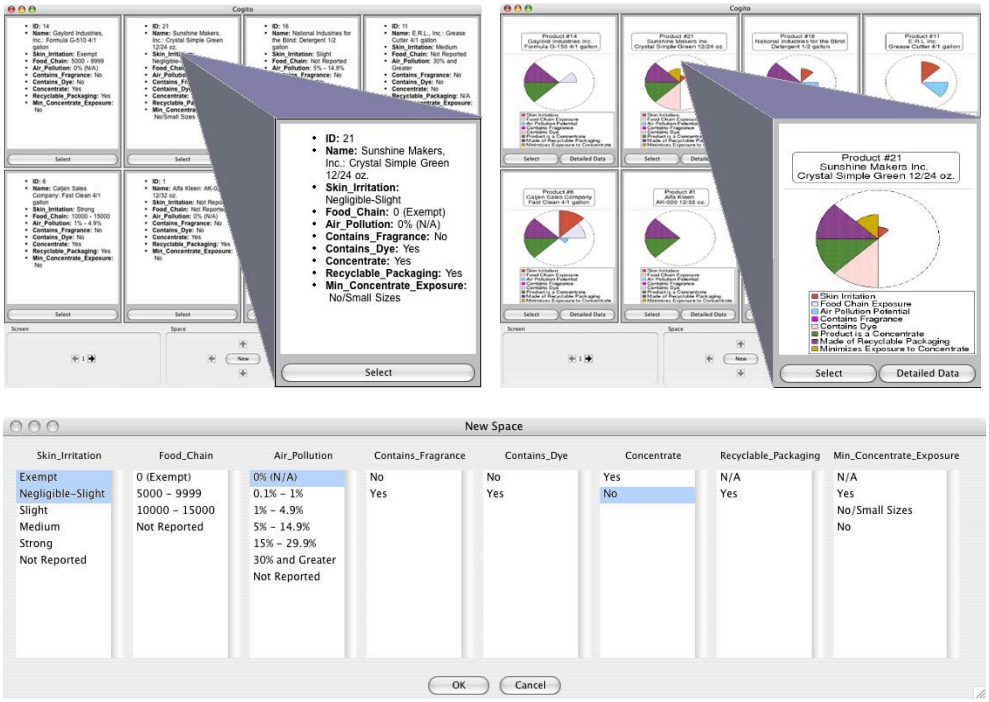


Figure 3. Cogito-based frameworks – Textual-based representation (upper-left), graphical-based representation (upper-right), and the search/query mechanism (bottom-centre). Both the textual and graphical representations have a cell magnified for illustration purposes.

3. EVALUATION

We used the US-EPA EDSS to conduct an analytical comparison with our *cogito*-based framework. We acknowledge that both EDSSs may enable some degree of reflective activity and that both provide quality representations of system objects. In this regard, the US-EPA EDSS provided three separate tools that enable reflection through a tabular representation. The *cogito*-based tools are similar to each other, but differ in terms of how the system objects are represented – textual versus graphical representations. Although it may be that all of these different representations provide reflection, for our examination, we were most interested in determining the degree at which each type of representation was successful in aiding the users. As such, it was subjective opinions of user satisfaction that we were most interested in acquiring and analysing.

For our evaluation, we recruited 28 participants from the University of Regina Computer Science Participant Pool [Hepting, 2006]. The participants were asked to perform a variety of reflective activities on the EDSSs and afterwards asked to relate their experiences by completing a questionnaire [Maciag, 2007]. In trying to obtain useful information from our participants, we asked questions relating to how they perceived the quality of their overall reflective experience – such as whether the tabular representation of the US-EPA EDSSs provided an adequate basis for reflection or whether the textual or graphical *cogito*-based representations provided a more preferred reflective environment.

Furthermore, we wanted to evaluate the participants' comprehension of certain criteria within the problem domain. We wanted to examine the correlation between participant responses and their intended perceptions. This was an attempt to rate the quality of the chosen features used to represent the cleaning products. Expert users may be able to satisfactorily define the given criteria with relative ease. However, we were interested if the same was possible for those who may be less experienced – for as previously mentioned, we hypothesize that this type of EDSS should provide support for both experts and non-experts alike. As such, we elicited the participants' interpretations of what we

considered were the more complex, or potentially problematic features representing the cleaning products. These included: *food chain exposure (fce), volatile organic compound (VOC)* - in relation to *air pollution potential, concentrated packaging (con), and minimizes exposure to concentrate (exp).* We also made note of any open-ended comments given by the participants. Results of this analysis proved quite interesting.

4. RESULTS AND DISCUSSION

When we asked participants to state their degree of agreement in terms of how the objects were represented by each EDSS and whether the representation enabled reflective activities, results indicated a slight preference for the graphical-based *cogito* EDSS. When asked whether the tabular representation provided by the US-EPA EDSS was conducive to reflective activities, 71% agreed (25% strongly agreed). When asked whether the textual representations provided by the *cogito*-based EDSS were conducive to reflective activities, only 64% of the participants agreed (14% strongly agreed). Finally, when asked whether the graphical representations provided by the *cogito*-based EDSS were conducive to reflective activities, 75% agreed (39% strongly agreed).

Looking back on the data collected we noted that many participants had a high exposure to tabular data on a monthly basis – 71%, with over half of the participants (57%) having weekly exposure. This could be why a higher percentage of participants felt that the tabular representations provided such an effective representation for reflective thought. In terms of the results obtained for the textual-based and graphical-based *cogito* EDSSs, one of the factors that may have contributed to the lower percentage of participants who thought the textual-*cogito* was less conducive to reflective activities may be that the textual representations provided a more realistic depiction of the cleaning products. For example, when shopping in a local market, a consumer can easily pick up a product and read its ingredient label. Here, there is only a slight difference between the represented world and the representing world – thus, the need for, and use of an external decision aid may be perceived as redundant. The tabular representations provided by the US-EPA tools may also evoke a similar response in such regard. However, given the participants' previous exposure to tabular data – results may have been skewed in its favour. Opinions relating to the graphical-based *cogito* EDSS provided insight into the power of refined illustrations. This was further indicated in open-ended comments given by the participants [Maciag, 2007] – that the graphical imagery was preferred, as upon first glance, you obtained an instant "feel" (stimuli) for the product. Many participants also commented on how the *cogito* EDSS provided a more conducive environment to conduct reflective activities given its cell-type user interface representation, as opposed to having to scan through the complex tabular display provided by the US-EPA.

When observing the results of our examination of user comprehension, results were interesting. One of the questions we asked the participants was whether they felt that the eight features representing the cleaning products (Section 2) were understandable and helpful in their reflective activities. 96% agreed (36% strongly agreed) that they were. However, when asked to define some of the *more complex* features, results contradicted the previous indicators as there was a rather large disconnect between how well the participants' defined the features and whether they thought they were understandable and helpful. The average participant score for each definition was below 50%, with only *fce*: 25%, *voc*: 29%, *con*: 14%, and *exp*: 36% of participants who correctly defined the respective definitions – thus, indicating the majority lacked a true understanding of the criteria. This provided a realization of the need to re-evaluate certain product feature representations. However, these results may indicate a larger issue, being that the participants almost unanimously stated their agreement that the features were understandable and helpful, yet were unable to successfully indicate their comprehension of them. *Could it be that the participant's simply did not consider these four features as being important in their reflective activities?*, or, *could it be that since these four representations were provided by the EDSS by default, that they were assumed to be important?* Many questions arise from these results. However, more research is required in this regard.

5. CONCLUSION

This paper examined the role of reflection and representation in EDSSs. Almost any EDSS will enable its users to conduct reflective activities. However, it is the degree to which reflective activities can be effectively, and satisfactorily conducted that denote the success of the EDSS. Through our examination, we illustrated that for an EDSS to be effective in such a manner the support framework upon which the system is built must incorporate high quality system representations – ranging from the user interface display to how system objects are modelled and represented. We noted that designing quality representations is an ongoing practice – as what has meaning to some users may not have meaning to others. In this regard, from our examination it was shown that some users might prefer more refined representations, while others may prefer more realistic ones. Deciding which representation to model the system by may be unique to the decision domain. We hypothesize that designers need to continually examine the needs of their users' in order to ensure they can satisfactorily conduct their reflective activities and obtain satisfying decision solutions. Many questions still exist and there are many opportunities for future analysis. Future work will include further analyses on the role of reflection and representation and deeper analysis in understanding how to best to design EDSS to ensure user satisfaction in such regard.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Centre for Sustainable Communities at the University of Regina and the Natural Sciences and Engineering Research Council (NSERC) of Canada

REFERENCES

Blandford, A., Computer Support for the Development of Decision-Making Skills, *British Journal of Educational Technology*, 22(1), 48-59, 1991.

Courbasson C.M., Reflective Activity and its Measurement, *International Journal of Mental Health Addiction*, 4, 133-149, 2006.

Fischer, G., Social Creativity: Making All Voices Heard, In Proc. HCI International Conference, 2005.

Frysinger, S.P., Human Factors of Environmental Decision Support Systems, In Proc. International Symposium on Environmental Software Systems (ISESS), 2003.

Frysinger, S.P., Cognitive Considerations in Environmental Decision Support, In Proc. International Symposium on Environmental Software Systems (ISESS), 2005.

Hepting, D.H., Towards a Visual Interface for Information Visualization, In Proc. Information Visualization, 295-302, 2002.

Hepting, D.H. and Maciag, T., Consumer Decision Support for Product Selection, In Proc. International Symposium on Environmental Software Systems (ISESS), 2005.

Hepting, D.H., Ethics ad Usability Testing in Computer Science Education, *ACM Special Interest Group on Computer Science Education*, 36(2), 76-80, 2006.

Hoyrup, S., Reflection as a Core Process in Organisational Learning, *Journal of Workplace Learning*, 16(8), 442-454, 2004.

Maciag, T., Supporting Sustainable Decision-Making: Evaluation of Previous Support Tools with New Designs, VDM Verlag Dr. Muller, 2007.

Norman, D., The Design of Everyday Things, Basic Books, 1988.

Norman, D., Things That Make Us Smart, Basic Books, 1993.

Norman, D., Emotional Design, Basic Books, 2004.

Rosson, M.B. and Carroll, J.M., Usability Engineering: Scenario-Based Development of Human-Computer Interaction. Academic Press, 2002

Swayne, D.A., Denzer, R., Lilburne, L., Purvis, M., Qwinn, N.W.T., and Storey, A., Environmental Decision Support Systems: Exactly What Are They? *Environmental Software Systems*, 259-268, 2000.

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Monolith2 - An On-line Database for Cement/Waste Products

Martin O'Shea [a], **Julia A. Stegemann** [b] **and Mark Levene** [a]

[a] *School of Computer Science and Information Systems, Birkbeck, University of London, Malet Street, London WC1E 7HX (martin@dcs.bbk.ac.uk, mark@dcs.bbk.ac.uk)*

[b] *Department of Civil, Environmental & Geomatic Engineering, University College London, Chadwick Building, Gower Street, London WC1E 6BT (j.stegemann@ucl.ac.uk)*

Abstract: Monolith 2 is an on-line database written as part of a project that aims to develop process envelopes for generic treatment of problematic industrial wastes by stabilisation/solidification with cementitious binders. The front-end of Monolith 2 is web-based and is made up of a core application written in HTML with Java servlets and Java server pages using NetBeans. This front-end of the system connects to a back-end MySQL database, which implements a complex data model described in the paper.

Keywords: stabilisation; solidification; cement; industrial waste; ProCeSS project; web application.

1 INTRODUCTION

This paper describes the Monolith2 on-line database system, which aims to collect a wide variety of highly technical information about waste treatment with cementitious binders from multiple laboratories.

Monolith2 is a web-based application, whose front-end consists of a core application written in HTML with Java servlets and Java server pages using NetBeans. This front-end connects to a back-end database implemented in the open-source database management software MySQL. As well as possessing the characteristics of a conventional database application, Monolith2 has specifically written facilities for searching and querying data and uses third-party tools to plot data and to provide analyses of server log activity.

2 THE PROCESS PROJECT

Stegemann and Côté [1996] defines stabilisation/solidification (S/S) with cementitious or pozzolanic binders as an option for reducing leachability of contaminants from residual, predominantly inorganic, industrial wastes before disposal or reuse. Treatment by S/S is complicated by the fact that the presence of impurities such as the components of industrial wastes can have deleterious effects on cements. Therefore careful laboratory development and testing of S/S formulations are required prior to full-scale application, to avoid technology failures.

The ProCeSS (Process Envelopes for Cement-based S/S) project, Stegemann [2006], aims to support good practice in S/S by developing 'process envelopes' [1] for generic S/S of common residual

[1] The process envelope for a process is described in Stegemann [2006] as 'the range of operating conditions that result in an acceptable product. To determine the process envelope for any process, it is necessary to establish and quantify the key practical, technical and economic characteristics of both process and product'.

waste types. Data, gathered from a series of laboratory research projects, is being used to find relationships between engineering and leaching properties and S/S product composition in order to define process envelopes that describe the limits of applicability of S/S technology to the chosen waste types. Furthermore, such data will prove vital in developing scientific understanding of the mechanisms underlying successful S/S, and is required to predict the longevity of these materials and evaluate the long-term risk associated with their disposal.

Part of the ProCeSS project is a statistically-based experimental programme conducted by laboratories at four universities: University College London, Imperial College London, University of Surrey and University of Cambridge. The results from the experimental programme are being collected in the on-line Monolith Database and Interface for Cement Based Products, i.e. Monolith2, which has been written for data entry/search/retrieval/exchange.

3 MONOLITH2

Monolith2 is an enhancement of a Microsoft Access-based database, Monolith2001, created in a previous project, Stegemann et al. [2001], which collected data for nearly 8000 cement/waste products composed of materials such as incinerator ash and other baghouse dusts, solidified with Portland cement, fly ash, blast furnace slag and other binders, from more than 300 literature references. The on-line Monolith2 system has been developed to consolidate this wealth of information from the literature, with that generated in the ProCeSS project, in a quantitative, accessible format, which will be available to other scientists, internationally, at the completion of the project.

4 STRUCTURE

4.1 Architecture

Architecturally, Monolith2 forms a typical web-based application, where the client browser makes requests to the web server which, in turn, retrieves relevant data from the database and returns it to the client within a web page. The server comprises a core application written in HTML with Java servlets and Java server pages using integrated development environment NetBeans [2007]. Apache Tomcat [2007] is used as a 'container' to implement the run-time environment for the servlets and JSPs according to Sun Microsystems' (`http://www.sun.com`) specifications.

All aspects of Monolith2's functionality are implemented using servlets and JSPs including the basic data entry and maintenance pages, the search and query facilities and the connections to third-party products for plotting and reporting (Section 5). Its database is implemented with the open-source DBMS MySQL [2007]. The architecture of Monolith2 is shown in Figure 1.

4.2 Data model

Information regarding the handling, strength, durability and leaching characteristics of products in Monolith2 (Section 3) are defined in database tables, which, in turn, are used to define the cement/waste product composition. Other tables contain the results of various physical and leaching tests performed on the S/S products.

Monolith2's data model, described in Stegemann [2001], consists of the following components:

- **A: References** stores information about the scientific publications that data has been taken from.
- **B: Mix Components** stores information about the materials that are mixed together to create a cement-based product, i.e. binders, industrial byproducts, water and additions.

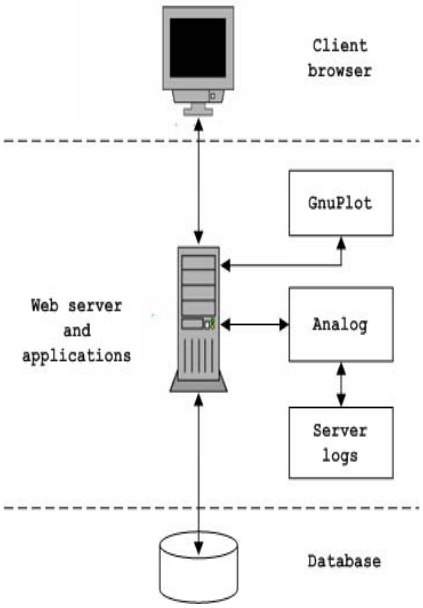


Figure 1: An abstract representation of Monolith2's architecture.

- **C: Phase and Oxide Composition** stores the phase or oxide composition of mix components.
- **D: Conventional Parameters** consists of four tables that store measurements such as pH and loss on ignition for components, products, leachants and leaching tests.
- **E: Chemical Composition** is made up of four tables storing the chemical analyses for mix components, products, leachants and leaching tests.
- **F: Product Formulation** specifies the quantity of each mix component in each cement-based product.
- **G: Product Mixing** stores information regarding the preparation of the cement-based products and their mixing, i.e. paste properties.
- **H: Physical Properties** consists of two tables storing physical properties of products (including hydration) and the physical properties of mix components.
- **I: Leaching Details** stores information regarding the leaching tests applied to the mix components and products.
- **J: Single Batch Leach** forms two tables storing data about leaching tests on mix components and products.
- **K: Other Information** stores any other information that was in the reference, but could not be accommodated elsewhere in the database, e.g. concerning matrix microstructure or transport characteristics.

The schema for the data model illustrating how the various components interact is shown in Figure 2. A number of metadata tables concerning users and privileges also exist.

5 FACILITIES

The following sections describe Monolith2's key features.

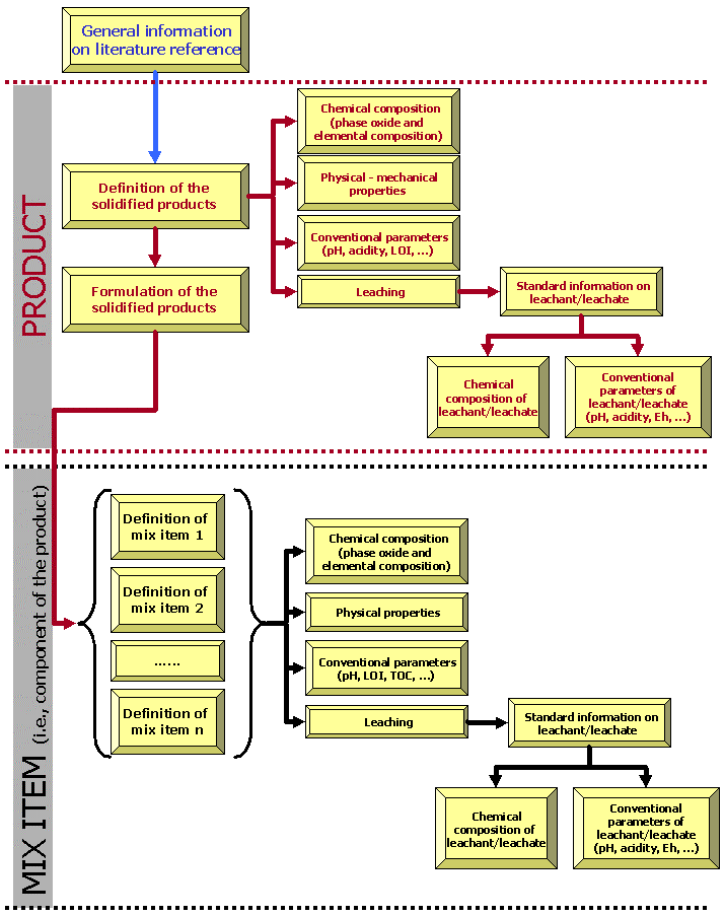


Figure 2: Schematic diagram of Monolith2's database structure, illustrated in Stegemann [2001].

5.1 Interface and Access

Upon logging in, the user is presented with the application's homepage. This forms a common point of entry and, as shown in Figure 3, the data model is displayed to allow simple navigation of the tables forming the database, each of which is independently maintained. From the horizontal banner strip at the top of the homepage, all users have access to the system components described below.

Currently, access to Monolith2 is restricted to a small user base whose user details are maintained by the database administrator. However, there are plans to change this so that any potential user will be able to register with the system on-line.

5.2 Data entry and update

Individual pages for each database table's insert, update and delete operations are implemented in Monolith2 using servlets and JSPs which are generated at run-time using embedded SQL and HTML. During data entry and maintenance, database keys are automatically generated and displayed in drop-down menus for ready selection; all units for value columns can also be chosen in this way. Additionally, Monolith2 has a built-in safeguard allowing the data entered by a particular user to be maintained only by that user or by those of that user's organisation, rather than by anybody else.

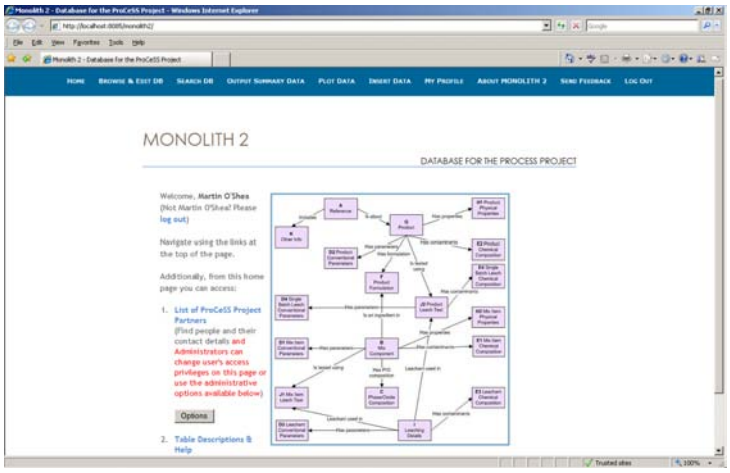


Figure 3: Navigation in Monolith2.

5.3 Searching

Two forms of searching are available. A 'table' search will display all the data in a table from a particular affiliation, and a 'value' search will retrieve data according to the value of a field in one of the tables.

5.4 Querying

Monolith2's principal method of database querying is via a pair of servlets, each of which serves the products or mix components types in the database. These types consist of several tables each (Section 4.2) and the dynamically generated HTML pages allow users to select from lists of these tables, their fields and field values for physical properties, e.g. unconfined compressive strength. Once selection is made, appropriate SQL is generated and temporary database views are created to retrieve the data and present it the user in the form of an HTML table.

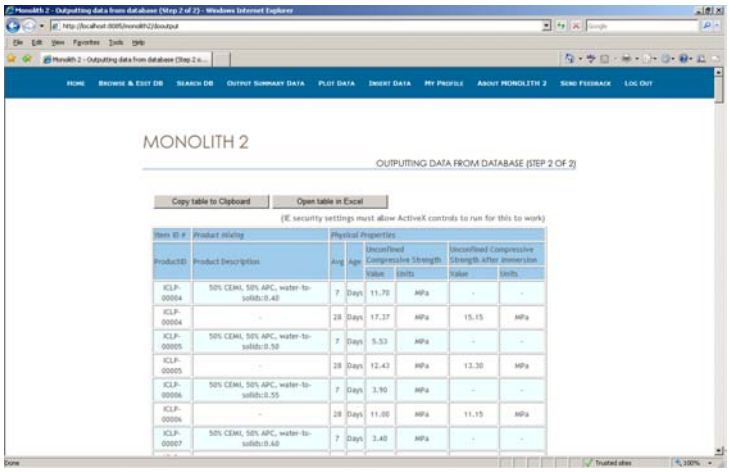


Figure 4: Querying unconfined compressive strength data in Monolith2.

Product data can also be displayed with age averages in rows or columns according to the criteria of physical properties entered. An example of query output is shown in Figure 4.

Additional data querying options allow for the generated output to be copied to the user's PC Clipboard or to be exported directly into Microsoft Excel.

5.5 Plotting

Monolith2 provides a tool for users to graphically represent types and ranges of the data entered. This facility consists of a set of web pages where users select items for the plot's X, Y and Z axes from drop down menus which correspond to tables, or individual columns of tables, from the database.

Users can also specify options for the plot including:

- The plot image's size and type, i.e. a fixed surface, a fixed surface with contours or a grid with contours.
- The plot's title and axes labels and font size to be used.
- Horizontal and vertical rotation.

The items and options chosen are then combined with system-generated SQL and temporary views to extract the relevant data from the database, 'plotted' as an image and returned to the user within a web page.

For each plot, all data and options are also permanently recorded in the database. Currently plotting is restricted to product data in Monolith2 but the facility has been written in a flexible manner to allow its future extension over mix item data.

The plots, generated by a third-party plotting tool, GnuPlot [2007], are designed to be easily comprehensible according to the principles of 'graphical excellence' and 'graphical integrity' defined in Tufte [2001]. An example plot is shown in Figure 5.

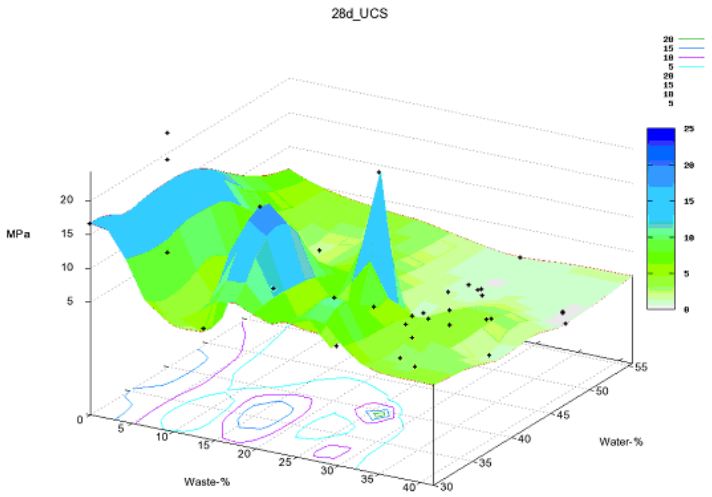


Figure 5: A plot generated by Monolith2 showing unconfined compressive strengths measured at 28 days for a variety of metal sludges treated with several different blended binder systems (the spike in strength occurs at 25% waste, 40% water and 75% binder, dry mass basis).

5.6 Reporting

A web server accepts HTTP requests from clients and serves those clients HTTP responses along with optional data contents, i.e. web pages, graphics et al. A web server also records data concerning user agents, i.e. browsers, date and times of visits, client IP addresses, quantity of data downloaded and so on. Data mining techniques are then employed to analyse this data to 'measure the "success" of the site with reference to its objectives', as described by Levene [2006].

In Monolith2, administrators can use the popular, open-source tool Analog [2007] to generate simple reports which include tables, bar- and pie-charts listing data concerning Monolith2's usage. Monolith2 also logs user requests for data from its database, updates, browsing and run-time exceptions in text files rather than the database. This data can also be used to model user activity within Monolith2 according to known techniques, e.g. Markov chains or forms of statistical analysis.

5.7 Other features

Monolith2 also includes a simple feedback facility to allow users to report comments or incidents directly to the database administrator by email. Extensive, on-line help text also describes each database table, their fields and their domain types, lengths and so on.

Finally, Monolith2 administrators have a servlet-based option to manually run a back up of Monolith2's MySQL database. This process emails the DBA so that the back up can be permanently archived.

6 APPLICATIONS

As stated in Stegemann [2001], 'Monolith2 was developed to catalogue the technical properties of cement-based products, not to make judgements about whether the properties or materials are good or bad, or where the line between "nonwastes" and "wastes" should be drawn. It is hoped that one of the uses of the database will be to make it easier for the appropriate authorities to make these distinctions in a sensible way, in order that more "waste" materials will be utilised, conserving natural resources and minimizing landfill while protecting the environment'. Stegemann [2006] further states that 'such data will be invaluable in developing a scientific understanding of the mechanisms involved in S/S, which is required to be able to predict the longevity of waste materials and evaluate the long-term risk associated with their management'.

7 FUTURE DIRECTIONS

The potential for future development of the Monolith2 database is defined in Stegemann [2001] 'by entry of information from users in both research and industry, as well as expansion of its structure to accommodate additional properties', e.g. related to contaminant transport and materials, and 'use of the data for development of additional models and knowledge-based systems'.

Interim plans for functionality include improving output overall and extending the plotting facilities. Monolith2's 'help' and feedback facilities are also likely to be enhanced with questionnaires. Longer term plans for Monolith2 include: improving the presentation of query and graphical output to match the specific requirements of the user base, providing data mining capabilities for modelling user activity within Monolith2, and personalisation to tailor pages to individual users' characteristics and preferences.

8 CONCLUSION

Stegemann [2001] states that the Monolith2 database 'can accommodate a wide variety of quan-

titative information about the characteristics of raw wastes, cement/waste product formulations, physical properties of cement/waste products, and single batch leaching test results for wastes and cement/wastes products'. The interface also 'provides flexible access to this information, allowing users to search for materials of interest, to aid in formulation design, evaluation, or process feasibility assessment'.

ACKNOWLEDGEMENTS

ProCeSS is co-funded by the Technology Strategy Board's Collaborative Research and Development programme, following an open competition. The Technology Strategy Board is a business-led executive non-departmental public body, established by the government. Its mission is to promote and support research into, and development and exploitation of, technology and innovation for the benefit of UK business, in order to increase economic growth and improve the quality of life. It is sponsored by the Department for Innovation, Universities and Skills (DIUS). Please visit *http://www.innovateuk.org* for further information.

ProCeSS is led by University College London, with the participation of 21 partners from academia (Imperial College London, University of Surrey, University of Cambridge, Birkbeck University of London), and industry (including British Cement Association, The Concrete Centre, UK Quality Ash Association, Cementitious Slag Makers Association, British Lime Association, Elkem Materials, Surface Engineering Association, SELCHP, Corus, Veolia Environmental Services, Grundon Waste Management, Sita UK, Scott Wilson, May Gurney Ltd, White Young Green Environmental, CIRIA and WRc). The project website is at *http://www2.cege.ucl.ac.uk/process*.

Monolith2 was originally adapted from Monolith2001 by Kevin Keenoy at the School of Computer Science and Information Systems, Birkbeck, University of London.

REFERENCES

Analog. *http://www.analog.cx/*. 2007.

GnuPlot. *http://www.gnuplot.info*. 2007.

Levene, M. *An Introduction to Search Engines and Web Navigation*. Addison Wesley, 365 pp., Harlow, Essex, UK, 2006.

MySQL. *http://www.mysql.com*. 2007.

NetBeans. *http://www.netbeans.org*. 2007.

Stegemann, J. A. *MONOLITH - A Database of Cement/Waste Properties*. ICMAT 2001, Symposium I, Singapore, July 1 - 6, 2001.

Stegemann, J. A. *Development of Process Envelopes for Cement-based Stabilisation/Solidification*. 6th International Conference on the Environmental and Technical Implications of Construction with Alternative Materials (WASCON), held in Belgrade, Serbia and Montenegro, 30 May - 2 June, M. Ilic, J. Goumans, S. Miletic, J. Heynen and J. Senden, Eds., ISCOWA/TTPI, pp. 783-790, 2006.

Stegemann, J. A., E. Butcher, A. Irabien, P. Johnston, R. de Miguel, S. Ouki, A. Polettini, and G. Sassaroli. *MONOLITH - A Database and Interface for Cement-based Products*. Developed under Contract No. BRPR-CT97-0570, Commission of the European Community, Brussels, Belgium, 2001.

Stegemann, J. A. and P. Côté. A proposed protocol for evaluation of solidified wastes. *Science of the Total Environment*, 178 (No.1-3):103–110, 1996.

Tomcat. *http://tomcat.apache.org*. 2007.

Tufte, E. *The Visual Display of Quantitative Information*. Graphics Press, 197 pp., Conn, USA, 2001.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Coupling empirical and deterministic models to assess surface water contamination

M.Terrado[a], M.P. Lavigne[b], D. Barceló[a], S. Duchesne[b], J.P. Villeneuve[b], A.N. Rousseau[b], S. Tremblay[b] and R.Tauler[a]

[a] Department of Environmental Chemistry, IIQAB-CSIC. Jordi Girona 18-26, 08034 Barcelona, Catalonia, Spain (mtcqam/dbcqam/rtaqam@iiqab.csic.es)
[b] Institut National de la Recherche Scientifique, INRS-ETE. 490, rue de la Couronne, Québec, G1K 9A9, Canada (martin-pierre_lavigne/sophie.duchesne/jean-pierre_villeneuve/alain_rousseau/sebastien.tremblay@ete.inrs.ca)

Abstract: Integration of chemometrics and deterministic modelling is proposed as a source for extracting useful information from data sets obtained in environmental monitoring studies. Organic contamination in the Llobregat river basin (Catalonia, NE Spain) was investigated for the period 2003-06. From the application of Multivariate Curve Resolution using Alternating Least Squares (MCR-ALS), 5 organic contamination sources were identified, explaining 87% of the total data variance. Pollution from chlorinated pesticides and alkylphenols showed a decrease in recent years, while contamination caused by PAHs had a rather constant distribution in space as well as in time. On the other hand, diffuse pollution from a contamination source described by contribution of terbutryn (triazine) and chlorpyrifos (organophosphate insecticide) presented an increase in 2006 with regard to 2005. Dynamics of this increasing identified pollution source was modelled using the system GIBSI. Knowledge about existing cultures, usages of the pesticides defining the contamination source and their application dates and quantities was necessary to be used as input for the model calibration. Therefore, the impact of existing agricultural practices over the basin can be simulated and compared with possible scenarios involving different practices. After watershed modelling, the effects of a proposed measure can be evaluated.

Keywords: chemometrics; deterministic modelling; hydrological modelling; contamination sources; pesticides

1. INTRODUCTION

The big challenge for the European water and environmental managers nowadays is the implementation of the Water Framework Directive (WFD). It is for this reason that new policy instruments have been introduced to protect and improve European waters. An ecological and holistic water status assessment approach, a river basin water planning and a strategy for elimination of pollution by dangerous substances are included among them. To achieve acceptance for environmental policies and successful implementation, trustworthy methods for the identification of contamination sources affecting the quality of surface water in a specific river basin are needed. In the same way, monitoring of the identified sources is also important in order to assess and control contamination distribution, since it allows establishing those parameters which mostly affect pollutants dispersion into the environment as well as calculating the expected effect of a proposed measure.

Chemometric data analysis methods [Massart et al., 1998] assume chemical concentrations measured in a particular sample as affected by multiple contributions coming from independent sources of different origin (industry, agriculture, urbanisation, nature, etc.). These sources are defined describing their chemical composition as well as their geographical and temporal distributions. Multivariate Curve Resolution using Alternating Least Squares (MCR-ALS) [Tauler, 1995] is a multivariate statistical method which allows

the transformation of complex data sets into a new perspective in which the more relevant information is made more obvious, decomposing the environmental data set in components describing the main sources of data variance. While chemometric methods are based on observation and analysis of experimental data, deterministic models are defined as numerical methods based on general laws or theoretical principles [Arheimer and Olsson, 2003]. Traditionally, scientific communities using both types of models have been clearly differentiated. Coupling of these two approaches is proposed so as to merge the information obtained from both methodologies, obtaining a better understanding of contamination problems affecting a river basin.

In this work, integration of the proposed methodologies was tested over a pilot study catchment: the Llobregat river basin (Catalonia, NE Spain). Draining an area of around 5000 km^2, the river is approximately 160 km long. It is born in Castellar de n'Hug (Berguedà) and goes by until reaching the Mediterranean Sea, in el Prat de Llobregat (Baix Llobregat) (see Figure 1). La Baells dam is located not very far from its source and regulates the river flow. Llobregat river has two main tributaries, Cardener and Anoia. The headwaters of Llobregat and Cardener are in a rather unpolluted area. However, the two river midwaters flow through a densely populated and industrialized region where salt and potash mines (in Cardona and Súria, respectively) exist. Agriculture is significant in this zone, mainly devoted to cereal farming, such as wheat, barley and fodder. Anoia headwaters are in an agricultural area. Downwaters flow through an industrialized area, and near the confluence with the Llobregat river the main activity is agriculture again (mainly vineyards), while industrial and urban contribution attains a lower importance. At its lower course, the river flows through one of the most populated areas in the Mediterranean region, waters receiving high inputs from industry and urbanisation. An intensive agricultural activity exists near the delta, with predominance of orchard and garden vegetables, such as lettuce, cabbage, green onions and artichokes.



Figure 1. Location map of the Llobregat river basin with identification of quality stations: LL005, Llobregat river at Guardiola de Berguedà; LL015, La Baells reservoir at Berga; LL020, Llobregat river at Olvan; LL030, Merlès stream at Santa Maria de Merlès; LL035, Llobregat river at Balsareny; LL040, Gavarresa stream at Sallent; LL045 Calders river at Navarcles; LL050, Or river at Sant Fruitós del Bages; LL055, Llobregat river at El Pont de Vilomara; LL065, Cardener river at Olius; LL075, Negre river at Clariana de Cardener; LL080, Cardener river at Cardona; LL085, Cardener river at Castellgalí; LL090, Llobregat river at Castellbell i el Vilar; LL095, Llobregat river at Olesa de Montserrat; LL100, Llobregat river at Abrera; LL105, Anoia river at Jorba; LL110, Anoia river at Vilanova del Camí; LL120, Riudebitlles stream at Sant Sadurní d'Anoia; LL125, Avernó river at Subirats; LL130, Anoia river at Martorell; LL135, Rubí stream at Castellbisbal; LL140, Llobregat river at Sant Joan Despí; LL145, Llobregat river at Barcelona.

2. METHODS

This section is divided in two parts. Part 2.1 describes empirical models applied in this work. Data pretreatment and analysis were performed under the Matlab 7.0 (2004; The Mathworks, MA, USA) computer and visualisation environment, using PLS Toolbox 3.5

functions (Eigenvector research Ltd., Manson, WA, USA). Part 2.2 describes deterministic models applied to the case of study. The physiographic database of the distributed hydrological model was prepared using the software PHYSITEL 3.0 (2006; INRS-ETE, Québec). The software HYDROTEL 2.6 (2002; INRS-ETE, Québec) allowed the hydrological calibration of the whole watershed. Eventually, GIBSI (*Gestion Intégrée des Bassins versants à l'aide d'un Système Informatisé*) (2007; INRS-ETE, Québec) was the integrated modelling system used for watershed management.

2.1 Empirical models

Structure of data and data pretreatment

From its extensive annual environmental monitoring program, the Catalan Water Agency (*Agència Catalana de l'Aigua*, ACA) is periodically obtaining large data sets with different parameters related to water quality. For the present study, different data matrices (one per year) were arranged with the concentration of 17 organic compounds in 24 sampling sites distributed all over the Llobregat river system for the period 2003-06 (see Figure 1 for sampling sites location). The four obtained data matrices had identical dimensions, variables and sampling sites coinciding every year. Analysed organic compounds (concentrations in ng/L) were triazines (atrazine, terbutryn and simazine), organophosphate plaguicides (chlorpyrifos and diazinon), chlorinated plaguicides (lindane, endosulphan I, endosulphan II, endosulphan sulphate and ΣHCHs), polyaromatic hydrocarbons (pyrene, fluoranthene and fenanthrene) and alkylphenols (nonylphenol etoxylate (e=0), nonylphenol etoxylate (e=1), nonylphenol etoxylate (e=2) and t-octylphenol). Compounds were selected on account of periodical data availability. Some of them, such as atrazine, chlorpyrifos, endosulfan, fluoranthene, HCHs, nonylphenol, octylphenol and simazine are substances proposed for regulation in the proposal of the prioritary substances list from the Water Framework Directive. Variables having too many values below the detection limit were excluded from the data set. For this reason, fenitrothion and fluorene, initially considered, were finally removed to perform the analysis. The four data matrices were concatenated one on the top of the other, keeping columns in common (variables) and giving an augmented data matrix **Daug**, of dimensions 96 rows x 17 columns. Because variable magnitudes differed considerably, scaling of **Daug** was considered to be necessary. Analysis without scaling was also performed to allow comparison, although results will not be displayed here for brevity. Scaling of individual matrices is neither presented in this work since it has been shown to result in a lost of information about the temporal component when identifying contamination sources distribution [Terrado et al., 2007].

Chemometric analysis

The multivariate data analysis method MCR-ALS was applied to data. It has de advantage of decomposing the data matrix by more natural constraints than those used in other frequently applied methods like Principal Component Analysis (PCA) [Joliffe, 1986], and environmental results can be interpreted straightforwardly. The drawback of MCR-ALS method is that solutions may not be unique and caution should be taken to check the reliability of the obtained results [Tauler, 2001]. MCR-ALS decomposes the augmented data matrix $\mathbf{D}_{\mathbf{aug}}$ by a constrained ALS algorithm using two alternating least-squares steps. Matrix decomposition gives a matrix of loadings ($\mathbf{V}^{\mathbf{T}}$), with the chemical composition of the main patterns of data variance, a matrix of scores ($\mathbf{U}_{\mathbf{aug}}$), describing the geographical and temporal distributions of the identified patterns and finally, a matrix of error ($\mathbf{E}_{\mathbf{aug}}$) containing the percentage of data variance not explained by the model. Information about geographical and temporal distribution of contamination sources appears intermixed in the columns of the resolved augmented scores matrix, being therefore, not directly available. However, MCR-ALS solutions can be additionally constrained to follow a trilinear model or to recover the information in the two modes, geographical and temporal [Tauler et al., 1998]. When trilinearity is applied, the profiles in the three different modes - $\mathbf{U}_{\mathbf{geo}}$ (scores matrix containing geographical information), $\mathbf{U}_{\mathbf{temp}}$ (scores matrix containing temporal information) and $\mathbf{V}^{\mathbf{T}}$ - are directly recovered and can be compared to obtained profiles

when using other methods based on a trilinear model, such as PARAFAC [Bro, 1997]. Constraints used for all data matrices in this study were non-negativity for the two modes (augmented scores mode and loadings in the second mode) and normalisation of the loadings matrix to equal length. Singular value decomposition was applied over data matrices when row-wise as well as column-wise augmentation was performed before applying any pretreatment on data. In row-wise augmentation, matrices were concatenated having rows in common (samples) while in column-wise augmentation, matrices had columns (variables) in common. This gave an estimation about how trilinear data matrices were. Since obtained results were quite similar, the application of the trilinearity constraint was considered to be suitable to resolve the main lines of data variation in a more simple way [Tauler et al., 1995, Felipe-Sotelo et al., 2006 and Peré-Trepat et al., 2007].

2.2 Deterministic models

Determination of the drainage structure of the basin

The complete drainage structure of the basin was obtained using PHYSITEL [Turcotte et al., 2001], a software program designed specifically to prepare the watershed data base for HYDROTEL (the hydrological model used as a second step for the watershed modelling). The most important data file for the determination of the drainage structure was a digital elevation model (DEM) based on a square grid with a specified spatial resolution (in this study, 200 m). Using the DEM, the drainage path from each cell to one of the cells surrounding it was defined. It was necessary to identify the cell corresponding to the outlet of the basin in order to define the complete area delimitation. After this, definition of the river network was possible, considering cells draining an area greater than a specified threshold. Finally, the determination of relatively homogeneous hydrological units (RHHU) was achieved, consisting on the creation of sub-watersheds associated to each river section of the modelled hydrological network. The software also allowed calculation of the percentage of land uses and soil types existing over each RHHU. Thus, a land use layer in raster format was obtained from the Catalan Government for year 2002. As no suitable soil map was available for the Llobregat basin, it was digitized for the present work taking an existing map of Catalonia and the Balear Islands form the Institute of Catalan Studies (IEC) as a reference.

Simulation of hydrological processes over the basin

For the hydrological simulation, the physically-based, distributed, hydrological model HYDROTEL was used [Fortin et al., 1995]. It consists of 6 computational models which are run in a cascade at each time step: 1) estimation of meteorological variables, 2) snow accumulation and melt, 3) potential evapotranspiration, 4) vertical water budget, 5) flow on RHHUs and 6) channel routing. For every model, the software offers the choice of different algorithms. Hydrological, meteorological and physiographical data were needed. Daily flow measures at different stations over the river network were furnished by the Water Catalan Agency for the different hydrological years of the studied period. Measures of flow at the outlet of the basin (or close to it) were essential. Meteorological data were obtained from the Catalan Meteorological Service (*Servei Meteorològic de Catalunya*, SMC) also at a daily frequency. Parameters used in the model calibration were established adopting the works from Lavigne, 2006 and Turcotte et al., 2007 as reference, and taking into account the regional differences existing between them and the region considered in this work. Parameters were adapted as much as possible to the Llobregat river basin [Bathurst et al., 2006]. Although the model was able to simulate the effect of dams, they were not considered in this work for reasons of data requirements.
Model calibration for the Llobregat river basin was performed in two different stages. In the first stage, a calibration of the whole basin was done focussing in the final part of the river. Abrera hydrological station (see Figure 2.a) was selected to start the calibration for i) being located sufficiently away from dam effects (as said, these effects were not modelled in the present study); ii) having a complete meteorological and hydrological data register for the studied period of time; iii) being placed in an area affected by the different contamination

sources existing in the basin (industrial, agricultural and urban); and iv) being the closer station to the outlet over whom no flow diminution was observed. Martorell and Sant Joan Despí stations were found Abrera station downstream. Despite being closer to the outlet, they were not chosen for calibration due to an observed flow decrease probably caused by water collecting at these sites, which made hydrological simulation more difficult. Once the first stage was finished and calibration values were obtained for Abrera station, a second simulation was performed dividing the basin in 6 different sub-basins (groups of RHHUs) (see Figure 2.a) and using results from the previous calibration as initial calibration parameters. In this way, the basin is not considered as a whole, and values of the different parameters can be adjusted independently for each group of RHHUs. This offers a higher flexibility to the model which sometimes can be necessary in order to obtain a more accurate calibration. Data about 14 hydrological stations were used, each station being associated to a river section (see Figure 2.b).

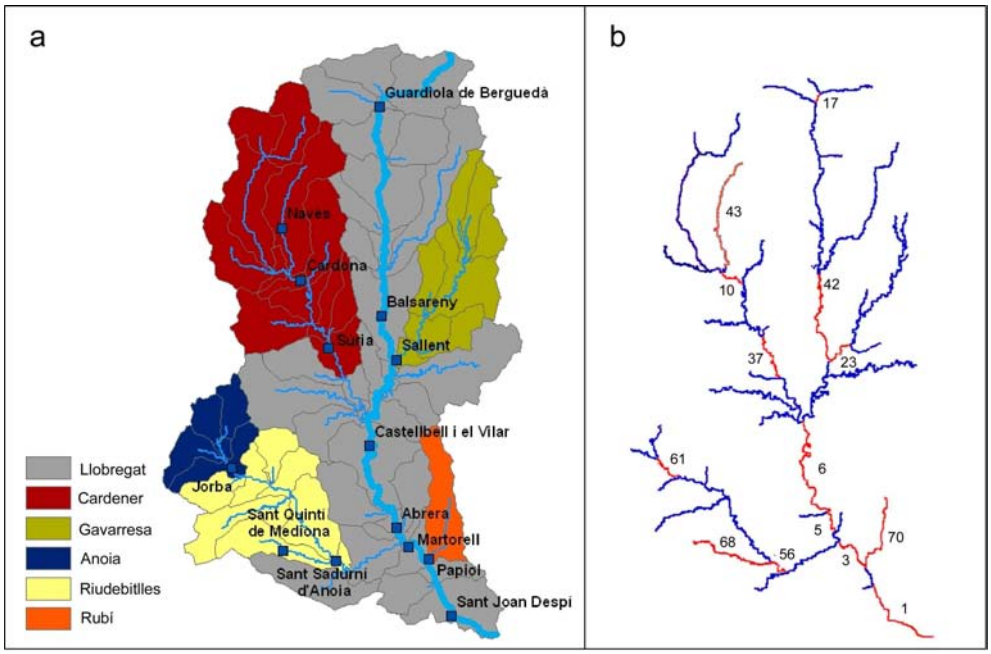


Figure 2. HYDROTEL calibration steps: a) Division of the Llobregat basin in UHHRs and sub-basins (groups of UHHRs). Spatial distribution of hydrological stations; b) Association of the different hydrological stations to a specific river section.

HYDROTEL calibration for the Llobregat basin was based in the strategy proposed by Turcotte et al., 2003. Parameters being more often modified during calibration were the recession coefficient (RC), the 2nd and 3rd soil layers depth (Z2 and Z3 respectively) and the potential evapotranspiration multiplication factor (FETP). RC, Z2 and Z3 are calibration parameters controlling the vertical water budget module of stream flow forecasting over the basin while FETP controls potential evapotranspiration. RC regulates contribution of the 3rd soil layer to flow. It is associated with prolonged summer drought recession objectives. Z3 depth affects the water quantity of this layer, horizontal flow and vertical flow between this layer and Z2. FETP is related to variables such as soil humidity content and meteorological conditions. As its relative sensitivity is high for all the periods of the year, flow is rapidly affected by FETP modification. That makes FETP and Z3 suitable to regulate annual and summer flow volumes. Z2 has effects over flow simulation during the summer and fall periods (usually, the rainy season) [Lavigne, 2006].

Integrated modelling of the basin

The integrated modelling system for watershed management GIBSI (*Gestion Intégrée des Bassins versants à l'aide d'un Système Informatisé*) was used for the Llobregat river basin modelling. The system can either be used as a data management system or as an impact assessment tool to study the effect of management scenarios on surface water quality using

mathematical models [Villeneuve et al., 1998 and Quilbé et al., 2007]. It is comprised of a relational database management system (RDMS), physically-based simulation models (hydrology, soil erosion, chemical pollutants transport and water quality), management models (point source, diffuse source and reservoir management models) and a geographic information system (GIS). Evaluation is performed via scenarios comparison. Two types of data, spatial and alphanumerical, are needed to be integrated in GIBSI. Spatial data integration was done by means of the GIS GRASSLAND. In this study, raster and vector data layers with a 200 m resolution and UTM projection European Datum 1950 Zone 31N were used. Required layers were a DEM of the whole area, the basin contours, municipalities, land uses, pedology, slope, sub-basins, SSUs (which corresponded to UHHRs in HYDROTEL) and river sections. Digital cartography was obtained from the Catalan Government (*Generalitat de Catalunya*) web site. Alphanumerical information needed as input depended on the simulation models to run. The aim of the present work was the implementation of the pesticides transport model, for which turning of the hydrological and soil erosion models was required, since outputs of one model were used as inputs for the next one. The water quality model was not used in this work owing to a lack of data availability. The hydrological model integrated in GIBSI was HYDROTEL, which has already been described in the previous section. Using data calculated by the hydrological model for each specific SSU, the erosion model derived from the algorithms of the Universal Soil Loss Equation (USLE), calculates the sediment yield transported to the river system corresponding to every land use over a SSU. Values of the different parameters used for calibration were extracted from the literature [Bathurst et al., 2006, Rawls and Brakensiek, 1989, Wischmeier, 1978].

The chemical transport model is based on the nitrogen, phosphorus and pesticide transport algorithms of the Soil and Water Assessment Tool (SWAT) [Arnold et al., 1998]. Transport model is used to calculate the agricultural pollutant loads exported to the river flow. Nitrogen and phosphorus algorithms (PO module) were not used in this study since attention was only focussed on pesticides transport (PES module). Pesticides algorithm simulates application, degradation and transport of pesticides by water and eroded soil particles over each agricultural land use in every SSU of the basin. Information about products cultivated in the region of study was obtained from a Vegetal Defence Association (*Associació de Defensa Vegetal del Baix Llobregat).* As information about agricultural practices was not directly available, it was extracted from the fitosanitary warnings published by the Catalan Department of Agriculture, Alimentation and Rural Action (DARP). Pesticides application taxes were obtained from the *vademecum* of fitosanitary products, document collecting recommendations of suitable doses to use at each pesticide application.

3. RESULTS

3.1. Chemometric results

MCR-ALS application on the organic contamination data matrix identified 5 principal contamination sources existing in the Llobregat river basin for the period 2003-06. Together, these 5 identified sources explained 87.08% of the total variance of data (see Figure 3). The major part of data variance (66.40%) was explained by a contamination source with industrial origin describing PAHs contributions. Distribution of this contamination source was rather constant in space as well as in time (see Figure 5). Spatially, the outstanding sampling site in comparison to the rest of locations was LL085, emplaced in the Cardener river at Castellgalí (see Figure 1). It is located at the central part of the basin, downstream of Manresa, which is one of the most populated municipalities in the basin, having an industrial activity especially important in the sectors of metallurgy, chemistry and leather. Manresa has also the condition of being a radial centre for road links, a key factor for its important commercial activity. The second identified contamination source explained 15.77% of data variance (a significant difference with respect to the first described source). It was loaded by lindane and total HCHs (a sum of different compounds, in which lindane isomer has the major proportion) as well as some contribution of PAHs, a group of compounds which appeared already in the previous identified source of

contamination (see Figure 3). This pattern corresponded to a mixture of pollution with agricultural (for lindane presence) as well as industrial origin (due to PAHs presence). It was a contamination source showing a clear decrease in time, for which its contribution to samples in year 2006 was half the contribution presented in 2003 (see Figure 4). However, for all the studied years, the more affected sampling sites were LL135, located in Rubí stream at the down part of the basin, followed by LL100, in the Llobregat river at Abrera station (see Figure 1). In the region near the confluence of the Llobregat river with the Anoia river (the second important tributary by the left side), there exists an important industrial activity located in Martorell. A third contamination source was identified explaining 8.19% of the total data variance. It was a contamination originated from agricultural practices and described by contribution of the group of analysed endosulphans (endosulphan I, endosulphan II and endosulphan sulphate) (see Figure 3). The highest levels of this contamination in samples were detected in year 2004. It presented a reduction during recent years, attaining its lower contribution in 2006 (see Figure 4). For all the studied period, LL135 followed by LL105, were the locations showing the highest levels of this contamination pattern. LL105 was placed at the upper part of the Anoia river, in an agricultural area mainly devoted to cereal farming. After reaching this point, the Anoia river flows through a zone of vineyards before intersecting the Llobregat river. As site LL135 is found downstream the intersection (see Figure 1), pollutants transported by the Anoia river when passing through this vast agricultural zone, are transported reaching the Llobregat river and downstream. Fourth and fifth identified contamination sources explained a very similar amount of data variance (see Figure 3). A clearly agricultural source defined by terbutryn and chlorpyrifos contribution was identified, explaining 5.30% of data variance. This source showed a decrease in its contribution to samples from 2003 to 2005, but an important increase was observed in 2006 (see Figure 4). LL135 and LL145 sampling sites (close to the river mouth) were identified as the most affected locations by this source. Cultures having a major predominance in the final part of the basin are orchard and fruit trees. Fifth pollution source explained 5.04% of the variance of data and it was loaded by the group of alkylphenols. These compounds can act as endocrine disruptors and are indicators of a type of contamination coming from industry and urbanisation. A clear decrease of this contamination source was identified in recent years for the analysed sites in the Llobregat river basin (see Figure 4).

From the identified contamination sources obtained by chemometric analysis of data matrices, two were found to have an agricultural origin: one describing a contamination by endosulphans and the other due to terbutryn and chlorpyrifos application. This last contamination source was selected for subsequent modelling since terbutryn and chlorpyrifos were found at concentrations above the detection limit in a 20% of the analysed samples, while endosulphan concentrations were below the detection limit in the major part of the analysed samples. Besides, the contamination source defined by terbutryn and chlorpyrifos was the only one showing an increase in the recent years (from 2004 to 2006), so a major attention should be put on it.

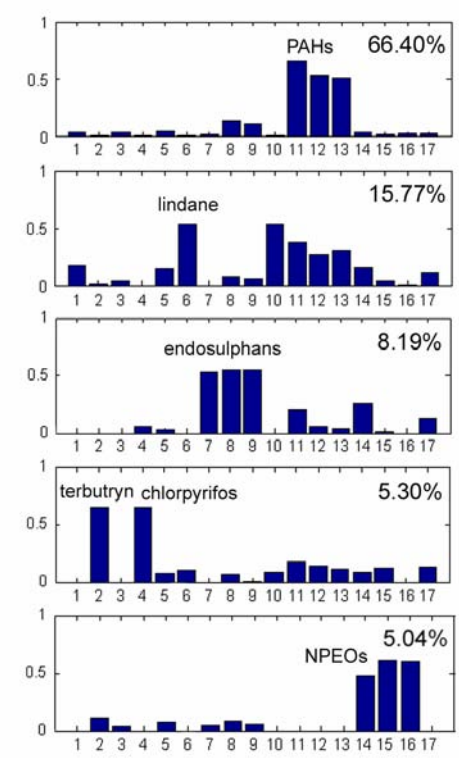


Figure 3. Contamination sources of organic compounds identified in the Llobregat river basin. Analysed compounds: 1, atrazine; 2, terbutryn; 3, simazine; 4, chlorpyrifos; 5, diazinon; 6, lindane; 7, endosulphan I; 8, endosulphan II; 9, endosulphan sulphate; 10, sum of hexachlorocyclohexane; 11, pyrene; 12, fluoranthene; 13, fenanthrene; 14, nonylphenol etoxylate (e=0); 15, nonylphenol etoxylate (e=1); 16, nonylphenol etoxylate (e=2); 17, t-octylphenol

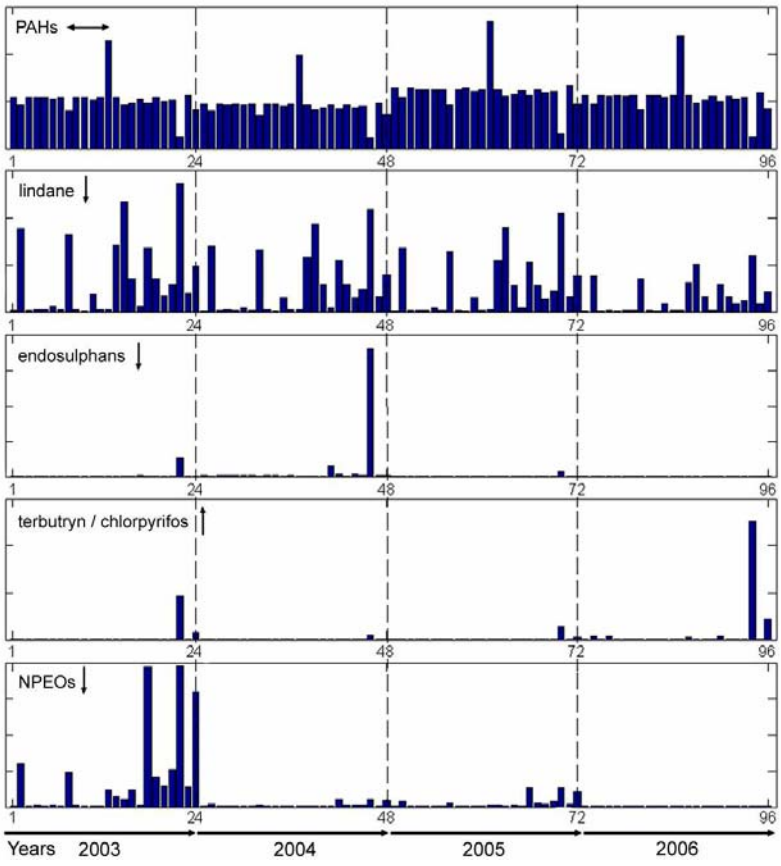


Figure 4. Spatial and temporal distribution of organic contamination sources identified in the Llobregat river basin. The 24 sampling sites are ordered in the X axis from north to south for every sampled year (2003 to 2006).

3.2. Deterministic results

As explained in the previous section, Abrera station was focussed for calibration of the hydrology model HYDROTEL. Total observed flow in Abrera for all the analysed period, $1.52 \cdot 10^9$ m^3, was very close to the simulated flow by the software HYDROTEL, $1.55 \cdot 10^9$ m^3. Obtained values of Nash-Sutcliffe coefficient and Root Mean Square Error (RMSE) were 0.61 and 0.125, respectively. These coefficients gave and idea of how good the simulation was. Nash-Sutcliffe coefficient can vary between one and minus infinite. In general, the closer it is to one, the better the simulation, since simulated flow is more similar to measured flow. For RMSE, simulation is better when its value tends to zero.
Figure 5 displays the results from the hydrologic simulation for the station located in Abrera. The area graphic in the first plot (right axis) represented daily temperatures. In this case, temperatures followed a similar pattern (shape) during the 4 years of study. The 0ºC temperature was indicated in the graphic by a horizontal dark line. On the left, daily precipitation (mm) as well as snow depth on soil were represented by a graphic of bars. Snow was only calculated for those days in which temperature fell below zero and a rain episode took place. However, in the case of the Llobregat basin, few days with this characteristics existed, being therefore, a parameter having a more limited importance. Temperature and precipitation corresponded to average values calculated over the set of RHHUs being placed upstream the simulated station, once interpolation between stations was performed. The central plot in Figure 5 showed the quantity of water available for runoff in the area over the station where flow was simulated. Finally, the plot at the bottom of Figure 5 was a hydrogram representing observed and simulated flows (using a dark area and a light line, respectively) in the river section corresponding to Abrera station. In this work, high flows originated by rain episodes occurring in spring and fall were clearly observed. There was a good synchronisation of high flows and a correct base flow simulation. On the other side, in the case of consecutive rain episodes, simulation of the retarded flow (water abandoning the soil slowly after rain) was not accurate enough. Retarded flow is related to the recession coefficient, a parameter which was difficult to calibrate for the case of study (sometimes simulated flow was higher than observed flow and other times it was lower). Thus, finally, a balance had to be achieved.

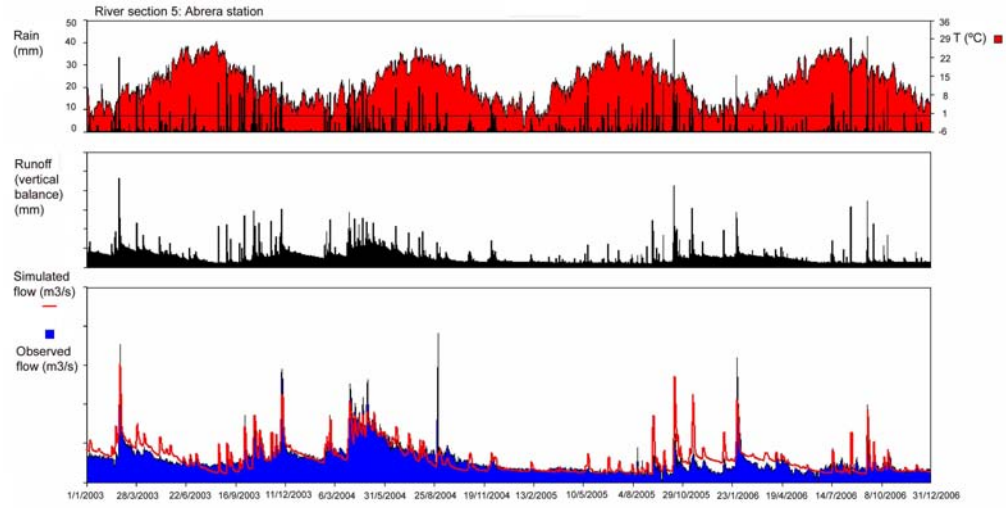


Figure 5. Results of HYDROTEL simulation for Abrera station (period 2003-06)

Application of the transport model was done for the case of a diffuse agricultural contamination source identified in the Llobregat river basin. As the source was loaded by a similar contribution of terbutryn and chlorpyrifos (see Figure 3), it was introduced to the data base of the GIBSI system as a compound X, integrating the properties of the different compounds defining the source. Thus, compound X had a sorption and a half-life time being the average of the two compounds.

ACKNOWLEDGEMENTS

This work is supported by the European Union FP6 Integrated Project MODELKEY (SSPI-CT-2003-511237-2), the Spanish Ministry of Education and Science (CTQ2006-15052) and the Catalan Government (2007BE-100162).

REFERENCES

- Arheimer, B., and J. Olsson, *Integration and coupling of hydrological models with water quality models: applications in Europe*, Swedish Meteorological and Hydrological Institute, Sweden, 2003
- Arnold, J.G., R. Srinivasan, R.S. Muttiah, J.R. Williams, Large area hydrologic modelling and assessment part I: model development, *Journal of the American Water Resources Association*, 34 (1), 73-89, 1998
- Bathurst, J.C., A. Burton, B.G. Clarke, F. Gallart, Application of the SHETRAN basin-scale, landslide sediment yield model to the Llobregat basin, Spanish Pyrenees, *Hydrological Processes*, 20, 3119-3138, 2006
- Bro, R., PARAFAC. Tutorial and applications, *Chemometrics and Intelligent Laboratory Systems*, 38, 141-171, 1997
- Felipe-Sotelo M., L. Gustems, I. Hernàndez, M. Terrado, R. Tauler, Investigation of geographical and temporal distribution of tropospheric ozone in Catalonia (Nort-East Spain) during the period 2000-2004 using multivariate data analysis methods, *Atmospheric Environment* 40, 7421-7436, 2006
- Fortin, J.P., R. Moussa, C. Bocquillon, J.P.Villeneuve, Hydrotel, a distributed hydrological model compatible with remote sensing and geographical information systems, *Revue des Sciences de l'Eau*, 8, 97-124, 1995
- Joliffe, I.T., *Principal Component Analysis*, Springer, New York, 1986
- Lavigne, M.P., Mémoire M.Sc. en Science de l'Eau, *Modélisation du régime hydrologique et de l'impact des coupes forestières sur l'écoulement du ruisseau des eaux-volées à l'aide d'Hydrotel*, 2006
- Massart, D.L., B.G.M. Vanderginste, L.M.C. Buydens, S. de Jong, P.J. Lewi, and J. Smeyers-Verbeke, *Handbook of chemometrics and qualimetrics: parts A and B*, Elsevier, Amsterdam, 1998
- Peré-Trepat E., A. Ginebreda, R. Tauler, Comparison of different multiway methods for the analysis of geographical metal distributions in fish, sediments and river waters in Catalonia, *Chemometrics and Intelligent Laboratory Systems* 88, 69-83, 2007
- Quilbé, R., A.N. Rousseau, GIBSI : an integrated modelling system for watershed management – sample applications and current developments, *Hydrology and Earth System Sciences Discussions*, 4, 1301-1335, 2007
- Rawls W.J., D.L. Brakensiek, *Estimation of soil water retention and hydraulic properties*, Unsaturated flow in hydrologic modelling, theory and practice, Kluwer Academic Publishers, 275-300, 1989
- Tauler, R., Multivariate Curve Resolution applied to second order data, *Chemometrics and Intelligent Laboratory Systems*, 30, 133-146, 1995.
- Tauler, R., A. Smilde, B. Kowalski, Selectivity, local rank, 3-way data-analysis and ambiguity in multivariate curve resolution, *Journal of Chemometrics*, 9, 31-58, 1995
- Tauler, R., I. Marques, E. Casassas, Multivariate curve resolution applied to three-way trilinear data: study of spectrofluorimetric acid-base titration of salicylic acid at three excitation wavelengths, *Journal of Chemometrics*, 12, 55-75, 1998
- Tauler, R., Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution, *Journal of Chemometrics*, 15, 627-646, 2001
- Terrado, M., M. Kuster, D. Raldúa, M.L. de Alda, D. Barceló, and R. Tauler, Use of chemometric and geostatistical methods to evaluate pesticide pollution in the irrigation and drainage channels of the Ebro river delta during the rice-growing season, *Analytical and Bioanalytical Chemisty* , 387, 1479-1488, 2007
- Turcotte, R., J.P. Fortin, A.N. Rousseau, S. Massicotte, J.P. Villeneuve, Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network, *Journal of Hydrology*, 240, 225-242, 2001

- Turcotte, R., A.N. Rousseau, J.P. Fortin, J.P. Villeneuve, A process-oriented, multiple-objective calibration strategy accounting for model structure, *Water Science and Application*, 6,153-163, 2003
- Turcotte, R., L.G. Fortin, V. Fortin, J.P. Fortin, J.P. Villeneuve, Operational analysis of the spatial distribution and the temporal evolution of the snowpack water equivalent in Southern Québec, Canada, *Nordic Hydrology*, 38 (3), 211-234, 2007
- Villeneuve, J.P., C. Blanchette, M. Duchemin, J.F. Gagnon, A. Mailhot, A.N. Rousseau, M. Roux, J.F. Tremblay, R. Turcotte, *Rapport final du projet GIBSI*, INRS-ETE, 1998
- Wischmeier W.H., *Predicting rainfall erosion losses: a guide to conservation planning*, US Department of Agriculture, 1978

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A data mining approach to discover weather patterns contributing to PM10 exceedances

Sfetsos A. [(a)], **Vlachogiannis D.** [(a)]

(a) Environmental Research Laboratory,
Institute of Nuclear Technology - Radiation Protection, NCSR "DEMOKRITOS",
15310 Aghia Paraskevi Attikis, Athens, Greece
(ts, mandy@ipta.demokritos.gr)

Abstract: The present study presents a data mining methodology to discover weather characteristics that are associated with exceedances of particulate matter (PM) with diameter less than 10 μm. The proposed approach is an eigenvector one consisting of a rotated principal components algorithm coupled with the fuzzy c-means clustering algorithm. This study examines only those days where at least one station exhibited an exceedance of the limit of a daily average value of 50 $\mu g/m^3$. The analysis was conducted using daily data from the monitoring network of the Grater Athens Area that were complemented with meteorological data from the NCEP Global Forecasting System (GFS). The analysis concluded that the PM10 exceedances in the Athens Area can be classified into 10 distinct types with varying spatial characteristics and weather contribution.

Keywords: data mining; climatology; PM10; exceedances.

1. INTRODUCTION

According to recent EU position papers (EC Directive 99/30/EC) PM is one of the top priority air pollutants that adversely influence human health (e.g. Kappos et al. [2004]). Climatological approaches have been reported with considerable success in relating atmospheric conditions to air pollution concentrations (e.g. Davis and Gay [1993]; Comrie [1994]. Eigenvector approaches have been applied in similar application that aim to identify group of days with similar characteristics, relating weather types to air quality patterns, such as those of Kasomenos et al. [1998] for the Attica basin, for Moscow (Shahgedanova et al. [1998]). Greene et al. [1999] developed the "Temporal Synoptic Index" to categorize weather in four US cities and related the resulting weather types to air pollution concentrations and human morbidity.

The Greater Athens Area (GAA) has a population of over 4.5 million inhabitants. The main pollution problem is attributed to the heavy traffic conditions (over 2.5 million registered vehicles and steadily decreasing infrastructures) taking into account that a large number of these vehicles are non-catalytic or old technology diesel engines (Grivas et al. [2008]). Space heating is the primarily done by fuel oil and only recently natural gas has been introduced in the energy mix with low penetration levels. Three main industrial zones exist in the region, which are located to the West (Thriassion plain including two oil refineries, cement plants and a steelworks factory), S–SE (electricity generating unit) and the North (secondary industrial operations and electricity production). According to a recent study by Mirasgedis et al. [2008], the external costs associated with industrial activities amount to 211M€ per year, representing 0.5% of the GDP and 4.1% of the Gross Value Added by the industrial activities in the area of interest in 2000 and are mainly associated with human mortality and morbidity primarily due to PM10 emissions.

This study presents the application of an eigenvector methodology for the identification of weather characteristics and the spatial analysis of exceedances of PM10 occurring in at least one monitoring location in the Greater Athens Area. According to Yarnal et al. (2001) eigenvector-based approaches consists of two phases. Once the available data are selected, a dimension reduction algorithm is applied so that only useful information from the original data set is conveyed. Then a clustering algorithm is applied to determine groups of data that have common variational properties.

2. METHODOLOGY DESCRIPTION

In the present study, the selected variables were transformed to have zero mean and untiy variance and thus common scaling. The scaling factors were determined using only those values that are above the regulatory limit, since these constitute the focus of the study and their potential impact is far more stressed.

The dimension reduction is achieved using a varimax rotated Principal Components Analysis (rPCA), where the number of new independent components is determined by setting a cut-off threshold in the estimated eigenvalues. In this study the *eigenvalue greater than one rule* suggested in Shahgedanova et al. [1998], was applied in order to determine the number of variables used. Several other alternatives were tested such as factor analysis and the positive matrix factorization, but did not return significant differences in the number of independent components nor their properties.

The selection of the clustering algorithm is very important in related studies, as it effectively identifies weather types. Following the work of Kalkstein et al. [1987] the average linkage clustering was favoured, although in recent years many authors shifted to the *k*-means algorithm and more recently to variations of the subtractive clustering algorithm. In the present study the fuzzy c-means algorithm (e.g. Besdek [1981]) was applied and the number of clusters was found using the "compactness and separation" criterion (Kim et al. [2001]).

3. AREA OF APPLICATION AND DATASETS USED

The weather conditions in GAA are mainly the results of interaction of large and local scale circulation systems. The climate of the GAA is typically Mediterranean with hot dry summers and wet mild winters. According to Ministry of Environment annual reports particulate matter is the only pollutant that exhibits persistent exceedances in GAA. The table below shows the annual mean concentration of PM10 in the GAA monitoring locations (MinEnv 2007).

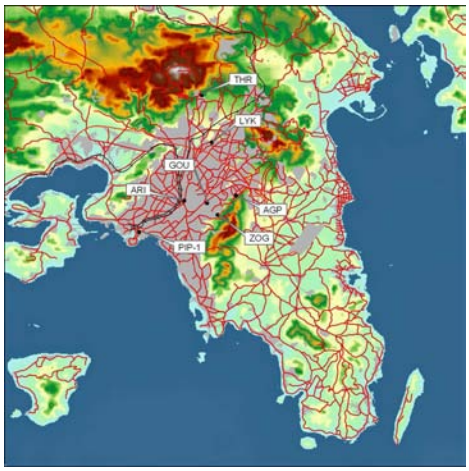


Figure 1. Monitoring Locations and GFS grid points.

Table 1. PM10 Annual Mean Concentration in μg/m^3 (MinEnv 2007)

YEAR	PIR-1	MAR	ZOG	LYK	AGP	ARI	GOU	THRA
2001	57	55	35	60	47	55	50	31
2002	63	69	35	62	38	55	53	34
2003	54	38	34	59	37	56	22	32
2004	56	29	33	63	39	58	18	33
2005	54	46	29	53	41	53		32
2006		48	26	59	34	57	34	27

Data for a six-year period (from 2001–2006) were used in this study. The air quality data stations, portrayed in Figure 1, are operating on a continuous basis under the supervision of the Greek National Air Pollution Monitoring Network. Only data from the stations of (i) Aristotelous (ARI), (ii) Agia Paraskevi (AGP), (iii) Lykovrisi (LYK), (iv) Thrakomakedones (THR) and (v) Zografou (ZOG) were used because they contained an adequate number of samples to conduct the analysis. The meteorological data were obtained from the NCEP reanalysis data set (http://dss.ucar.edu/datasets/ds083.2/) over the respective period of time on a daily basis at 00:00 and 12:00 hours. These were selected to have a general overview of the weather patterns and midday is period were highest values are observed at the suburban stations (MinEnv 2007). The selected variables are presented in Table 2 and their correlation coefficients in Table 3 are: the mixing layer height (MLH) which is crucially associated with dispersion of air pollutants, temperature (T) at 2m and the E-W (U) and N-S (V) components of wind speed

Table 2. Selected variables, and statistical properties

		Mean	STD	Units
00:00	***T***	290.2504	5.7588	K
	MLH	507.2484	384.8502	M
	U	1.3705	2.3071	m/s
	V	-1.6306	3.421	m/s
12:00	**T**	293.9041	5.9127	K
	MLH	1013.012	435.6832	M
	U	-1.3874	2.2443	m/s
	V	-0.5165	4.7	m/s
Daily	**ARI**	64.9458	25.9312	μg/m^3
	AGP	49.2586	27.4576	μg/m^3
	LYK	70.9209	27.3346	μg/m^3
	THR	36.2391	28.4849	μg/m^3
	ZOG	39.137	23.5763	μg/m^3

4. RESULTS

The application of the r-PCA approach returned 4 new variables (each named as PCi, i=1…4) that explained approximately 77.64 % of the variance of the initial data. The variables that had a significant contribution to each loading are described in Table 4. It may be observed that the resulting loadings are organized into well related variables at every interval on the day. PC1(define) is dominated by the PM. PC2 is related to the T parameters at both time intervals, whereas PC3 is dominated by the N-S and the E-W wind components. Finally, PC4 is dominated by the MLH parameters at both intervals and the E-W (*U*) component at 12:00. Table 3 provides an explanation of the grouping of similar variables into a new loading since they exhibit the highest correlation, in excess of 0.7. The correlation coefficient of the resulting factors is below 0.2 for every possible combination of the resulting four variables.

Table 3. Correlation Coefficient of examined variables

	T	**MLH**	**U**	**V**	**T**	**MLH**	**U**	**V**	**ARI**	**AGP**	**LYK**	**THR**	**ZOG**
00:00													
T	1.00	-0.19	-0.01	-0.17	0.94	-0.04	-0.30	-0.19	-0.15	0.19	-0.12	0.07	0.15
MLH	-0.19	1.00	-0.16	-0.31	-0.31	0.47	0.20	-0.38	-0.18	-0.11	-0.01	-0.07	-0.05
U	-0.01	-0.16	1.00	0.31	0.08	0.03	0.31	0.28	-0.06	-0.12	0.00	-0.02	-0.17
V	-0.17	-0.31	0.31	1.00	-0.07	0.00	0.34	0.72	0.28	0.16	0.24	0.20	0.12
12:00													
T	0.94	-0.31	0.08	-0.07	1.00	-0.01	-0.27	-0.04	-0.15	0.21	-0.09	0.10	0.14
MLH	-0.04	0.47	0.03	0.00	-0.01	1.00	0.32	-0.07	-0.18	-0.03	-0.05	0.00	-0.02
U	-0.30	0.20	0.31	0.34	-0.27	0.32	1.00	0.28	0.07	-0.12	0.17	-0.02	-0.12
V	-0.19	-0.38	0.28	0.72	-0.04	-0.07	0.28	1.00	0.29	0.17	0.27	0.27	0.13
Daily													
ARI	-0.15	-0.18	-0.06	0.28	-0.15	-0.18	0.07	0.29	1.00	0.58	0.79	0.58	0.62
AGP	0.19	-0.11	-0.12	0.16	0.21	-0.03	-0.12	0.17	0.58	1.00	0.66	0.73	0.87
LYK	-0.12	-0.01	0.00	0.24	-0.09	-0.05	0.17	0.27	0.79	0.66	1.00	0.65	0.73
THR	0.07	-0.07	-0.02	0.20	0.10	0.00	-0.02	0.27	0.58	0.73	0.65	1.00	0.79
ZOG	0.15	-0.05	-0.17	0.12	0.14	-0.02	-0.12	0.13	0.62	0.87	0.73	0.79	1.00

Table 4. Rotated Factor Loadings

		PC1	**PC2**	**PC3**	**PC4**
00:00	T	0.0052	-0.654	-0.0542	-0.0184
	MLH	-0.0311	0.1418	-0.2626	-0.5576
	U	0.1618	-0.1205	0.4638	-0.103
	V	-0.0707	0.0236	0.5402	0.0435
12:00	T	0.0117	-0.6731	0.0458	-0.0017
	MLH	0.0011	-0.1238	0.0487	-0.6791
	U	0.0134	0.1244	0.3688	-0.4241
	V	-0.0823	0.0361	0.522	0.1217
Daily PM10	ARI	-0.521	0.1104	0.0372	0.1061
	AGP	-0.4509	-0.1659	-0.0428	-0.0241
	LYK	-0.5673	0.0373	0.0415	-0.0837
	THR	-0.2281	-0.0555	0.0118	-0.0267
	ZOG	-0.3348	-0.0882	-0.0566	-0.0294

The number of clusters that were found using the compactness and separation criterion was equal to 10 (Figure 2). Then depending on the resulting cluster centers, all data were attributed to the nearest center using the Euclidean distance. The resulting percentage appearance of each resulting cluster is estimated and displayed on Figure 3.

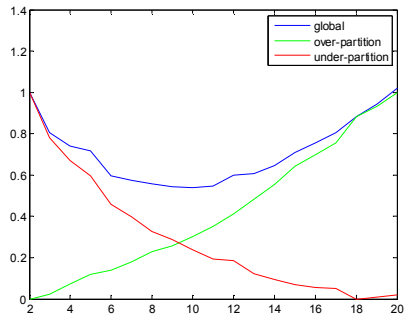


Figure 2. Number of clusters

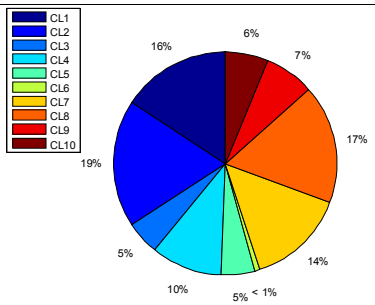


Figure 3. Data distribution in clusters

The resulting clusters are presented in Table 5. The Table shows the corresponding values of the examined variables of each cluster and the nearest real date from the modeling data. These dates are only indicative and they closely, but not entirely, match the observed conditions, and they are intended to be used in future detailed studies with numerical weather and atmospheric pollution models.

Table 5. Cluster characteristics

	00:00				12:00			
	T	**MLH**	**u**	**v**	**T**	**MLH**	**u**	**v**
	C	m	m/s	m/s	C	m	m/s	m/s
CL1	22.95	451.45	0.96	-3.60	26.42	988.04	-2.83	-3.17
CL2	21.82	233.49	2.40	-0.17	26.27	931.95	-1.67	1.57
CL3	16.99	856.40	4.28	1.50	21.57	1825.04	2.68	2.67
CL4	11.17	980.44	-0.80	-5.34	13.25	1134.71	-1.64	-5.72
CL5	16.82	361.70	1.51	1.44	21.12	967.05	-0.84	3.93
CL6	16.18	963.77	-2.79	1.48	19.29	1434.79	-0.85	4.24
CL7	10.78	422.38	1.95	0.53	14.74	859.71	-0.31	2.55
CL8	15.66	273.84	0.85	-1.60	19.29	627.40	-2.51	-0.01
CL9	22.02	1082.28	-0.59	-7.28	24.16	1508.53	-2.40	-8.76
CL10	10.55	743.61	3.18	0.97	14.64	1385.54	1.81	2.44

Table 5. Continued

	ARI	**AGP**	**LYK**	**THR**	**ZOG**	**"Matching Day"**
	$\mu g/m^3$	$\mu g/m^3$	$\mu g/m^3$	$\mu g/m^3$	$\mu g/m^3$	
CL1	51.38	48.39	57.32	30.57	37.07	18-Jun-02
CL2	65.25	57.15	71.25	42.71	43.40	31-Aug-04
CL3	53.42	42.52	71.98	36.74	32.22	15-Nov-04
CL4	58.31	34.62	63.99	21.77	30.63	9-Mar-03
CL5	107.98	96.09	119.23	84.89	80.46	21-Apr-01
CL6	228.84	248.52	267.71	236.54	221.95	17-Apr-05
CL7	73.06	39.55	74.68	31.09	31.26	29-Nov-04
CL8	69.90	46.32	68.95	32.18	36.44	12-Nov-04
CL9	43.03	48.83	57.68	30.46	40.94	12-Jun-03
CL10	57.67	27.38	66.43	22.32	20.63	10-Jan-01

The following paragraph shows a detailed analysis of the resulting clusters

CL1: Corresponds to warm conditions with low to moderate winds that change from NE to NW. High pollution levels are located in the city center (ARI) and LYK, which is located near the main highway, that are mainly from primary emissions.

CL2: Corresponds to sea breeze conditions which are frequently occurring during summertime. Low morning winds change into south components. High values are found in the city center (primary emissions) and in the stations nearest the mountains (THR, AGP) from transportation.

CL3: Corresponds to clear days (high MLH during midday) and moderate winds from SE. The peak is located near LYK, which is a combination of primary emissions and secondary from transportation from the city center.

CL4: Corresponds to cold days with predominant strong NW winds and relatively high MLH. The higher pollution values are found in LYK and ARI, which is mainly associated with primary production of PM10 from diverse sources.

CL5: Corresponding to very high PM10 levels in all stations, approximating a value between 80 and 100 $\mu g/m^3$ for all stations in the GAA. The other important is the dominant south wind component during all hours.

CL6: Corresponds to extremely high PM10 concentrations which occurs only a very few number of times (percentage < 1%). However pollution values in all stations is in the area of 200 $\mu g/m^3$.

CL7: Corresponds to cold days that are partly or totally clouded (low MLH values during all day). Also stationary conditions (low winds) with a mildly increasing SW component in the direction. The hotspots in the GAA are the city center (ARI) and LYK.

CL8: Corresponds to mild days that are partially cloudy or overcastted. The wind is low and blowing from NE in the morning changing to W. The associated with primary emissions are above the limit of 50 $\mu g/m^3$, but also AGP has values near that due to transportation.

CL9: Corresponds to warm days with clear skies (high MLH). The wind has strong component from the North, which is typical weather during summer days (etesian winds). The peak is found in the eastbound part of city (AGP, ZOG) and LYK.

CL10: Corresponds to cold and clear days with high MLH values. The winds are low and have an east component in addition to an increasing south one. The locations of primary emissions (ARI, LYK) are the only locations with high concentrations.

5. CONCLUSIONS

This study presented a data mining approach for analyzing the correspondence between exceedances of the national regulatory limit of PM10 and weather conditions in the Greater Athens Area. A rotated PCA approach was used coupled with the fuzzy c-means algorithm to discover closely matching patterns of air pollution and weather conditions, in addition to spatial diversifications between stations. The analysis revealed the presence of ten distinct weather / air pollution patterns.

REFERENCES

Bezdec, J.C., *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, 1981

Comrie, A.C., A synoptic climatology of rural ozone pollution at three forest sites in Pennsylvania, *Atmospheric Environment*, 28 (9), 1601-1614, 1994.

Davis, E.R., and D.A. Gay, A synoptic climatological analysis of air quality in the Grand Canyon national park. *Atmospheric Environment,* 27 (5), 713-727, 1993.

Grivas, G., A. Chaloulakou, and P. Kassomenos, An overview of the PM10 pollution problem, in the Metropolitan Area of Athens, Greece. Assessment of controlling factors and potential impact of long range transport, *Science of the Total Environ*ment, 389(1), 165-77, 2008.

Greene, J.S., L.S. Kalkstein, H. Ye, and K. Smoyer, Relationships between synoptic climatology and atmospheric pollution at 4 US cities. *Theoretical Applied Climatology*, 62, 163-174, 1999.

Kalkstein, L.S., G. Tan, and J.A. Skindlov, A comparison of three clustering procedures for use in synoptic climatological classification. *Journal of Climatology & Applied Meteorol*ogy, 19, 717-730, 1987.

Kappos, A.D., P. Bruckmann, T. Eikmann, N. Englert, U. Heinrich, P. Höppe, E. Koch, G.H.M. Krause, W.G. Kreyling, K. Rauchfuss, P. Rombout, V. Schulz-Klemp, W.R. Thiel, and H.E. Wichmann, Health effects of particles in ambient air, *International Journal of Hygiene and Environmental Health*, 207 (4), 399-407, 2004.

Kassomenos, P.A., H.A. Flocas, S. Lykoudis and M. Petrakis, Analysis of mesoscale patterns in relation to synoptic conditions over an urban Mediterranean basin, *Theoretical and Applied Climatology,* 59, 215-229, 1998.

Kim, D.J., Y.W. Park, and D.J. Park, A novel validity index for determination of the optimal number of clusters. *Institute of Electronics, Information and Communication Engineers Transactions of Information Systems*, E84-D, 281-285, 2001.

MinEnv 2007, The Air Pollution in Athens 2006 Annual Report, published by Ministry of Environment, March 2007, (in Greek).

Mirasgedis, S., et al. Environmental damage costs from airborne pollution of industrial activities in the greater Athens, Greece area and the resulting benefits from the introduction of BAT, *Environmental Impact Assessment Review*, 28 (1), 39-56, 2008.

Shahgedanova, M., T.P. Burt, and T.D. Davies, Synoptic climatology of air pollution in Moscow, *Theoretical Applied Climatology,* 61, 85-102, 1998.

Yarnal, B., A.C. Comrie, B.J. Frakes, and D.P. Brown, Developments and prospects in synoptic climatology, *International Journal of Climatology*, 21, 1923-1950, 2001.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Virtual Sources for Spatio-temporal Monitoring Data Analysis

A. Nuzhny, E. Saveleva, S. Kazakov, S. Utkin

Nuclear Safety Institute Russian Academy of Sciences, 52 B. Tulskaya, Moscow, 113191, Russia (nuzhny@inbox.ru, esav/kazak/uss@ibrae.ac.ru)

Abstract: The current work deals with application of independent virtual sources as a tool for analysis and storing groundwater monitoring data. Estimation of virtual sources is based on learning from data principle. Virtual sources allow to analyse the main features of the process. They can be useful for both data exploratory analysis and prediction. This work considers two methods for independent virtual sources construction: (1) based on using Central Limit Theorem and (2) based on the mixture of Gaussians. The methods were applied to real spatio-temporal data on groundwater level dynamics (2D spatial case) and groundwater contamination by radioactive nitrates (3D spatial case). These data sets are characterized by heterogeneous sample distribution both in space and time.

Keywords: Data mining, Independent sources, Gaussian distribution, Spatio-temporal data, Monitoring.

1. INTRODUCTION

Real monitoring systems usually deal with a problem of data analysis in spatio-temporal continuum One of possible ways is a data mining methodology featuring adaptive learning from data principle [Cherkassky and Muller, 1998]. Data mining provides a set of methods to handle, to model and to make predictions basing on measured data and additional knowledge somehow connected with data (physical laws, empirical rules, etc.). Diversity of data mining methods is very high covering all kinds of data analysis problems: regression, classification, clustering, etc.

The current work considers a data mining approach based on independent sources (independent components) [Lee T.-W.1998]. Generally independent component analysis solves the blind source separation problem by decorrelating the signals and reducing some higher-order statistical dependencies. Independent sources store enough information on the process for modelling and making predictions. Thus independent sources also can be useful as a compressing tool to store the monitoring data more effectively. Independent sources can be considered as one of possible non-linear generalizations of principle component analysis.

Classical independent sources problem is an inverse problem. A measured value (a receiver) is treated as a superposition of signals from some sources. The problem is to establish sources. As any inverse problem this one is an ill-posed problem with unlimited number of possible solutions. In this work independent sources are considered as virtual (not real) sources responsible for a process, thus any reliable solution is equally good.

This work discusses two approaches to find the independent virtual sources. Both of them deal with Gaussianity: (1) consider superposition more normally distributed in comparison with single signals (based on using Central Limit Theorem) and (2) consider the receiver

as a weighted superposition of Gaussians. Both methods have some theoretical limitations complicating their application to real data.

Methods were applied to analysis of real data from a groundwater monitoring system studying the vicinity of nuclear waste facilities.

2. THEORETICAL BACKGROUND

2.1 Method using the Central Limit Theorem

To formulate an independent sources problem let us consider values measured at spatio-temporal points x_i as receivers obtaining signals $c_i=c(x_i)$, which are a superposition of initial signals ($b_j=b(r_j)$) from some sources r_j. The task is to restore initial signals and their sources. Otherwise it is to find the coefficients w_{ij} in:

$$b_j = \sum_i w_{ij} c_i \quad (1)$$

under the additional constraint:

$$\sum_j w_{ij}^{\,2} = 1 \quad (2)$$

Such formulation of the independent sources problem requires the following assumptions: the sources are points, no spatial size; influence of sources on receivers is immediate; distribution of signals from sources is isotropic. The possibility to work under these assumptions should be considered before application of the method.

One of the possible ways to solve this problem is to use the Central Limit Theorem stating that the superposition of sources presents more normal distribution then single signals. The level of normality can be characterized by the excess kurtosis (the function of the fourth standardized moment) [Joanes and Gill, 1998]. Kurtosis equals to zero for normally distributed data, the positive kurtosis indicates more narrow probability density function in comparison with the normal one, and the negative kurtosis is due to wider probability density function then the normal one. Kurtosis (K) can be written in terms of matrix $\mathbf{W}=(w_{ij})$ from (1) as

$$K = \frac{\frac{1}{n}\sum_j \left(\sum_i w_{ij} c_i \right)^4}{\left(\frac{1}{n}\sum_j \left(\sum_i w_{ij} c_i \right)^2 \right)^2} - 3 \quad (3),$$

where n is a number of sources. Thus the problem is stated as the minimization of squared kurtosis (3) under the constraint (2).

Numerical optimisation under a constraint uses a penalty function. This function has zero value within the allowed range of values and it tends to infinity in the forbidden area. The penalty function for maximization under the constraint (2) looks as

$$\left(\sum_j w_{ij}^{\,2} - 1 \right)^2 \quad (4).$$

The final expression for optimization is thus:

$$\Phi = K^2 - \frac{1}{\mu}\left(\sum_j w_{ij}^{\,2} - 1 \right)^2 \quad (5),$$

with parameter μ monotonically decreasing during the optimization.

A number of parameters of the optimisation problem equals to a number of elements in matrix **W,** which strongly depends on the dimension of the initial data set. Together with the number of parameters grows the number of local optimisation solutions. The number of functional parameters of the optimisation problem can be reduced by a preliminary compressing data transformation (**AC**), for example, by transformation to principle components. The final transformation of data to independent sources in this case is performed by matrix

$$\boldsymbol{B} = \boldsymbol{WA} \quad (6)$$

Coefficients of matrix **B** describe the credit-debit dependencies between sources and measurement wells. A positive coefficient indicates a positive dependence: the increase of a source value causes an income of a value at the receiver. A negative coefficient is when a source value is increasing on account of the receiver. Thus matrix **B** is an important tool for analysis of the influence between sources and monitoring wells, especially in the case when sources can be associated with some facilities (for example, lakes or rivers in the case of water monitoring).

The matrix **B** also allows to estimate the locations of independent sources [Üzümcül et al., 2003]:

$$\boldsymbol{r}_j = \sum_i \boldsymbol{b}_{ij}(\boldsymbol{x}_i - \boldsymbol{x}_0) + \boldsymbol{x}_0 \quad (7),$$

here x_0 is a centre of mass over all measurement locations.

2.2 Gaussian mixture approach

The process under study can be approximated as a weighted superposition of Gaussian-like signals:

$$\boldsymbol{c}(\bar{\boldsymbol{r}},\boldsymbol{t}) = \sum_{t-1}^{n} \boldsymbol{A}_i \, exp\left(-\frac{\alpha_i}{\boldsymbol{t}-\boldsymbol{t}_i} \left\| \bar{\boldsymbol{r}} - \bar{\boldsymbol{r}}_i \right\| \right) \quad (8),$$

where $\bar{r}$ indicates a spatial coordinate, t is a temporal coordinate, n is a number of sources (user-defined parameter), A_i, α_i, $\bar{r}_i$ and t_i, ($i=1,\ldots,n$) are fitting parameters. Parameters $\bar{r}_i$ and t_i indicate a spatio-temporal coordinates of sources.

The model fitting procedure leads to an optimization problem. The model parameters (weights) are estimated so as to minimize the sum of squared residuals at the sample locations:

$$\boldsymbol{D} = \sum_k \left(\boldsymbol{c}(\bar{\boldsymbol{r}}_k, \boldsymbol{t}_k) - \boldsymbol{y}_k \right)^2 = \sum_k \left(\sum_i \boldsymbol{A}_i \, exp\left(-\frac{\alpha_i}{\boldsymbol{t}-\boldsymbol{t}_i} \left\| \bar{\boldsymbol{r}}_k - \bar{\boldsymbol{r}}_i \right\| \right) - \boldsymbol{y}_k \right)^2 \quad (9)$$

here y_k are known values, $k=1,\ldots,K$ with K being a number of samples. Minimization was performed by RPROP [Riedmiller and Braun, 1993] method. RPROP is an effective and stable iterative optimization algorithm. The direction of a parameter's changes depends only on the sign of the objective function's derivative, and the learning rate depends on the previous state of the system. Such construction of iterations sometimes allows to jump over local extremes and significantly decreases the time of the approximation in the monotonic case.

The function (8) with fitted parameters allows to make predictions at any spatio-temporal locations. Any kinds of dependencies between sample locations and sources are also available. The main drawback of this approach is the assumption made on the form of

signals emitted by virtual sources. From this point of view the first approach is more general, as it does not require any assumption on the form of signals.

3. DATA DESCRIPTION

Groundwater monitoring in the region connected with radioactive storing facilities controls groundwater level dynamics and water quality. We consider two case studies describing two different events: a groundwater level dynamics and groundwater contamination by radioactive nitrate. The groundwater level is a unique value per a well at a time, thus it presents a two-dimensional in space problem. Nitrate contents also depend on the depth of a sample and consequently lead to a three-dimensional in space problem. The main problem is to extract the principle data features allowing their compact description.

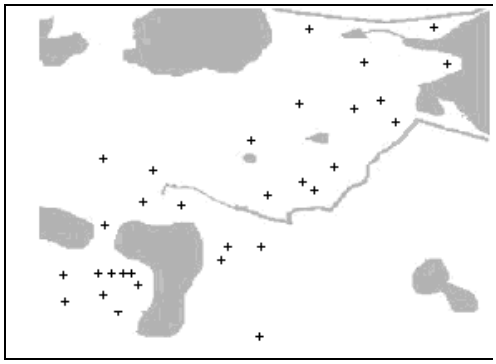


Figure 1. Spatial distribution of hydro-geologic wells

For our analysis we used 32 geological monitoring wells. Figure 1 presents spatial distribution of these wells together with the main hydrologic objects of the region. Each well contained at least one groundwater level measurement per month during period from April 1970 till January 2006. Total number of groundwater samples was 15673. In [Nuzhny et al, 2007] a linear feature extraction approach was discussed (principle component analysis). First principle component responsible for the averaging over the well gives satisfactorily good approximation for 32 spatially distributed wells. Other components did not significantly influence on that approximation. Thus, features with physical meaning were decided to be searched through non-linear description. Several not strong assumptions led us to a classical independent sources problem.

Nitrate samples were even sparser - an average period between two measurements in the same position is a year during the period with the highest sampling density (1964 - 2003 years). The total number of nitrate samples was 1576. Taking into the account the curse of dimensionality it is obvious that traditional data analysis fail for these data. The most reasonable approach in such case is a parametric one. Gaussian-like approximating functions are the most common for a parametric description.

4. RESULTS AND DISCUSSION

4.1 Virtual sources for groundwater levels

A temporal cut for a groundwater level can be approximated with 98% confidence using 10 principle components [Nuzhny et al, 2007]. Correlation of the first principle component with precipitation data shows the direct effect - instantaneous impact on the precipitation (maximum of correlation at zero). Other principle components present different structure of correlation, thus they are after other physical phenomena. So estimation of independent sources for groundwater level was performed by minimization of kurtosis (3) using 9 principle components. The purpose was to delete direct influence of precipitation and to focus on analysis of other dependencies.

As for theoretical assumptions of the method the rudest among them is the assumption of isotropy, as anisotropy is a usual case in hydrogeology. But still such assumption is often made for preliminary analysis and it does not seriously distort results. The immediate reply of the receiver on the source is fulfilled as we have a monthly step in our data, so instantaneous effect means the effect within the month.

Figure 2a shows spatial distribution of 9 independent sources, they are marked by pluses. All of them can be associated with surface water reservoirs (lakes and a river). Figure 2b shows histograms of lines from matrix **B** (6) corresponding to the wells marked in fig.2a as filled circles. Analysis of matrix **B** allows to get some information on possible influence of contaminated artificial reservoirs on the water in wells. In terms of groundwater levels problem matrix coefficients obtain real meaning: positive coefficient indicates that growing of water level in the source causes an income of water to the well; negative coefficient means that the source takes water from the well.

For example, the well number 18 gets impact from sources 2,3 and 7 associated with the same lake and the source 9 associated with the river. Sources 1 and 5 (both associated with contaminated lakes) take water from this well. So, we can conclude that water from contaminated lakes do not get to this well. But still sources 2,3 and 7 can be influenced by contamination via well 7 taking water from the source 5. Such results are the subject for further analysis by experts and decision makers working at the site.

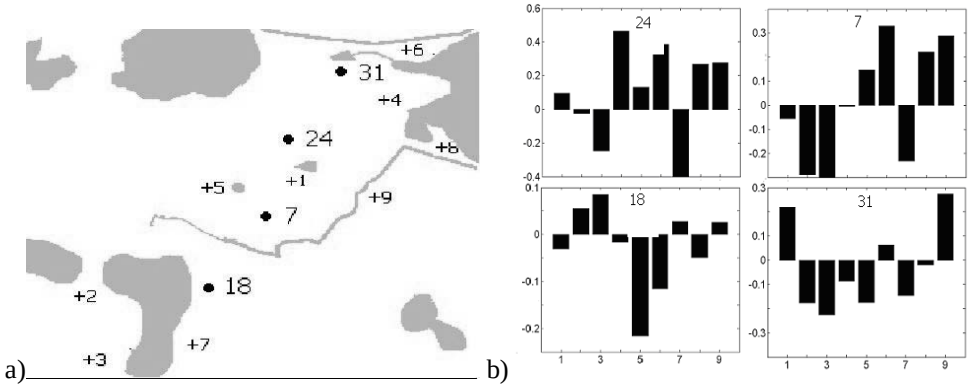


Figure 2. Results on the independent sources analysis: a) Schema of water reservoirs, 4 wells (filled circles) and 9 virtual sources (pluses); b) Histograms of lines of matrix **B** corresponding to 4 wells

So, we estimated virtual sources responsible for groundwater level dynamics. These sources appeared to be useful tools for studying influences between different water subsystems.

4.2 Virtual sources for nitrate contamination

Virtual sources for groundwater contamination by nitrates were estimated by Gaussian mixture approach (8). As initial values of parameters A_i and α_i we took small values (0.001), initial centres (sources) $\vec{r}_i$ were randomly distributed over the region. Examples of initial distributions of 8 and 4 centres are presented in figure 3 (rhombus), filled circles indicate spatial distribution of sample wells. Stability of RPROP algorithm guarantees the stability of the final result. Several initial random distributions of sources were used, and they led to very close results.

We started with 8 sources but the obtained result indicated surplus of such model - several sources united into one (fig. 4a). So finally the model with 4 sources (20 fitting parameters) was used (fig. 4b). The quality of the model can be checked by the relative error:

$$\boldsymbol{R} = \frac{\boldsymbol{D}}{\sum_{\boldsymbol{k}} \boldsymbol{y}_{\boldsymbol{k}}^2} \quad (10)$$

where ***D*** is defined by (9). Final relative error for model with 8 sources is ≈ 0.13, and with 4 sources - ≈ 0.2. Twice decreasing the complexity of the model we nearly preserved its quality.

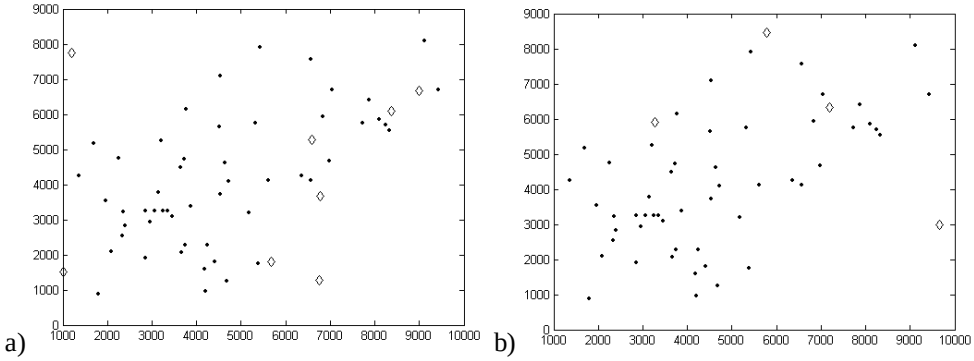


Figure 3. Examples of initial distribution of virtual sources (rhombus): 8 sources (a) and 4 sources (b). Filled circles indicate sample wells.

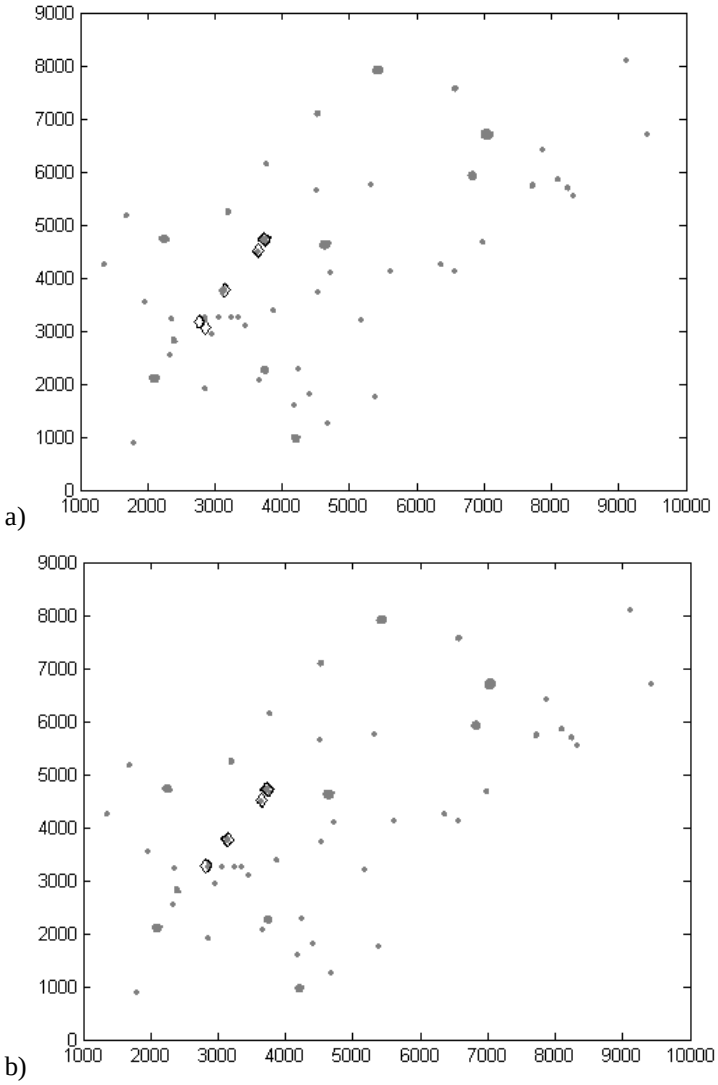


Figure 4. Final distribution of virtual sources (rhombus): 8 sources (a) and 4 sources (b). Grey filled circles present sample wells with diameter depending on local measurements' density

Positioning of virtual sources appeared to be very stable. But it is not because of denser sampling. In figure 4 monitoring wells are drawn as circles with diameters depending on the number of samples in the well. One can see rather frequent observed sample wells remote from virtual sources. Thus spatial location of virtual sources is due to the process.

Estimated virtual sources allow to describe the trend of the nitrate contamination by (8) and to make predictions on future development of the nitrate contamination. The 4 virtual sources indicate the zone responsible for the nitrate contamination process. Detailed monitoring in the vicinity of sources can be proposed to site administration.

4.3 Discussion on virtual sources

Two studied examples have different general features: groundwater level generally characterizes the water system; on the contrary the contamination process is caused by some real originating impact. This diversity arises within virtual sources analysis.

Considering obtained virtual sources due to these processes we find the approvement to such point of view. Groundwater level's virtual sources are distributed over the whole monitoring area somehow linking to surface water reservoirs and perhaps some groundwater ones.

Contamination virtual sources concentrate in the vicinity of the contaminating origin. All of these sources tend to the same area independently of their initial distribution. Geographically contamination virtual sources are located between two contaminated lakes, which can be the real origins of the groundwater contamination. Thus, virtual sources appeared to be not so virtual, but indicating to real features.

5. CONCLUSIONS

Independent virtual sources appeared to be useful tool allowing to describe

- dependencies between groundwater monitoring wells and surface water reservoirs. This information can be useful for analysis possibility of water migration in the neighbourhood of the contaminated water reservoirs.
- the localisation of the origin of the groundwater contamination.
- the model for groundwater contamination predictions.

Virtual sources also help to understand the general features on the process under study.

ACKNOWLEDGEMENTS

The work was partly supported by a grant 07-08-00257 of the Russian fund for fundamental researches (RFFI).

REFERENCES

Cherkasski V., Muller F., *Learning from data*, John Wiley Interscience, New York, pp. 437, 1998.

Joanes, D. N., Gill, C. A., Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society (Series D): The Statistician* V. 47: 183–189, 1998.

Lee T.-W., *Independent Component Analysis. Theory and Application*. Kluwer Academic Publishers, Boston, pp. 210, 1998.

Nuzhny A., Savelieva E., Jastrebkov A., Statistical Analysis for Extracting Features on the Groundwater Level Dynamics, in *Proc. of IAMG2007, Geomathematics and GIS analysis of resources, Environment and Hazards*, eds. P. Zhao, F. Agterberg, Q. Cheng, pp.723 - 726, 2007.

Riedmiller M., Braun H., A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm, in *Proc. of the IEEE Intl. Conf. on Neural Networks*, 1993

Üzümcül M., Frangi A.F., Reiber J.H.C., Lelieveldt B.P.F., Independent Component Analysis in Statistical Shape Models. Medical Imaging 2003, in *Proceed. of SPIE*, V.5032: 375-83, 2003.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Applying Data Mining Methods for Forest Planning Data Validation

A.Mäkinen[a], A.Kangas[a] and T.Tokola[b]

[a]*Department of Forest Resource Management, University of Helsinki, Latokartanonkaari 11 00014 University of Helsinki, Finland (antti.makinen@helsinki.fi)*

[b]*Faculty of Forestry, University of Joensuu (timo.tokola@joensuu.fi)*

Abstract: Decision making in forest planning is based mostly on simulated forest management scenarios. A fundamental tool in creating these scenarios is the forest planning system, which utilizes a set of models for projecting the future development of forests and assessing the effects of alternative management tasks, such as timber harvests [Burkhart 2003]. Input data for forest planning is obtained from several different sources, such as remote sensing and field measurements and visual assessments. All data collection systems include errors, both due to human and technical sources, which eventually affect the quality of forest plans. Part of the errors can be considered as next to impossible to detect but part of the errors are outliers and can be separated from the data. Statistical outlier detection methods have been used in data processing, although the statistical methods do not work well for multi-dimensional data. Data mining offers some interesting possibilities for the outlier detection task. The different data mining schemes for outlier detection include for example distance-, density- and clustering –based algorithms that have been proven to work with multi-dimensional data. In the field of forest research, data mining methods have not been studied almost at all. In this study we compared three different outlier detection schemes for finding the outliers in a large forest inventory data. The tested algorithms were Nested-Loop distance-based outlier detection [Knorr & Ng 1998], Simple-Pruning distance-based outlier detection [Bay & Schwabacher 2003] and Outlier Removal Clustering [Hautamäki et al. 2005]. The data included a total of 5090 field measured sample plots on 578 forest stands with a number of stand-level aggregate attributes representing different characteristics of the growing stock, the forest site and the surrounding region. Each of the examined methods has a number of parameters having a strong effect on the outlier detection result. Also the selection of the attributes which were used in the outlier detection strongly affected the results. None of the three methods proved to be superior compared to the others in finding the outliers. The large natural variation in the forest attribute values made the task of separating the outliers difficult. However, the examined data mining methods showed very promising results in finding outliers.

Keywords: *Forest Planning; Data Mining; Outlier Detection*

1. INTRODUCTION

The purpose of forest planning is to offer support for the decision maker [Buongiorno & Gilless 2003]. Generally the goal in forest planning is to find the best way to manage forest treatment units, forest stands. Forest planning systems are computer decision support tools that are used for evaluating alternative forest treatment scenarios. The planning systems include models for predicting the future development of forests, eg. forest growth models and models for predicting the possible outcomes from alternative forest treatment scenarios, eg. yield models [Burkhart 2003]. Usually some kind of optimization technique is applied for finding an optimal forest management scenario which will maximize the benefit of the decision maker. Errors in the inputs of a forest growth model will result in errors the predictions about the future. This in turn might lead to sub-optimal forest management, causing for example economical losses or losses in the social or ecological

values of the forest. This is why the quality and reliability of the forest planning data has a major effect on the quality of forest planning.

Most forest planning systems require some basic information about the forest site properties, the growing stock and the geographical area as inputs. Attributes that contain information about the forest site can include for example soil type and growth potential. Information about the growing stock is usually given as a list of trees and tree properties or some aggregated stand-level mean attributes such as mean diameter, mean height, volume of trees and main species.

Data for forest planning in Finland is collected mainly through partly subjective visual estimation of stand-level aggregate attributes. The method, though cost-effective, is highly subjective and prone to human errors [Haara 2003, Kangas et al. 2004]. In this method the expertise and experience of the forest planner has a strong effect on the quality of the data. The quality of the data can rarely be verified, except by labour intensive and costly reference measurements in the field. As the data for forest planning is acquired from multiple sources and in many cases without good estimates on the data quality, forest planning would benefit from tools that could be used to identify at least the coarsest errors in a data set.

The quality of the forest planning data varies between different forest types and inventory methods. In stand-wise visual estimation, the errors in the data can be caused by sampling, assessment and classification errors [Haara 2003]. According to different studies, the relative mean square errors (RMSE) for estimation accuracies of different aggregate attributes in stand-wise subjective field estimation range from 10 to 80 percent, with large variation between different attributes. For example the RMSE for mean height of trees has been reported to range from 12 to 17 percent whereas the RMSE for number of trees per hectare has been around 80 percent. [Poso 1983, Laasasenaho & Päivinen 1986, Kangas 2004]. The correlation of estimation error between different variables or the estimation error for categorical variables, such as soil type, has not been reported.

Recent developments in various forest inventory technologies, mostly in remote sensing, provide different types of data that varies in accuracy and reliability. With novel remote sensing techniques, such as airborne laser scanning, the RMSE's have been reported to range from around 5 to 50 percent. With laser scanning, the RMSE for mean height of trees has been reported to range from 3 to 15 percent and the RMSE for number of trees per hectare has been between 13 and 74 percent according to different studies [Naesset 2002, Maltamo et al. 2004]. The different remote sensing techniques have shown very promising results in forest inventory accuracy, but still at the moment the prevailing method for forest planning data collection in Finland is the stand-wise subjective field sampling.

Coarse errors in forest planning data might be detected by searching for outliers as the individual observations with large estimation errors might stand out from the rest of the data. Some of the problems with traditional statistical or distribution-based outlier detection methods are that they require a lot of manual work, demand prior information about data distributions and cannot handle multi-dimensional data [Barnett & Lewis 1994]. Detection of outliers has been studied also in the data mining research field and is considered to be an important data mining task [Knorr 2000]. Outlier detection is one of the many applications of data mining and it has been successfully utilized in many applications, such as fraud detection and computer network intruder detection [Hodge & Austin 2004, Han & Kamber 2006]. As the data sets in forest planning can be large and usually multi-dimensional, the data mining –based outlier detection methods could possibly provide good tools for finding errors in forest planning data.

This paper describes how different data mining –based outlier detection algorithms could be used to detect coarse errors from a typical forest planning data. A method that could automatically identify a significant proportion of the coarse errors in the input data would greatly improve the quality and reliability of forest planning.

2. OUTLIER DETECTION WITH DATA MINING

Data Mining (DM) is an umbrella term which covers a number of different methods for finding frequent patterns end previously unknown information from large amounts of data automatically. Data mining is considered to be a part of Knowledge-Discovery in Databases (KDD) process. Data mining methodologies are usually related to different scientific core fields such as computer science, statistics and machine learning. Common DM tasks include classifying, clustering, association rule mining and outlier detection [Piatetsky-Shapiro 1991, Han & Kamber 2006]. Different types of DM tasks can be divided into four main categories according to Knorr & Ng [1998]: 1. dependency detection, 2. class identification, 3. class description and 4. exception/outlier detection. In the first three categories the focus is on the majority of the objects in the data, where as the fourth category concentrates on a small portion of the whole data.

One definition for outliers is that an outlier is *an observation which appears to be inconsistent with the remainder of that set of data* [Barnett & Lewis 1994]. According to Papadimitriou [2002] different approaches for identifying outliers in a data set can be classified into distribution based, depth-based, clustering-based, distance-based and density-based techniques. The distribution-based techniques have their background in statistics and they utilize a distribution model fitted to the data and single out objects deviating from the distribution as outliers [Hawkins 1980, Barnett & Lewis 1994]. The distribution-based methods are usually univariate and need a priori assumptions about the data distributions. Depth-based techniques utilize computational geometry to compute layers of *k-dimensional* convex hulls and use the layers to divide the objects into outliers and normal objects [Johnson et al. 1998]. The depth-based methods suffer from the curse of dimensionality and cannot be used with high-dimensional data.

Usually the exceptions or outliers are considered to be mere annoyances, which the different data mining algorithms try to ignore. However, in many applications the outliers are actually the most interesting objects as they can lead to the discovery of new and meaningful knowledge [Knorr 2000]. Number of different approaches for identifying outliers in large data sets have been proposed in recent years [Hodge & Austin 2004].

Distance-based outlier detection introduced by Knorr & Ng [1998] utilizes a distance measure between objects to identify the outliers. The distance between two objects is usually computed as euclidean or mahalanobis distance in a k-dimensional feature space. The outlyingness is considered to be a binary property so that an object either is or is not an outlier. Distance-based outlier detection algorithms normally view the data objects at global scale and compare an object to all other objects in the data, which makes them computationally heavy for large data sets. A lot of research has been done in order to improve the efficiency of these algorithms [Ramaswamy 2000, Bay & Schwabacher 2003].

Density-based outlier definition offers a bit different approach for identifying the outliers. Algorithms for density-based outlier detection compute a continuous Outlier Factor (OF) locally for each object which can be used to divide objects into outliers and normal objects by setting a threshold value. The outlier factors are computed locally for example as a mean or maximum distance from object O to the object's k nearest neighbours. In the proposed density-based outlier detection algorithms, the outlier factor has been defined as Local Outlier Factor (LOF) [Breunig et al. 2000] or Connectivity Outlier Factor (COF) [Tang et al. 2002]. Many of the proposed distance- and density-based outlier detection algorithms handle well both *large* and *high-dimensional* ($n > 5$) datasets [Hodge & Austin 2004].

Clustering is a group of unsupervised classification methods used extensively for exploratory data analysis by many disciplines. Some of the most common clustering methods include hierarchical single- and complete-link clustering and partitional k-means and k-medoids clustering [Jain et al. 1999]. Many of the clustering algorithms, such as DBSCAN and it's many variations use a density-based notion of clusters to classify the objects in the data set [Ester et al. 1996, Brecheisen et al. 2006]. Most clustering algorithms consider outliers as noise that should be dealt with by identifying and removing the outlying objects. Because of this, some clustering algorithms provide outlier detection as a side product of clustering or are designed entirely for the purpose of identifying the outliers [Tang & Khosgoftaar 2004, Hautamäki et al. 2005].

3. MATERIAL AND METHODS

3.1. Stand-wise estimation data and reference measurements

The data in the study consisted of stand-wise visual estimates conducted from a helicopter by an experienced forest planning professional, and reference sample plot measurements conducted in the field. The total number of forest stands in the data was 578 and the number of reference sample plots was 5090, an average of 8.8 sample plots per stand. For each stand, a number of attributes describing the average properties of the growing stock, site quality and surrounding environment, were estimated. The sample plots were located inside the stands in a grid formation and the same set of attributes as for the stands were measured by field groups. In this setting the stand-level estimates for the different attributes represents the normal accuracy of stand-wise subjective sampling. The reference measurements in the sample plots can be considered very precise and they represent the true values for the different attributes.

Altogether six attributes, four of them numerical and two categorical, were selected for this study. These particular attributes were chosen because their values affect strongly the behaviour of the growth models in a forest planning system. Coarse errors in the values of these attributes will lead to notable errors in the growth and yield predictions. The selected attributes and their measurement units, abbreviations and some mean statistics are listed in Table 1.

Table 1. Stand-level mean attributes and abbreviations used in the analysis

attribute	abbreviation	mean	median	stdev
mean diameter [cm]	D_{MEAN}	16.1	-	8.0
mean height [m]	H_{MEAN}	13.6	-	6.4
basal area per hectare [m^2/ha]	BA	16.7	-	7.6
mean age [years]	A_{MEAN}	61.1	-	27.6
development class	DC	-	3	-
forest site class	SC	-	4	-

As the sample plot measurements were accurate, we could compute the true values for the stand-level attributes for each stand. This was done by calculating means for the attributes from the sample plots that were located inside a given stand. After we had the stand-wise, possibly erroneous, estimate x_{ij} and the true value y_{ij} for each attribute, the relative to true error δ_{ij} of the stand-wise estimation was calculated for each stand i and each attribute j. The error computed for the categorical attributes was a binary value; *TRUE* if the stand-wise estimate was correct and *FALSE* if not.

After calculating the relative estimation errors for each stand and numerical attribute, the errors were classified into four classes by the absolute magnitude of the error $|\delta_{ij}|$, so that in class I $|\delta_{ij}| < 0.25$, class II $0.25 < |\delta_{ij}| < 0.50$, class III $0.50 < |\delta_{ij}| < 0.75$ and class IV $|\delta_{ij}| > 0.75$. The percentage of observations in each of the four error classes for each numerical attribute is shown in Figure 1. Only a small proportion, under five percent of the stands had relative estimation errors in error class IV. The proportion of observations relative error greater than 0.25, classes II to IV, ranged from 27.6 percent to 45.4 percent depending on the attribute in question.

In addition to the four numerical attributes, D_{MEAN}, H_{MEAN}, BA and A_{MEAN} Figure 1 has also column *MAX,* where each stand was classified into one of the four error classes by the maximum error in any of their four numerical attributes. Most of the stands, around 60 percent, had at least one attribute in error class II and only in 13 percent of the stands the maximum error class was I.

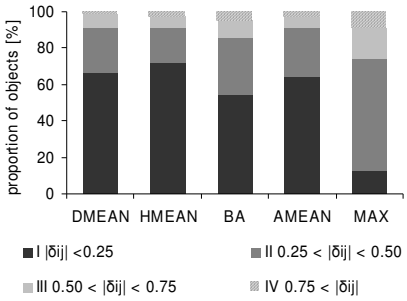


Figure 1. Proportion of objects in the four error classes for each of the four numerical attributes.

3.2. Outlier detection schemes

In this study, three different outlier detection algorithms were tested for identifying the coarse estimation errors in the visually estimated stand-wise data set. The applied algorithms were Nested-Loop distance-based outlier detection proposed by Knorr & Ng [1998], Simple Pruning distance-based outlier detection proposed by Bay & Schwabacher [2003] and Outlier Removal Clustering by Hautamäki et al. [2005]. These algorithms were selected because they are not univariate and because they do not require a priori information about data distributions. Forest stand data is multivariate and we do not necessarily have good a priori information about the data distributions.

The Nested-Loop distance-based outlier detection algorithm, *DB(p, D)* had parameters *p* and *D*, where an object *O* is considered to be an outlier if a proportion *p* of all objects are further than distance *D* away from *O*. The complexity of the Nested-Loop algorithm was $O(kN^2)$, where *k* is the number of attributes and *N* is the number of objects.

The Simple-Pruning distance-based algorithm was based on the Nested-Loop design, but randomization and a simple pruning rule were added for increased performance. The Simple-Pruning algorithm has showed nearly linear behaviour with large and multi-dimensional data sets. The parameters of the Simple-Pruning algorithm were *n* and *k*. The outliers were defined as the top *n* objects whose average distance to *k* nearest neighbours was greatest.

The Outlier Removal Clustering (ORC) consisted of two stages with were repeated multiple times. In the first stage, a k-means clustering was performed until convergence and in the second stage each object was assigned an outlyingness factor, which depended on the object's distance from the cluster centroid. The algorithm searched for the object with the maximum distance to the cluster centroid d_{MAX} and after that assigned the outlyingness factor o_i for each object x_i (1). Outlier Removal Clustering algorithm had parameters *k* which was number of cluster centroids and *T* which was the outlier threshold for o_i between 0..1.

$$o_i = \left\| x_i - c_{p_i} \right\| / d_{MAX} \tag{1}$$

The distance d_{xy} between two objects *x* and *y* were calculated as normalized euclidean distances (2) for the continuous attributes and as Hamming's distance between categorical attributes.

$$d_{xy} = \sqrt{\sum_{i=1}^{P} (x_i - y_i)^2 / \sigma^2} \tag{2}$$

The outlier detection algorithms were run multiple times with changing parameter values for the visually estimated stand-wise data. During each run, a number of stands were classified as outliers and the list of outliers was stored for later analysis. As the aim of this study was to test if the outlier detection algorithms could identify the stands with coarse estimation errors, the outlier lists from the different runs were examined against the classified errors. With this kind of cross-examination we could determine what proportion of the stands with coarse estimation errors, ie. errors in classes II-IV, had been defined as outliers and what proportion of the outliers did not have any coarse errors.

4. RESULTS

We executed several runs with different parameter values for all three outlier detection methods in order to see how the parameters affected which stands were classified as outliers. With Nested-Loop algorithm, increasing the value of parameter *p* from 0.9 to 0.995 decreased the proportion of outliers in all objects from 96.2 to 4.8 percent. Also the value of the parameter *D* had a strong effect on the proportion of outliers so that when the value of *p* was fixed to 0.95, increasing the value of *D* from 0.5 to 1.5 decreased the proportion of outliers from 74.7 to 6.2 percent. All three algorithms were affected strongly by the chosen parameter values.

The proportion of outliers in different error classes depended on the different parameter values in all algorithms. This was the most interesting part as the main idea of the tests was to find out if the stands with coarse errors have a higher probability to be classified as an outlier. A notable trend in the results was that the percentage of outliers was clearly higher in the stands where at least one variable had an estimation error in class III or IV. This behaviour can be seen in Figure 2, where the proportion of outliers in error class IV increases sharply as the proportion of outliers in all stands increases. Very similar behaviour was found in the performance of all three algorithms. Individual attributes seemed to affect the results to some extent. In all of the tests, the coarse, class III and IV, errors in the variables D_{MEAN} and *BA* increased the stand's probability to be classified as an outlier.

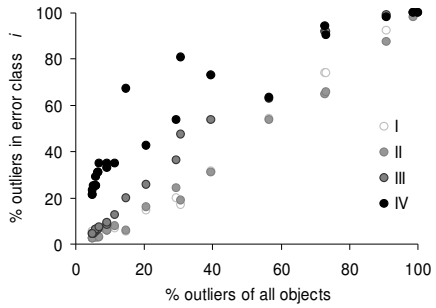


Figure 2. Relation between the percentage of outliers in all objects and percentage of outliers in each error class. Nested-Loop algorithm, 6 attributes.

We ran the outlier detection algorithms with two attribute combinations, where in the other combination the four continuous attributes were included in the outlier detection and in the other all six attributes were included. The difference in the results between the two attribute combinations was minimal as the proportion of outliers was almost exactly the same with both attribute combinations.

5. CONCLUSIONS

The objective of this paper was to test if data mining –based outlier detection algorithms could be used to identify coarse assessment errors from a typical forest planning data set. The errors were classified into magnitude classes in order to examine how the magnitude of the error affected the probability of a stand to be classified as an outlier. The probability for a stand to be an outlier was notably higher when at least one of the attributes had a coarse estimation error in class IV. This suggests that coarse individual estimation errors will be likely to stand out from the rest of the data as outliers.

All of the three algorithms had different parameters that had a strong effect on what proportion of the whole data was classified as an outlier. This is quite obvious as the parameters were simply defining what an outlier is. The higher the proportion of the outliers was, the more stands with no coarse estimation errors were classified as outliers. Also, when only a small proportion of all stands were classified as outliers, then the number of outliers in the stands with coarse estimation errors was also smaller. In this kind of a trade-off situation, it was difficult to find a balance so that the proportion of outliers in the stands with no coarse errors would be minimized and the number of outliers in the stands with coarse errors maximized. Also, ranking the different algorithms in some supremacy order was difficult as the results were very similar. We used only a single data set in this study and it would be valuable to test these methods with different forest stand data sets to find out if the optimal parameter values depend on the data set as suggested by Hautamäki et al. [2005].

Normally, different forest attributes correlate with each other to a certain degree. Tall trees tend to have larger diameter than short trees and in a forest with a lot of tall and broad trees, the total volume of timber is normally high due to so called allometric correlations. Because of these correlations, observations in a forest stand data set tend to bunch together into groups or clusters. When we are measuring forest attributes and we do a severe error in the measurement of one of the attributes, this observation might stand out from the rest of the data if it does not fit these correlations. This would be the case for example when mean diameter and height would be measured correctly and total volume would be coarsely underestimated. These kind of outlying observations can possibly be detected with outlier

detection methods even though the methods are not based on analysing the correlations. However, if we systematically overestimate the values of most of the attributes, the correlation between the different attributes might fit the normal correlation and that observation would not necessarily stand out as an outlier. Because of this, all estimation errors can not be considered to be outliers or vice versa. This has a lot to do with the actual data collection method. When forest planners are estimating the forest properties with a highly subjective method, human errors are possible. Classification errors or data input errors could produce individual coarse errors which could be fairly easy to identify as outliers. As in future the forest data may originate from very many sources, from old data to new measurements, the possibility of independency between errors increases and the usefulness of datamining probably increases.

The size of the data was fairly small compared to many data mining –related studies. Yet, the dimensionality of the data is quite high. The problem was that as precise reference measurements require a lot of manual work and expertise, the costs for acquiring this kind of data are high. However, as this study was the first step in applying data mining methodology in a forestry setting, we decided that this kind of fairly small data set should be sufficient.

In spite of the problems with the tested methods, the outlook for applying data mining –based methods for identifying coarse estimation errors in forest planning data is promising. If the forest planning system could identify the outliers, which might be due to coarse estimation errors, this information could be used to provide some estimate about the likely quality of the results. It is possible that the detected possible errors could be used for instance as a basis for planning field checking [Store 2007], for detecting illogical treatment recommendations made in field or even for imputing more logical data to replace the most susceptible observations. However, this is all left to future work.

The next step in studying data mining –based outlier detection methods for identifying coarse estimation errors in forest planning data should be testing of different types of approaches, such as density –based outlier detection. Also, the algorithms should be tested with larger data sets. Another interesting research question would be to study what are the optimal parameter values for the algorithms and how they depend on the particular data set.

ACKNOWLEDGEMENTS

The authors would like to thank the Foundation for Research of Natural Resources in Finland for funding this research project. Initial steps in this research topic were taken as a part of SIMO (SIMulation and Optimization for next generation forest planning) project (http://www.simo-project.org)

REFERENCES

Barnett, V. & Lewis, T. Outliers in statistical data. John Wiley & Sons, Chichester, 3rd edition, 584 pp. 1994.

Bay, S.D. & Schwabacher, M. Mining distance-based outliers in near linear time with randomization and a simple pruning rule. SIGKDD ’03, August 2427, Washington, DC, USA. 2003.

Brecheisen S., Kriegel H.-P. & Pfeifle M. Parallel density-based clustering of complex objects, Proceedings of the 10th Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD'06), Singapore, 2006, in: Lecture Notes in Artificial Intelligence (LNAI), Springer, Vol. 3918, 179-188. 2006.

Breunig, M.M., Kriegel, H-P., Ng, R.T. & Sander, J. LOF: Identifying density-based local outliers. Proc. ACM SIGMOD 2000 Int. Conf. On Management of Data, Dalles, TX, 2000.

Buongiorno, J. and Gilles, J.K. Decision methods for forest resource management. Academic Press. 439 p, 2003.

Burkhart, H. Suggestions for choosing an appropriate level for modeling forest stands. In: Amaro, A., Reed, D. & Soares, P. (eds.) Modelling forest systems. CABI Publishing, 398 p. 2003.

Ester, M., Kriegel, H-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the 2nd Int. Conf. on Knowledge Discovery and Data Mining (KDD'96), Portland, OR, 1996, pp. 226-231. 1996.

Haara, A. Comparing simulation methods for modelling the errors of stand inventory data. *Silva Fennica 37*(4): 477–491. 2003.

Han, J. & Kamber, M. Data mining: concepts and techniques, 2nd ed. Morgan Kaufmann Publishers, 2006.

Hautamäki, V., Cherednichenko, S., Kärkkäinen, I., Kinnunen, T. & Fränti, P. Improving k-means by outlier removal. In H. Kalviainen et al. (Eds.): SCIA 2005, LNCS 3540, pp. 978–987, 2005.

Hawkins, D. Identification of outliers. Chapman and Hall, London. 1980.

Hodge, V. J. & Austin, J. A Survey of outlier detection methodologies. *Artificial Intelligence Review 22*: 85–126, 2004

Jain, A.K., Murty, M.N. & Flynn, P.J. Data clustering: a review. *ACM Computing Surveys 31*(3) (1999) 265–323. 1999.

Johnson, T., Kwok, I. & Ng., R. Fast computation of 2-dimensional depth contours. Proc. KDD, pp. 224-228. 1998.

Kangas, A., Heikkilä, E. & Maltamo, M. Accuracy of partially visually assessed stand characteristics: a case study of Finnish forest inventory by compartments. *Canadian Journal of Forest Research 34*, 916–930, 2004.

Knorr, E.M. & Ng, R.T. Algorithm for mining distance-based outliers in large datasets. VLDB Conference Proceedings, September 1998

Knorr, E.M., Ng, R.T. & Tucakov, V. Distance-based outliers: algorithms and applications. *The VLDB Journal 8,* 237–253, 2000.

Laasasenaho, J., and Päivinen, R. On the checking of inventory by compartments. *Folia Forestalia 664*, 1986. (In Finnish)

Papadimitriou, S., Kitagawa, H., Gibbons, P.B. & Faloutsos. C. LOCI: Fast outlier detection using the local correlation integral. Proceedings of the 19th International Conference on Data Engineering (ICDE), p. 315-326, 2003.

Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J. & Vehmas, M. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sensing of Environment 90*, 319–330, 2004.

Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment 80*, 88– 99, 2002.

Piatetsky-Shapiro, G. & Frawley, W. J. Knowledge discovery in databases. AAAI/MIT Press. 1991.

Poso, S. Basic features of inventory by compartments. *Silva Fennica 17*, 313–349, 1983. (In Finnish)

Ramaswamy, S., Rastogi, S, R. & Shim, K. Efficient Algorithms for mining outliers from large data sets. Proceedings of the ACM SIGMOD Conference on Management of Data, Dallas, TX, 427-438, 2000.

Store, R. Erilaisten tavoitteiden ja maastotyötarpeiden huomioon ottaminen metsäsuunnittelun maastoinventoinnin suunnittelussa. In: Tikkanen, J. & Lappalainen, S. (toim.). Asiakaslähtöisyys metsäsuunnittelun kehittämishaasteena. Metlan työraportteja 65. (http://www.metla.fi/julkaisut/workingpapers/2007/mpw065.htm) 125 s, 2007.

Tang, J., Chen, Z., Fu, A.W. & Cheung, D. A robust outlier detection scheme for large data sets. PAKDD. 2002.

Tang, W. & Khosgotaar, T.M. Noise Identification with the k-means Algorithm. Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'04) - Volume 00, 2004.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Entropy-based fuzzy set optimisation for reducing ecological model complexity

A. M. MOUTON[a], B. De Baets[b], A. Peter[c], G. Holzer[c], R. Müller[c] and P.L.M. Goethals[a]

[a] *Laboratory of Environmental Toxicology and Aquatic Ecology, Ghent University, J. Plateaustraat 22, B-9000 Ghent, Belgium*

[b] *KERMIT: Research Unit Knowledge-based Systems, Ghent University, Coupure links 653, B-9000 Ghent, Belgium*

[c] *Swiss Federal Institute for Environmental Science and Technology (EAWAG), Forschungszentrum für Limnologie, 6047 Kastanienbaum, Switzerland*

Abstract: In recent years, fuzzy logic has been acknowledged as a suitable approach for species distribution modelling due to its transparency and its ability to incorporate the ecological gradient theory. Specifically, the overlapping class boundaries of a fuzzy model are similar to the transitions between different environmental conditions. However, the need for ecological expert knowledge is an important constraint when applying fuzzy species distribution models. Recent research has shown that data-driven fuzzy models may solve this 'knowledge acquisition bottleneck' and this paper is a further contribution. The aim was to reduce the complexity of a data-driven fuzzy habitat suitability model for European grayling (Thymallus thymallus) in the Aare River (Thun, Switzerland). Therefore, we applied an entropy-based fuzzy set selection algorithm, which allowed minimisation of the number of fuzzy sets needed for data-driven fuzzy model development. Comparison of the presented model with a previously developed model revealed that the entropy-based algorithm reduced model complexity substantially without a significant decrease in predictive accuracy. The results of this study could minimise monitoring costs and efforts, and enhance communication between water managers and stakeholders due to increased model transparency.

Keywords: Fuzzy logic, grayling, species distribution modelling, model complexity, data driven.

1. INTRODUCTION

Ecological data often is imprecise and characterized by high uncertainty. Therefore, instead of using crisp values for quantification of variables, many experts apply linguistic descriptions such as "low", "high" or "moderate". Fuzzy systems allow transforming these linguistic descriptions into a mathematical framework in which suitable data processing can be performed [Kampichler, et al., 2000]. The main advantages of fuzzy classifiers are their simplicity and linguistic interpretability, which are important factors for the usability and acceptance of a model [Adriaenssens, et al., 2004, Van Broekhoven, et al., 2006]. This turns fuzzy systems into a popular technique for ecological modelling, resulting in numerous applications. Yet, a purely knowledge-driven approach, aiming at formalizing problem-relevant human expert knowledge, is difficult and tedious. Recent research has shown that complementing fuzzy systems by data-driven techniques can solve this "knowledge acquisition bottleneck" [Žnidaršic, et al., 2006]. For example, the induction of fuzzy rule-based models by heuristic search algorithms is often used in the field of fuzzy rule learning [Hüllermeier, 2005].

To preserve the interpretability of the fuzzy model, however, the number of parameters applied in the model should be limited. Therefore, the aim of this paper was to reduce the complexity of a data-driven fuzzy habitat suitability model for European grayling (Thymallus thymallus) in the Aare River (Thun, Switzerland). We applied and evaluated an entropy-based fuzzy set selection algorithm, which allowed minimisation of the number of fuzzy sets needed for data-driven fuzzy model development. Specifically, the hypothese was tested that the entropy-based algorithm reduced model complexity substantially without a significant decrease in predictive accuracy. The results of this study could minimise monitoring costs and efforts, and enhance communication between water managers and stakeholders due to increased model transparency.

2. MATERIAL AND METHODS

2. 1 Study area

The studied site is a 1300 m stretch of the Aare river in the Bern department, Switzerland, and is situated along the city of Thun. Up to this point, the Aare river is draining an area of about 2490 km² and is classified as a 7th order stream [Strahler, 1957]. The average flow is 111 $m^3.s^{-1}$, with respective base and peak flows of 23 and 570 $m^3.s^{-1}$. The Aare river at the studied site was originally a braided river with large gravel banks. However, since the beginning of the 18th century anthropogenic disturbances were introduced for flood control and hydropower generation [EAWAG, 2002]. Hence, the flow regime is altered and controlled by flood control weirs. Nevertheless, the studied site contains some of the major spawning habitats for European grayling in Switzerland.

To allow for the development of the habitat suitability model, an intensive monitoring campaign was set up. In the studied stretch, 50 cross-sections were defined and water depth was measured along each cross-section at equal distances of about 1 m using a Raytheon 760 depth measuring device (Raytheon, MA, USA). Flow velocity was measured with a Flo-Mate 2000 flow meter (Marsh-McBirney Inc., MD, USA) at 40 % of the water column height in 14 of these cross-sections at equal distances of about 25 m, resulting in 63 measurements. The substrate composition was assessed by underwater photography and visual assessment with DIN A4 frames. Hence, the dominating substrate of the different patches in the studied stretch could be defined. If a patch was covered by macrophytes, both substrate percentages were set to 99 %. This substrate combination can not be observed in the river stretch and hence the definition of this specific situation will not affect optimisation results. All data were collected at a flow of ca. 100 $m^3.s^{-1}$ and no significant flow changes were observed during the measurements.

A finite element grid of the studied stretch with 5625 elements and 22500 nodes was generated using SMS (surface water modelling system, Brigham Young University) software, while the size of the grid cells was adjusted depending on river geometry. Flow velocity and depth values were calculated at each node by a 2-dimensional hydraulic model which was generated using FESWMS (Finite Element Surface Water Modelling System, U.S. Geological Survey). Additional measurements of depth and flow velocity were performed in the whole stretch and more specific in the spawning areas to validate the hydraulic model. The hydraulic modelling was conducted by Schneider & Jorde Ecological Engineering in cooperation with the Swiss Federal Institute of Aquatic Science and Technology (EAWAG) and is described in detail in [EAWAG, 2002].

European grayling spawns in faster flowing patches (0.1 – 0.4 $m.s^{-1}$) with fine to medium-sized gravel substrate. During egg deposition, the trembling female grayling is pushing its abdomen in the gravel substrate, hereby creating small grooves [Fabricus and Gustafson, 1955]. These light-coloured grooves can easily be distinguished from the substrate which is mostly covered with dark brown algae. Hence, the spawning grounds of grayling were visually identified and localised using GPS (Garmin 12X). Each grid cell in the studied stretch was defined as suitable for spawning or not by combining the results of the hydraulic simulations with the observations of the spawning grounds. The resulting dataset

contained 22510 grid cells, 1 output variable indicating whether the habitat was suitable for spawning or not and 4 input variables characterising each grid cell (Table 1).

2.2. Fuzzy set parameter optimisation

The values assigned to the four input variables depth, flow velocity, percentage of fine gravel and percentage of medium-sized gravel, were defined by fuzzy sets [Zadeh, 1965] and not by conventional sets with crisp boundaries (hereafter called crisp sets). When using these crisp sets, for instance depths below 1.5 m would be considered 'low', depths between 1.5 and 3 m 'moderate' and depths higher than 3 m 'high'.

A given depth either belongs to a set (it has a membership degree of 1 to this set) or it does not. A fuzzy set is described by its membership function, indicating the membership degree for each variable value to this set. As the boundaries between these sets are overlapping, an element can partially belong to a fuzzy set and thus have a membership degree to this set ranging from zero to one. Hence, the linguistic statement 'the depth is quite low but tending to be moderate' can be translated into a depth which has a membership degree of 0.4 to the 'low' fuzzy set and of 0.6 to the 'moderate' set. In this study, all membership functions had trapezoidal shapes and were defined by four parameters (a_m, b_m, c_m and d_m): the membership degree linearly increases between a_m and b_m from 0 to 1, is equal to 1 between b_m and c_m and linearly decreases from 1 to 0 between c_m and d_m. A triangular membership function is obtained when b_m equals c_m.

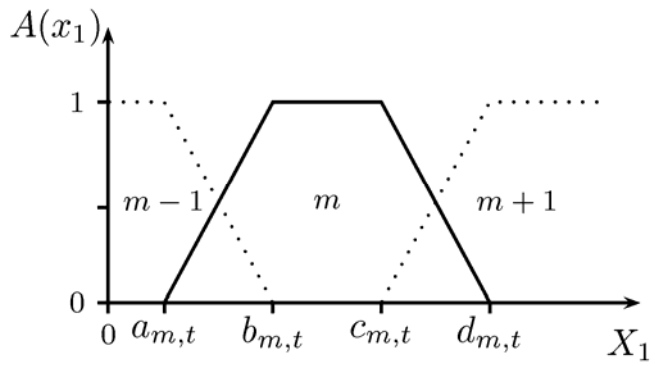


Fig. 1. The parameters of the fuzzy set m of variable X_1.

The parameters of the membership functions corresponding to the fuzzy sets of the input variables have often been derived from an expert knowledge. However, if a fuzzy set of an input variable contains very few training instances, rules which apply to this fuzzy set will be trained inadequately. Therefore, a uniform distribution of the input variables over the fuzzy sets was needed to generate reliable rule bases. The Shannon–Weaver entropy [Shannon and Weaver, 1963] indicated this uniformity and was applied to optimise the parameters of the membership functions of the input variables. The fuzzy sets were converted into crisp ones whose boundaries were the points having a membership degree of 0.5 to the corresponding fuzzy set. The entropy E is given by (convention $0 \log_2 0 = 0$):

$$E = -\frac{1}{\log_2 n} \sum_{i=1}^{n} p_i \log_2 p_i \qquad \textbf{Eq. 1}$$

where n is the number of classes and p_i is the proportion of data points belonging to class i. The algorithm for parameter optimisation starts with n equal to 2, and then the parameters of the fuzzy sets are adjusted in steps of $\frac{r}{2s}$ with r the range of the variable of which the fuzzy sets are optimised, and s the fixed stepsize. For each variable, parameter optimisation started by creating two crisp sets with boundary at r/n with n the number of sets, which is equal to two. The crisp sets were transformed into fuzzy sets by setting the parameters of each set m, $a_{m,t}$, $b_{m,t}$, $c_{m,t}$ and $d_{m,t}$ (Fig. 1) as follows:

$$c_{m,t} = a_{m+1,t} = (m+1) \cdot \frac{r}{n} - t \cdot \frac{r}{2s} \quad , \qquad \textbf{Eq. 2}$$

$$d_{m,t} = b_{m+1,t} = (m+1) \cdot \frac{r}{n} + t \cdot \frac{r}{2s} \quad , \qquad \textbf{Eq. 3}$$

with t indicating the iteration of the sets optimisation and thus $t = 1$ in this situation. Next, the entropy $E_{n,t}$ of this fuzzy set configuration was calculated as described in Eq. 1 and this entropy was set to E_{best}. This entropy E_{best} is compared with the entropy fixed threshold E_{thres} and the algorithm was terminated if $E_{best} > E_{thres}$.

If $E_{best} < E_{thres}$, the algorithm searched for the fuzzy set of which an expansion of the set boundaries could lead to the greatest increase in entropy. For each set m, the upper boundary was expanded as follows:

$$c_{m,t+1} = a_{m+1,t+1} = (m+1) \cdot \frac{r}{n} - (t+1) \cdot \frac{r}{2s} \quad , \qquad \textbf{Eq. 4}$$

$$d_{m,t+1} = b_{m+1,t+1} = (m+1) \cdot \frac{r}{n} + (t+1) \cdot \frac{r}{2s} \quad . \qquad \textbf{Eq. 5}$$

The entropy of this new fuzzy set configuration in iteration $t+1$, $E_{n,t+1}$ was calculated and compared to E_{best}. If $E_{n,t+1} > E_{best}$, the algorithm continued with this new configuration and the upper boundary of the set m was further adjusted according to Eq. 4 and Eq. 5. If $E_{n,t+1} < E_{best}$, the last adjustment of the fuzzy set was aborted and the algorithm continued with adjusting the boundaries of the next fuzzy set.

To avoid that the algorithm would generate erroneous fuzzy sets in which the total membership degree of a variable value would exceed one, the boundary adjustment of the fuzzy sets in each iteration t was limited as follows:

$$(t+1) \cdot \frac{r}{2s} \leq \min(\min(c_{m,1} - b_{m,1}, c_{m+1,1} - b_{m+1,1}), \min(c_{m,t} - b_{m,t}, c_{m+1,t} - b_{m+1,t})) \qquad \textbf{Eq. 6}$$

If the boundary adjustment of fuzzy set m reached this limit and the entropy was still lower than E_{thres}, the set was split into two symmetric sets. This split was created in the same way as the first split of the total variable range into two fuzzy sets, but the variable range was now replaced by the range of the fuzzy set m. The entropy of this new set configuration, $E_{n+1,1}$, was calculated and compared to E_{best}.. If $E_{n+1,1} < E_{best}$, the algorithm did not split the sets and continued with the next fuzzy set. If $E_{n+1,1} > E_{best}$, the algorithm restarted from this new configuration. If an entropy which exceeded E_{thres} was obtained during sets optimisation, the boundaries of the remaining sets were still optimised, but without splitting the sets.

The applied fuzzy sets optimisation method ensures that the distribution of the training data instances over the fuzzy sets is optimal. In a situation with n fuzzy sets, sets which contain less than $100/n$ % of the data will be expanded as far as possible, whereas sets which contain more than $100/n$ % of the data will be reduced to or split into smaller sets. This method avoids that empty or poorly represented sets are included in the model and increases model efficiency by deleting redundant sets. However, a more uniform distribution of the input data over the fuzzy sets does not guarantee that each fuzzy rule is represented uniformly in the input data. Water depth values, for instance, can be uniformly distributed over a ‘low’ and a ‘high’ fuzzy set, but this does not imply that many sampling sites with a high flow velocity and a high depth are represented in the training data set.

In this paper, the impact of the fuzzy sets on the model training result was analysed by comparing two training scenarios. In the first scenario, models were trained based on fuzzy sets which were derived from expert knowledge and were described in an ecological study of spawning grayling in the Aare river [EAWAG, 2002]. In the second scenario, models were trained based on the fuzzy sets which were derived from the fuzzy set parameter optimisation approach described in this paper.

2.3. Fuzzy rule-based modelling and rule base optimisation

The fuzzy rule base combines the input variables into the habitat suitability for spawning grayling and consisted of if-then rules, such as ‘IF depth IS moderate AND flow velocity IS high AND percentage of fine gravel IS high AND percentage of medium-sized gravel IS moderate THEN habitat IS suitable’. The if-part of the rule, the antecedent, describes in which situation this rule applies, while the then-part, the consequent, indicates whether the

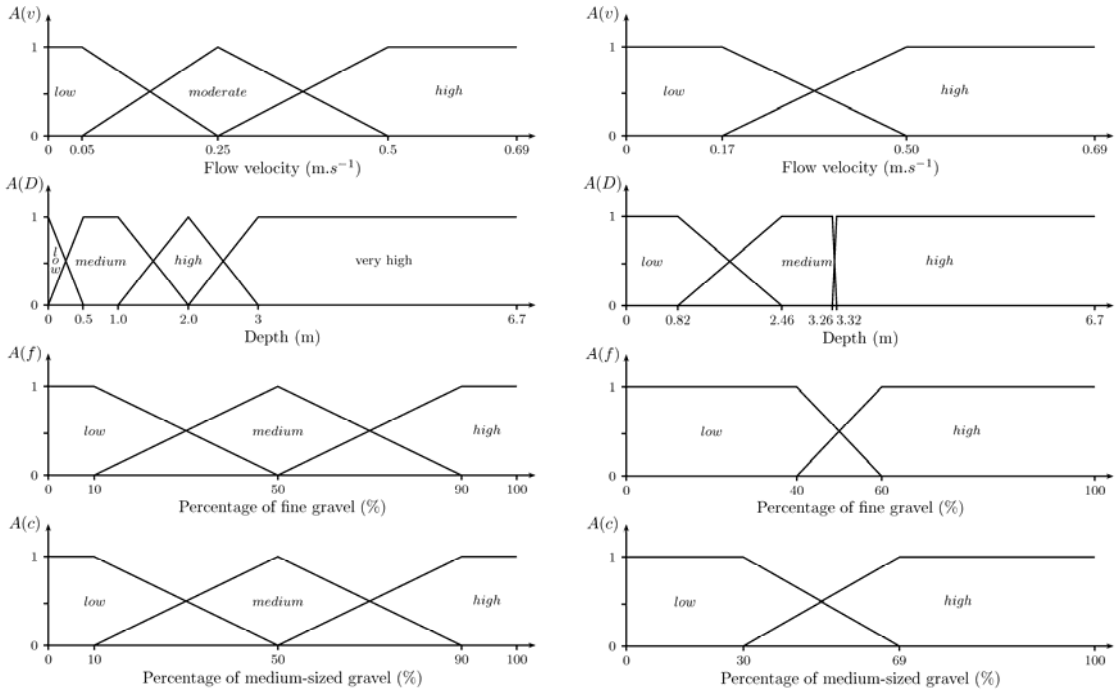
habitat in this situation is suitable or not for spawning grayling. Given crisp values of the four input variables, the output of the fuzzy model is calculated as described by Van Broekhoven et al. [2006]. For each instance, its membership degrees to the membership functions (Fig. 1) of each input variable is calculated. The degree of fulfilment of each rule is then calculated as the minimum of the membership degrees in its antecedent. Finally, to each linguistic output value a fulfilment degree is assigned equal to the maximum of the fulfilment degrees of all rules with the output value under consideration in their consequent. The approach is similar to the Mamdani-Assilian procedure [Assilian, 1974, Mamdani, 1974] in which the fuzzy output is defuzzified in a crisp one based on the fuzzy sets of the output variables. However, the output variable in this work was already defined by two crisp sets: present and absent. Therefore, a different type of model was applied: a fuzzy classifier. The model output was assigned to the fuzzy set with the highest fulfilment degree, which allowed comparison of the modelled output with the observed output and calculation of performance measures. If the output variable consists of two sets, this approach is very similar to the defuzzification procedures used in Mamdani-Assilian models.

To generate a reliable habitat suitability model, the consequents of the fuzzy rules were optimised using a nearest ascent hill-climbing algorithm. Starting from fixed fuzzy sets and a randomly selected rule base, the consequent of one rule is changed into its neighbouring fuzzy set and the impact on model performance is calculated. If model performance increases, the algorithm continues with the adjusted rule base, if not, it continues with the original one. Ten-fold cross-validation was applied to indicate the robustness of the optimisation results. The folds were constructed by randomising the original data set and assigning each data point to one fold without replacement. The species prevalence (i.e. the frequency of occurrence) was constant for all ten folds and equal to the prevalence of the original dataset (0.203 = 4579/22510).

Models were trained based on Cohen's Kappa [Cohen, 1960], which is derived from the confusion matrix [Fielding and Bell, 1997] and ranges from -1 to 1. This measure was selected as a training performance criterion because model training based on Kappa showed good results in a previous study [Mouton, et al., accepted]. Each training iteration was stopped when no further increase of the performance measure on the test fold was observed. Each training iteration was repeated and the obtained rule base was compared to each rule base obtained in previous iteration steps. The resulting rule base similarity indicates the percentage of rule consequents that were different between two rule bases. If the rule base with the highest performance on the test fold was obtained 3 times, training continued on another fold. The training algorithm is described into detail in Mouton et al. (accepted).
.

3. RESULTS AND DISCUSSION

The results show that entropy-based fuzzy set optimisation leads to fewer fuzzy sets than fuzzy set construction based on expert knowledge (Fig. 2). Moreover, the entropy of the optimised fuzzy sets is higher than the entropy of the expert knowledge-based sets.



(a) (b)

Fig. 2. The fuzzy sets which were derived from expert knowledge (a) and from entropy-based fuzzy set optimisation (b).

Consequently, entropy-based fuzzy set optimisation leads to sets over which the input data is distributed more uniformly (Table 1). Specifically, the entropy values show that the distribution of the input data over the expert-based fuzzy sets is less uniform than the distribution over the entropy-based sets. Due to this lack of uniformity, the likelihood increases that some rules of the model do not represent an environmental condition which is present in the studied data set or even in the natural environment (Mouton et al.,).

Such "ghost rules" increase model complexity without contributing to the model performance and should thus be avoided. Although there is no universal relation between the number of ghost rules and the uniformity of the input data distribution over the fuzzy sets, the results in this paper show that the fuzzy set optimisation approach leads to fewer ghost rules in the model (Table 2). Table 2 shows that for the expert-knowledge based model, 50 % of the rules are ghost rules which are not represented in the dataset, whereas for the model with the optimised fuzzy sets, 75 % of the rules represents at least 0.5 % of the sampling points. River managers may thus get the false impression that the model with 108 fuzzy rules is more accurate than the model with 24 rules.

Table 1. The entropy of the different input variables of the fuzzy model. This entropy reflects the uniformity of the distribution of the input data over the fuzzy sets which were derived from expert knowledge (EK) and from entropy-based optimisation.

Variables and their units	Entropy	
	EK-based	Optimised
Flow velocity (m.s^{-1})	0.818	0.94
Depth (m)	0.808	1
Percentage of fine gravel (%)	0.822	0.99
Percentage of medium-sized gravel (%)	0.940	1

The fuzzy set optimisation approach also significantly reduces the model complexity. Specifically, the expert knowledge-based model contained 108 different fuzzy sets and thus the rule base of this model consisted of 108 different rules, whereas the rule base of the entropy-based optimised model contained 24 different rules.

Table 2. Distribution of the samples in the data set over the environmental situations considered in the fuzzy habitat suitability models. The model with the expert knowledge (EK)-based sets and the model with the optimised sets considered respectively 108 and 24 different environmental conditions.

Percentage of samples represented by rule	EK-based		Optimised	
	Number of rules	%	Number of rules	%
0 %	54	50	1	4
0 - 0.5 %	22	20	5	21
0.5 - 5 %	26	24	12	50
> 5 %	6	6	6	25

This model simplification significantly reduced the computational time of the consequent rule base optimisation process. However, a more important advantage of this approach is the improvement of model transparency and thus user friendliness. Although a fuzzy model with 108 rules may still be more transparent than some black box data-driven approaches such as Artificial Neural Networks, the high number of different environmental situations which are represented by the fuzzy rules may confuse river managers. Moreover, Table 3 illustrates that the significant reduction of model complexity does not result in a substantial decrease in model performance. Although the Kappa and TSS values decrease slightly, the percentage of correctly classified instances (CCI) is higher for the model with the optimised fuzzy sets.

Table 3. The performance of the fuzzy models obtained after rule base optimisation based on fuzzy sets which were derived from expert knowledge (EK) and from entropy-based fuzzy set optimisation. The model performance was calculated for the complete dataset and is quantified by the percentage of correctly classified instances (CCI), Cohen's Kappa and the true skill statistic (TSS).

Performance criterion	EK-based	Optimised
CCI	0.71	0.79
Kappa	0.34	0.32
TSS	0.33	0.31

Not only do these results show that different performance criteria focus on different aspects of model performance, but they also confirm that reducing the model complexity by fuzzy set optimisation does not necessarily lead to a substantial reduction of model performance. Moreover, the predictions of both models agree for 82 % of the data points. Further analysis of the differences between the predictions of both models revealed that most of the prediction differences occurred at the edge of habitat patches. Specifically, in some cases one model predicted a smaller patch to be suitable for spawning grayling than the other model. This difference in model predictions was caused by the different fuzzy sets and consequently the complex model could predict the suitability of some habitat patches more accurately than the simplified model could. However, analysis of the model performance showed that in general, the predictions of the simplified model were even more accurate than those of the complex model.

Finally, sampling costs and efforts increase substantially if a higher number of different environmental conditions is represented in the model because several authors suggest that the distribution of the input data over the different environmental conditions should be as uniform as possible[Guisan and Zimmermann, 2000, Hirzel and Guisan, 2002].

4. CONCLUSION

In this paper, an entropy-based fuzzy set optimisation algorithm was applied in order to reduce the complexity of a fuzzy rule-based species distribution model. The results showed that model complexity could be reduced significantly without a substantial loss in model

performance. Although more complex models may be appropriate for specific situations, application of more simplified model may reduce sampling costs and efforts. The reduction in complexity also enhances the transparency of the model and may thus be a valuable contribution to integrated river management by improving communication between modellers, river managers and different stakeholders.

ACKNOWLEDGEMENTS

This research was sponsored by the Special Research Fund (BOF) of Ghent University. The habitat demand data for the presented case study were developed by Schneider & Jorde Ecological Engineering and the Swiss Federal Institute for Environmental Science and Technology (EAWAG).

REFERENCES

Adriaenssens, V., B. De Baets, P.L.M. Goethals and N. De Pauw, Fuzzy rule-based models for decision support in ecosystem management, *Science of the Total Environment*, 319, 1-12, 2004.

Assilian, S., Artificial intelligence in the control of real dynamical systems, Ph.D. Thesis, London University, London, 1974.

Cohen, J., A Coefficient of Agreement for Nominal Scales, *Educational and Psychological Measurement*, 20, 37-46, 1960.

EAWAG, Fishereiliches Gutachten über die Aarebaggerung in Thun, Forschungsanstalt des ETH-Bereichs, EAWAG (Fish Federal Institute for Environmental Science and Technology), Kastanienbaum, Switzerland, 2002.

Fabricus, E. and K.J. Gustafson, Observation of the spawning behaviour of the grayling, *Thymallus thymallus*, *Report of the Institute of Freshwater Research, Drottningholm*, 36, 75-103, 1955.

Fielding, A.H. and J.F. Bell, A review of methods for the assessment of prediction errors in conservation presence/absence models, *Environmental Conservation*, 24, 38-49, 1997.

Guisan, A. and N.E. Zimmermann, Predictive habitat distribution models in ecology, *Ecological Modelling*, 135, 147-186, 2000.

Hirzel, A. and A. Guisan, Which is the optimal sampling strategy for habitat suitability modelling, *Ecological Modelling*, 157, 331-341, 2002.

Hüllermeier, E., Fuzzy methods in machine learning and data mining: Status and prospects, *Fuzzy Sets and Systems*, 156, 387-406, 2005.

Kampichler, C., J. Barthel and R. Wieland, Species density of foliage-dwelling spiders in field margins: a simple, fuzzy rule-based model, *Ecological Modelling*, 129, 87-99, 2000.

Mamdani, E., Application of fuzzy algorithms for control of a simple dynamic plant, *Proceedings IEE*, 121, 1585-1588, 1974.

Mouton, A.M., M. Schneider, A. Peter, G. Holzer, R. Müller, P.L.M. Goethals and N. De Pauw, Optimisation of a fuzzy habitat model for spawning European grayling (*Thymallus thymallus* L.) in the Aare river (Thun, Switzerland), *Ecological Modelling*, accepted.

Shannon, C.E. and W. Weaver, Mathematical theory of communication, University of Illinois Press, Urbana, Illinois, 1963.

Strahler, A.N., Quantitative analysis of watershed geomorphology, *Transactions of the American Geophysical Union*, 38, 913-920, 1957.

Van Broekhoven, E., V. Adriaenssens, B. De Baets and P.F.M. Verdonschot, Fuzzy rule-based macroinvertebrate habitat suitability models for running waters, *Ecological Modelling*, 198, 71-84, 2006.

Zadeh, L.A., Fuzzy Sets, *Information and Control*, 8, 338-353, 1965.

Žnidaršic, M., A. Jakulin, S. Džeroski and C. Kampichler, Automatic construction of concept hierarchies: The case of foliage-dwelling spiders, *Ecological Modelling*, 191, 144-158, 2006.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Identifying Biological Echoes in Radar Scans Using Machine Learning

Reginald Mead[a], John Paxton[a] and Richard S. Sojda[b]

[a] Department of Computer Science, Montana State University, Bozeman, MT 59717 USA (mead/paxton@cs.montana.edu)

[b] Northern Rocky Mountain Science Center, U.S.D.I.-Geological Survey, Bozeman, MT 59717-3492 USA (sojda@usgs.gov)

Abstract: Radar ornithology using data from the NEXRAD weather radar system has given scientists new tools for studying bird migration in the United States. Unfortunately, the process of identifying echoes from birds in radar data still largely requires that trained technicians spend hours manually scouring radar scans. This paper provides some background for understanding biological echoes in NEXRAD data and then describes our initial investigations of the use of machine learning techniques to help automate the process of echo classification. Doppler data allows researchers to look at bird migration by examining large clusters of birds that could be observed approaching and descending at stopover points where they would rest until they began the next leg of their journey. One of the NEXRAD system's greatest strengths has also been a significant obstacle for researchers studying bird migration: with 154 radar stations across the United States, each often producing hundreds of volume scans per day, the amount of data to sort through is staggering. The real problem is that classifying birds in radar scans currently requires a skilled technician who has been trained in visually identifying the tell-tale signs that distinguish biological echoes from non-biological echoes. Consequently, the task of plotting a specific migration over any significant amount of space and time quickly becomes a difficult and resource intensive problem. We have begun by using a K-nearest neighbour classifier, a naïve Bayes classifier, and a neural network to classify the echoes. Early validation results using tenfold cross-validation procedures are hopeful and indicate that machine learning techniques could be well suited for this task. Accuracy rates have exceeded 98 percent. Although these early results are encouraging, it is important to keep in mind that each of the training sweeps used in these experiments was selected by an expert because it could be considered a prototypical example of one particular echo type dominating a sweep. Intuitively, these sweeps are the easiest to classify, which may explain the results. The real test will be to apply these methods to more complex data, including ambiguous data as well as mixed sweeps containing both types of echoes. Our next efforts will concentrate on acquiring and experimenting with such data. Our eventual goal is to use machine learning methods to map bird migration pathways.

Keywords: Machine learning, Artificial intelligence, Radar ornithology, NEXRAD, Bird migration

1. Introduction

In the 1940s, researchers discovered that radar could detect the position and movements of birds and other flying creatures such as insects [Gauthreaux and Belser, 2003]. With this discovery, the field of radar ornithology was born. Researchers used this technique to the best of their ability with whatever data was available to them. The real breakthrough came in the 1990s when the United States began replacing its WSR57 weather surveillance radars with WSR-88D Weather Surveillance Radar 1988 Doppler [Diehl and Larkin, 2002]. The WSR-88D or NEXRAD NExt Generation RADar system provided freely available radar data to researchers that covered a broad geospatial range.

By far, one of the most important aspects of the new system was the addition of a Doppler component to supplement the reflectivity data. Doppler data allowed researchers to examine large clusters of migrating birds approaching and descending at stopover points where they would rest until they began the next leg of their journey.

One of the NEXRAD system's greatest strengths has also been a significant obstacle for researchers studying bird migration; with 154 radar stations, each producing hundreds of volume scans per day, the amount of data to process is staggering. The real problem is that classifying birds in radar scans currently requires a skilled technician who has been trained in identifying the tell-tale signs that distinguish biological echoes from non-biological echoes. Consequently, the task of plotting a specific migration over any significant amount of space and time quickly becomes a difficult and resource intensive problem.

This paper describes an approach to this problem that leverages the power of machine learning techniques to automate the process of echo classification and provide a means by which researchers can automatically detect scans of interest. We provide a brief technical background for the problem, describe several of the classifiers investigated, and conclude with some early empirical results and directions for the future.

2. Biological Pattern Recognition

2.1 Data Format

At its most abstract level, NEXRAD data is hierarchically organized into four basic structures: volumes, sweeps, rays, and pulse volumes [Diehl and Larkin, 2002; Klazura and Ima, 1993]. The volume is the highest level structure, representing a snapshot in time of the entire three dimensional space around the radar station. Volumes are composed of sweeps. Sweeps are essentially two dimensional structures containing echoes at the same elevation angle. Sweeps are further subdivided into rays. Like the spokes on a tire, rays extend outward from the radar station and contain all of the pulse volumes that share the same azimuth. This leads to the lowest level of the NEXRAD hierarchy, the pulse volume. A pulse volume is essentially a rectangular volume of space (technically, it is conical in shape) for which the radar produces reflectivity, velocity, and spectrum width values. Figure 1 illustrates the three dimensional organization of these components.

2.2 Data Preprocessing

Before training and classification can take place, six preprocessing steps are performed on the data. Some of these steps are obvious; others not so.

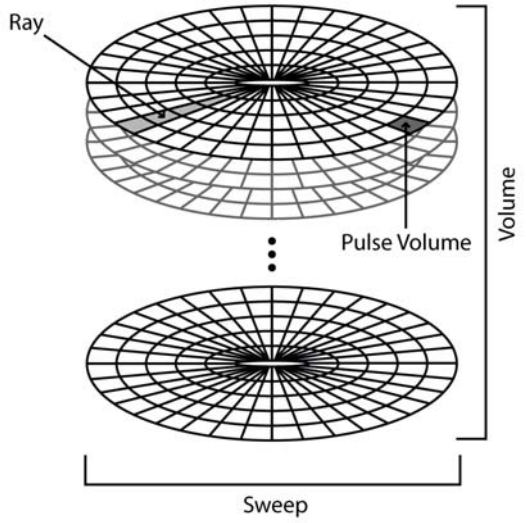


Figure 1. WSR-88D Hierarchical Data Structure

Our first step is to select the sweep or sweeps that we want to use. When studying bird migration, this means using the lowest (.5 degree) elevation sweep. This is the elevation at which most birds will appear.

Our second step is to remove untrustworthy data. This includes both pulse volumes that are very near and pulse volumes that are very far from the radar station. Due to the curvature of the earth and the elevation of the instrument, the radar beam nearest to the ground in the vicinity closest to the radar station. As a result, the echoes nearest the radar

station are also the echoes most likely to be the result of ground clutter. To avoid this problem, we discard the first 20 km of data. At the other extreme is data that is very far from the radar. While reflectivity data can be trusted at distant ranges, velocity data past a certain range cannot be. This is due to what has been coined the *Doppler Dilemma* or the *Range-Velocity Ambiguity* [Doviak and Zrnic, 2006]. The *Doppler Dilemma* states that there is a connection between maximum range and maximum velocity: extending one decreases the other. The balance between range and velocity is determined by the pulse repetition frequency of the radar. To avoid using ambiguous velocity data, we discard pulse volumes farther than 145 km (the unambiguous velocity range) from the radar.

Third, after removing untrustworthy data in the second step, pulse volumes with bad or range-folded reflectivity data are removed. Bad values typically indicate empty space resulting from returned echo strength being less than the signal to noise ratio for the radar.

Fourth, bad or range-folded velocity or spectrum width values are set to zero. Unlike bad valued reflectivity data, these values do not typically represent empty space. Therefore, these values are simply zeroed rather than removed so that the algorithm can exploit the uncorrupted reflectivity information.

Fifth, a set of second order features are calculated for each of the three base values. These second order features include variance, kurtosis (a measure of peakedness), and skewness (a measure of asymmetry) [Joanes and Gill, 1998].

Sixth, insignificant or deleterious features are removed from the data. Features, in this context, include the three base values, azimuth, range, and any second order features calculated in the previous step. A feature is classified as insignificant or deleterious depending on the classifier. In the same way, the rationale for removing a feature varies with the classifier. For classifiers that give more weight to features that have a higher correlation with the classification, removing unimportant features can improve the system's execution time. For a classifier that weights all features equally, removing excess features can have a significant effect on classification accuracy.

2.3 General Methodology

Our classification system starts with supervised training data that is comprised of a number of sweeps that have been selected by an expert and each sweep has been classified as either dominated by biological or non-biological echoes. These sweeps are broken into individual pulse volumes and each pulse volume is assigned the classification of its containing sweep.

For validation, the training data is separated into folds, where a fold consists of data from several sweeps. Due to practical computer processing (hardware) limitations, the sweeps are not used in their entirety. Instead, roughly 5% of each sweep is randomly sampled. Our system performs ten-fold cross validation using this training data [Kohavi, 1995].

3. Machine Learning Approaches

3.1 K Nearest Neighbour Classifier

The K nearest neighbour classifier [Wu et al, 2008] is an instance based learning approach. This type of classifier is known as a "lazy" learner because it delays most of the computational workload until the classification phase. The K nearest neighbour classifier treats each training value as a point in n-dimensional space, where n is equal to the number of features being used. In simple systems, training values are stored and no other processing is performed during the training phase. Later, during the classification phase, a distance function is used to find the K nearest neighbours to an unknown instance and the classes of those neighbours are used to assign a class to the new instance. Figure 2 provides an example of a K nearest neighbour classifier trained on two dimensional data. In this example, the unknown instance represented by the gray circle would be classified as non-biological because two of the three closest training instances are non-biological.

Instance based learning classifiers have been successfully applied to a number of problem domains, but there are consequences that must be considered when using this type of classifier [Aha et al., 1991]. When dealing with radar data, the three characteristics of KNN classifiers with the biggest impact are intolerance to noise, intolerance to irrelevant features, and computational complexity. Fortunately, a number of advances have been made in mitigating these problems.

Intolerance to noisy data can often be assuaged by increasing *k*. The idea is to increase the number of neighbours and thereby decrease the effect of noisy training data. The obvious caveat is that increasing k too much can cause the neighbourhood to cross the concept boundary resulting in reduced performance [Wu et al., 2008].

Simple KNN classifiers suffer when irrelevant, or less important, features are included because they can have a large impact on the distance given by the distance function. A related problem arises when features with large ranges of values begin to dominate the distance function. Both of these issues can be addressed by weighting the features [Wu et al., 2008]. Despite the benefits of weighting, lower priority features can still be a problem because weighting them correctly requires specific knowledge regarding the relative importance of features.

High computational cost is the result of calculating distances to every training instance for each new instance during the classification phase. One improvement that is effective for certain applications is to only save training instances that are misclassified [Aha et al., 1991]. These instances generally define the concept boundary.

For our experiments, we used the IBk classifier provided by the Weka machine learning library [Witten and Frank, 2005]. This KNN implementation was used with the classic Euclidean distance function.

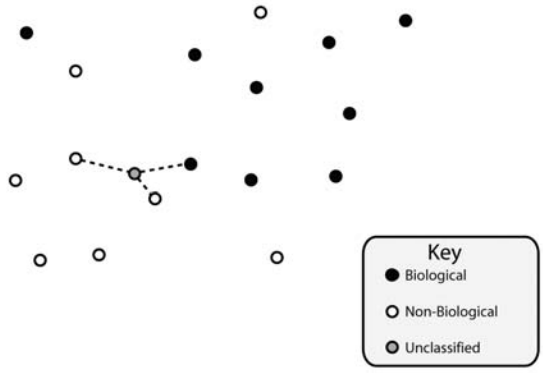


Figure 3. K Nearest Neighbor Classifier Example (K = 3)

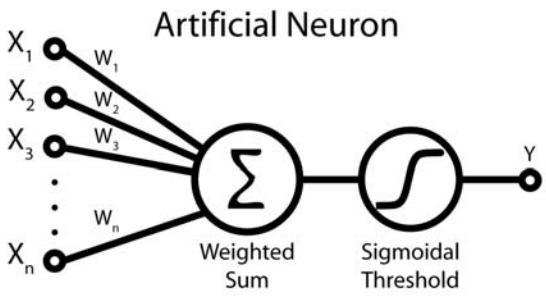


Figure 3. Artificial Neuron

3.2 Naïve Bayes Classifier

The naïve Bayes classifier [Mitchell, 1997] has become popular due to its simplicity, efficiency, and effectiveness. It is often used as a benchmark when experimenting with other, more complex classifiers. This classifier is built based on Bayes' theorem, which states that the probability of seeing event A, given that event B has occurred, can be written as follows:

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)}$$

In our classification problem, Bayes' theorem is used to determine the probability of a pulse volume belonging to the biological class given the pulse volume's features. An unknown pulse volume is classified by maximizing this probability across all the possible classes.

In practice, the theoretical Bayes approach needs an enormous amount of training data and takes an impractical amount of computation time due to the P(B|A) factor in the previous theorem. For our problem, P(B|A) is the probability of seeing the set of features given a specific classification. Calculating the joint probability for a set of features given a classification can be a daunting task due to the number of possible combinations that arise when a nontrivial number of features and classes are being considered. We considered up to 14 features and two classes. Each feature had a wide range of possible values. The combinatorial explosion for this feature set makes calculating the joint probability an intractable problem.

The naïve Bayes classifier avoids this problem by making the profound assumption that attributes are independent of each other. Under this assumption, the following conclusion is reached when calculating the joint probability for features A, B, and C given class D:

$$P(A, B, C \mid D) = P(A \mid D) * P(B \mid D) * P(C \mid D) \tag{2}$$

This product is computed by simply counting the frequencies of the various attributes for each of the candidate classes. Bayes' theorem can now be written as follows:

$$P(D \mid A, B, C) = \frac{P(A \mid D) * P(B \mid D) * P(C \mid D) * P(D)}{P(A, B, C)} \tag{3}$$

Bayes' theorem can be further reduced by eliminating the denominator. This is valid because we are only interested in maximizing the probability across the possible classes. The denominator is a normalization factor and remains constant for each of the various possible values (classes) of *D*. For a set of features A, B, and C, a classification Y can be made by maximizing the probability of the class D given the features.

$$Y = MAX\,(P(D \mid A, B, C)) \tag{4}$$

After applying the independence assumption, (4) can be rewritten as:

$$Y = MAX\,(P(A \mid D) * P(B \mid D) * P(C \mid D) * P(D) \tag{5}$$

The independence assumption is controversial within the scientific community. Although this assumption is invalid in most real world applications, studies have shown that making this assumption is largely inconsequential for many problems. As a result, the naïve Bayes classifier has proven to be remarkably efficient and effective [Mitchell, 1997].

The classifier implementation used in this investigation was the naïve Bayes classifier provided by the Weka 3 machine learning library [Witten and Frank, 2005].

3.3 Neural Network Classifier

Neural networks are a computational framework loosely based on the biological neurons contained in the brain. Basheer and Hajmeer [2000] provides an excellent overview of the fundamentals of neural networks. These networks are composed of a number of simple processing units called neurons. Neurons typically map a number of inputs to a single output. A typical neuron associates a set of weights with the inputs such that when a specific input is presented to the neuron, it multiplies each input by its respective weighting. These weighted inputs are then summed and passed through a threshold function to achieve a final output. Typical output bounds are [-1, 1] and [0, 1]. A common threshold function is the sigmoid function [Mitchell, 1997]:

$$S(x) = \frac{1}{1 + e^{-x}} \tag{6}$$

Figure 3 illustrates a typical artificial neuron. In this case a weight vector **W** is used to map an input vector **X** to an output Y.

The specific mapping, or function, that a neuron executes is determined by its set of weights. Learning algorithms have been developed to learn a mapping by adjusting the weights until a set of inputs properly produces a set of outputs. A single neuron can learn

simple functions. More advanced, often non-linear, functions require a network of neurons working together.

Neural networks typically group neurons into a number of layers, including an input layer, output layer, and any number of hidden layers. An example of a feed-forward neural network can be seen in Figure 4. Feed-forward refers to the unidirectional nature of communication between nodes in this network. A more complex form of neural network is the backpropagation network. This form of network uses the backpropagation learning algorithm to propagate weight changes back through the network [Du, 2006]. Backpropagation networks are one of the most common forms of neural networks used in practical applications. The backpropagation network implemented by the multilayer perceptron classifier in Weka [Witten and Frank, 2005] was used for this investigation.

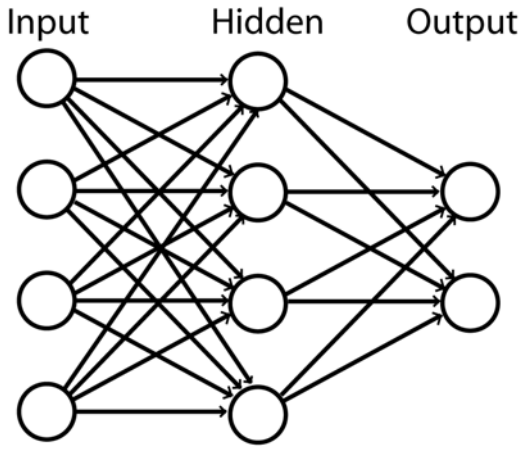


Figure 4. Typical Feed-Forward Neural Network

4. Empirical Validation

Our experiments used ten-fold cross validation [Kohavi, 1995] to evaluate each of the three classifiers. We used 40 sweeps categorized beforehand by radar ornithologists, and each fold consisted of data from four sweeps. Due to computational limitations, only 5% of each sweep was used during training. This resulted in each fold containing approximately 29,000 pulse volumes. The decision to use 5% of the data was based solely on practical considerations. A formal sensitivity analysis examining this decision will be conducted in the future.

After training a classifier on 36 sweeps from nine folds, each of the remaining four sweeps in the tenth fold was categorized by that classifier. Such classification was done at the pulse volume level at this stage. Every pulse volume in the four unclassified sweeps was presented to the classifier as a set of features. Each of the classifiers in the Weka framework produce a classification probability distribution. Although this is natural for naïve Bayes, constructing distributions for the neural network and KNN are not as obvious. The neural network contains output nodes for each of possible classes and each output node produces a numeric score. These scores are then normalized to produce a distribution. In a similar fashion, the class frequencies present in the *k* nearest neighbours are normalized to provide a distribution for KNN. For our problem domain, the classifiers produce a probability distribution over the two possible classes: non-biological and biological echoes. The distribution [.20, .80], for example, would signify that the classifier gives a 20% chance of the pulse volume being non-biological and an 80% chance of being biological. After this distribution was produced, we assigned each pulse volume a class by majority rule.

Next, an entire sweep was assigned a class. Unlike pulse volume classification, sweep classification was not by majority rule. One goal of our work is to reduce the number of sweeps that researchers must consider in their search for biological echoes. As such, eliminating false positives is more important than eliminating false negatives. We found through a series of experiments that using a 70% threshold did not significantly effect classification accuracy, but did substantially bias the algorithm towards producing false negatives rather than the alternative. Therefore, a sweep was only classified as biological if at least 70% of its pulse volumes had been so classified.

Subsequently, each of the remaining nine folds was iteratively processed, resulting in all 40 sweeps being classified once. It is at this stage that we compared the sweep's classification with that provided by the expert. This entire process was repeated for each of the three classification algorithms using the same 40 sweeps.

The following table summarizes the experimental results. The K Nearest Neighbour approach and the Naïve Bayes approach classified sweeps with the same level of accuracy. When time is taken into account, however, the K Nearest Neighbour classifier required nearly 40 times as long to complete.

Table 1. Cross Validation Classification Results

Classifier	Correctly Classified Sweeps	Time to Train and Classify
K Nearest Neighbour	39 / 40	191 min
Naïve Bayes	39 / 40	5 min
Neural Network	40 / 40	17 min

The neural network took three times longer than Naïve Bayes, but made up for the extra time by classifying all 40 sweeps correctly. We used a midrange PC running Ubuntu Linux as our test machine. The test machine contained 1GB of RAM and had a 3.0 GHz Pentium 4 processor.

5. Conclusions and Future Directions

All of the classifiers investigated in this paper performed well on the training data, and the naïve Bayes and neural network classifiers required a fraction of the time used by the K Nearest Neighbour classifier.

These early results are encouraging, but more research is needed. It is important to keep in mind that each of the training sweeps used in these experiments was selected by an expert because it could be considered a prototypical example of one particular echo type dominating a sweep. Intuitively, these sweeps are the easiest to classify, which may explain the results. The real test will be to apply these methods to more complex data. This includes ambiguous data, i.e., data that is not clearly of one type of echo, as well as mixed sweeps containing definitive examples of both types of echoes.

We are currently working with domain experts to acquire training data that covers a broader range of possibilities. Testing the system against more complicated data should bring to light areas in the system that can be improved. If harder data does reduce classification, we will also examine how modifying the basic configuration of some of our classifiers affects accuracy. For example, we will study how varying k, or using a different distance function affects the accuracy of the *k* nearest neighbour classifier.

Our future goal is to develop a system capable of spatially and temporally tracking groups of migrating birds. Such a system could identify migration corridors that are likely to contain migrating birds at certain times of the year as well as day. Applications of this system would include consultation to lessen the environmental impact by development of such things as wind power generation facilities and power lines. In the scientific community, this system would be a tool for scientists researching migration behaviour, allowing them to sift through enormous data sets to study both local bird movements as well as regional migration. Such information can be used by wildlife managers to target their efforts in those areas where birds are either prevalent or are not, depending on their objectives. Likewise, such information can inform development of natural resources to minimize environmental impacts.

To meet this goal, the current system will need to be expanded in a number of ways. Segmentation and boundary detection for classifying groups of echoes at the sub-sweep level will allow us to extract valuable information from sweeps that are not dominated by biological echoes. Another possible improvement would be the inclusion of additional

classes or subclasses (i.e., rain, snow, or dust, rather than simply non-biological). The system might also benefit from clustering algorithms that could group similar echo types in an unsupervised way. Our system will also need the ability to consider inter-sweep relationships between pulse volumes. Knowledge of relationships between sweeps (i.e., elevation and acquisition time) would allow the system to track targets as they move through time and space (even traveling into the range of another radar station). Finally, the weather radar system in the United States is being upgraded to a NEXRADII system which will provide an additional four moments of data as well as increase the precision of the data, itself.

Acknowledgements

We would like to thank the United States Department of Interior's Geological Survey and Fish and Wildlife Service (FWS) for facilitating and funding this research. Specifically, we wish to thank J.B. Bortner and M. Green of the FWS for believing in the early stages of this project and recognizing the possibilities for mapping bird migration. Similarly, R.H. Diehl of the University of Southern Mississippi provided early conceptual ideas and critical insights into the problem domain. Subsequently, he and his students provided the training data that we used.

References

Aha, D., D. Kibler, and M. Albert, Instance-based learning algorithms, *Machine Learning*, 6(1), 37-66, 1991.

Basheer, I., and M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, *Journal of Microbiological Method*, 43(1), 3-31, 2000.

Diehl, R.H., and R.P. Larkin, Introduction to the WSR-88D (NEXRAD) for ornithological research, *Proceedings of the Third International Partners in Flight Conference*, 876-888, Albany, California, 2002.

Doviak, R.J., and D.S. Zrnic, *Doppler radar and weather observations: second edition*, Dover Publications, 562 pp., Mineola, New York, 2006.

Du, K. –L, and M. N. S. Swamy, *Neural Networks in a Softcomputing Framework*, Springer, 566pp., London, 2006.

Gauthreaux, S., Jr., and C. Belser, Radar ornithology and biological conservation, *The Auk*, 120(2), 266-277, 2003.

Joanes, D., and C. Gill, Comparing measures of sample skewness and kurtosis, *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), 183-189, 1998.

Klazura, G.E., and D.A. Ima, A description of the initial set of analysis products available from the NEXRAD WSR-88D system, *Bulletin of the American Meteorological Society*, 74(7), 1293–1311, 1993.

Kohavi, R., A study of cross-validation and bootstrap for accuracy estimation and model selection, *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, 2, 1137-1143, San Mateo, California, 1995.

Mitchell, T.M. *Machine learning*, McGraw-Hill, 414pp., Jurong, Singapore, 1997.

Witten, I. H., and E. Frank, *Data mining: practical machine learning tools and techniques, second edition*, Morgan Kaufmann, San Francisco, California, http://www.cs.waikato.ac.nz/ml/weka/, 2005.

Wu, X. D., V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P.S. Yu, Z. H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, Top 10 algorithms in data mining, *Knowledge and Information Systems*, 14(1), 1-37, 2008.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Statistical analysis of spatial plant patterns under the effect of forest use

I. López[a], T. Standovár[b], J. Garay[b], Z. Varga[c] and M. Gámez[a]

[a] *Department of Statistics and Applied Mathematics, University of Almería. La Cañada de San Urbano 04120 Almería, Spain (milopez@ual.es , mgamez@ual.es)*

[b] *Department of Plant Taxonomy and Ecology, L. Eotvos University, H-1117 Budapest, Pazmany Peter setany 1/c. , Hungary (garayj@ludens.elte.hu , standy@ludens.elte.hu)*

[c] *Institute of Mathematics and Informatics, Szent István University, Páter K. u. 1., H-2103 Godollo, Hungary. (Varga.Zoltan@gek.szie.hu)*

Abstract: The analysis of the consequences of land use (in particular forest use) may be considered as a partial step towards an integrated modelling of a land system. Let us consider a forest territory, where a gap-cut is made, and after a given time period the eventual change in the spatial distribution of undergrowth plants and tree seedlings is to be detected (see Mihók et al., 2005 and Gálhidy et al., 2006). If floristic data are collected along a line transect, we can try to detect the change in the plant distributions along the transect, the so-called *change-point*, and see whether this occurs at the geometric frontier of the human intervention.
The problem, on a theoretical level, can be addressed using the methodology of *change-point analysis* which is a technically involved branch of mathematical statistics (see e.g. Brodsky and Darkhovsky, 1993; Csörgö and Horváth, 1997), widely used to explore the possible temporal or spatial structure of local homogeneity from collected data. (The main application fields of change-point analysis include meteorology, hydrology, or environmental studies, economy, quality control in industry, biology and medicine.) In this paper we propose a practical, operative approach, using only technique of classical statistics.
In our case, given a plant species, along a line transect quadrats have been located and in each quadrat the individuals have been counted. We consider these data as samples of two distributions of the same type but with different parameters, separated by a change-point *K*. Based on the maximum likelihood approach, an algorithm is given to estimate *K*.

Since the distribution of the change-point estimate is not known, as a substitute of its confidence interval, the so-called *change-interval* will be calculated, using an adaptation of the *bootstrap* method. (For this widely applied simulation method see e.g. Efron and Tibshirani, 1993, a justification of the use of bootstrap in this case can be found in Ferger, 1993.) The implementation of the above algorithms was realized with the application of the statistical software "R". As an illustration, for a concrete plant species, the maximum likelihood estimation of the change-point and the calculation of the above mentioned change-interval will be presented.

Keywords: forest use, plant patches, change-point

1. INTRODUCTION

The presence of spatial patterns is commonly detected in vegetation studies. Differences in site conditions, outcome of interspecific interaction, growth characteristics of dominant species (e.g. clonality) can all play their roles in the development of these patterns depending on the spatial scale considered. Deliberate management activities can also change these vegetation patterns. There is a huge variety of methods that have been developed for pattern detection, ranging from analyzing spatial point patterns of individuals of a single species, to applying multivariate techniques like numerical classification (cluster

analysis) to distinguish compositionally more or less homogeneous vegetation patches (e.g. Grieg-Smith [1983], Jongman et al. [1987], Podani [2000]).

Here we apply a new approach for detecting spatial patterns of plant distribution.

2. STATISTICAL ANALYSIS OF HUMAN EFFECTS ON A FOREST

In forests, human management is aimed at only a few – though dominant – components (trees, wild game species) of the whole ecosystem. However, nowadays an increasing attention is paid to the loss of biodiversity. As a consequence there is a need to assess the effects of different land use activities – in our case forest management – on original biodiversity. Experience shows that not all plant species show a clear, easily detectable reaction to management activities, even if they create steep gradients (like the opening of a small canopy gap in an old forest). Often one can only find difference in the distribution of a species among patches of different quality. Consequently, new methods – capable of detecting such minor changes – could be used to detect minor, not readily detectable causes of human management.

There is a very wide range of applications of statistical methods of change-point analysis in ecology (see Becker et al. [2007], Schleip et al. [2006], Reed [2000] and their references). Our method not only provides an estimation for the location of a change-point and the different distributions laying in different patches, but also provides a so-called *change-interval* (C.I. for brevity) which localizes the distribution change with a high probability level. This interval can be considered as an estimate of a transient zone between patches. The latter has a particular importance in plant ecology since the change between patches usually is not point-like. In a transient zone there may be a mix of two patches, or a special plant composition. Our aim is to study case when the transient zone is small and contains only a mix of the patches.

3. THE FIELD EXPERIMENT

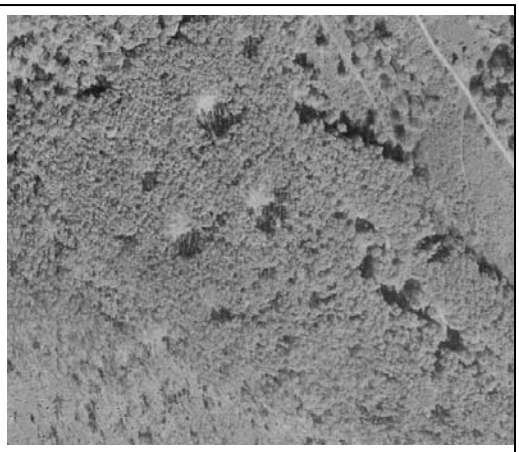


Figure 1. Aerial photograph showing the 3 large and 5 small artificial gaps.

To test the applicability of our method, we used data that were collected in the framework of a study aimed at investigating the effects of canopy gap size on the resulting spatial distributions of key abiotic environ-mental variables (light and soil moisture) in gaps, and at studying how light and soil moisture affect the abundance and distribution of herb layer species. The study site is located in the Börzsöny Mountains, northern Hungary (47.9º N, 18.9º E). Mean annual temperature is 8 ºC, mean monthly temperature is -3.5 ºC and 18 ºC in January and in July, respectively. Annual precipitation is 700–800 mm. Bedrock is andesite, on which medium deep brown forest soil has developed. The study area is located at 540–610 m elevation, on a relatively steep east-northeast facing slope, that is covered by an almost pure stand of European beech (*Fagus sylvatica* L.). Average tree height is 25 m, mean diameter at breast height is 30 cm. Detailed site description is given in Gálhidy et al. (2006).

The selected stand was a good representative of even-aged, mature (86 year old), and dense forests – typical products of the common contemporary silvicultural system (uniform shelter-wood) [Matthews 1991]. Because of the dense tree canopy, understorey vegeta-tion was extremely sparse before opening the experi-mental gaps. Three large gaps (the proportion of tree height of surrounding stand (H) to gap diameter (D) was 1:1.5) and five small gaps (H:D was 1:0.5) were created in February 2001 (Figure 1). We used a systematic sampling design with 5-meter grid resolution and 1x1 m quadrats. As Figure 2 shows, each large gap contained 123 quadrats, whereas small gaps contained 64 quadrats each. Vegetation data were recorded on seven occasions (in September 2000 – before gap creation – May/ September 2001, and May/September 2002, August 2004 and 2006). On each occasion we determined the cover of each herbaceous species using visual estimation in each quadrat. Among other environmental variable, light conditions were studied in each quadrat, so we could reliably decide if a quadrat was in a gap or non-gap environment.



Figure 2. The photograph take in a large gap in late October 2007 (© László Gálhidy)

For the present study we used the data of one species, bramble (*Rubus fruticosus* L.) collected in one of the large gaps in 2006, containing 25 quadrats, i.e. in the seventh growing season after the artificial gaps had been opened, represented in Figures 1 and 2.

4. MODEL DESCRIPTION, ALGORITHMS AND RESULTS

4.1 Model Description

We fix $0<<K<<N$ and suppose that the number of plants in quadrats 1, 2, 3,…, K are independent random variables with the same discrete probability distribution

$$\xi:\begin{cases} 0 & 1 & 2 & .... & k \\ p_0 & p_1 & p_2 & ... & p_k \end{cases}$$

whereas the number of plants in quadrats K+1, K+2, K+3, …, N are independent random variables with the same discrete probability distribution

$$\eta:\begin{cases} 0 & 1 & 2 & ... & l \\ q_0 & q_1 & q_2 & ... & q_l \end{cases}$$

1	2	…	K- 1	K	K +1	K +2	…	N
ξ	ξ	…	ξ	ξ	η	η	…	η

First, from a given sample vector $X:=(x_1, x_2,..., x_N)$, for each possible K, we estimate distributions of ξ and η, and the probability of "realization" of the given sample. Then, from the possible values of K we obtain the required estimate for K, applying the maximum likelihood approach.

4.2 Estimation of distributions ξ and η

For given $1 \leq K \leq N$-1, a possibility to estimate ξ in terms of relative frequencies may be the following: Let

$$k = \max_{j=1,\ldots,K} x_j ,$$

and for each $i = 0,1,\ldots,k$, we define the probability that the variable ξ takes each of its possible values:

$$\hat{p}_{K,i} = P(\xi = i) = \frac{\text{number of indices } j = 1,2,\ldots,K \text{ with } x_j = i}{K},$$

providing an estimate for the distribution of ξ.

In analogous way we estimate the probability distribution of η : let

$$l = \max_{r=K+1,K+2,\ldots,N-1} x_r ,$$

and for each $i = 0,1,\ldots,l$, we define

$$\hat{q}_{K,i} = P(\eta = i) = \frac{\text{number of indices } r = K+1,\ldots,N \text{ with } x_r = i}{N-K}.$$

Let P_K be the probability of "realization" of the sample X:=(x_1,x_2,...,x_N), calculated with the above estimated probabilities:

$$P_K := \left(\prod_{i=1}^{K} \hat{p}_{K,x_i} \right) \left(\prod_{s=K+1}^{N} \hat{q}_{K,x_s} \right),$$

considered the "goodness" of K. Based on the given sample X, our purpose is to find a K which maximizes P_K, providing the "best" (i.e. the "most probable") value of K . We shall deal with this in the next section.

4.3 Algorithms

Algorithm 1 (Estimation of the change-point *K*):

1. Introduce sample X. N:= Size (X).

2. FOR K=1 until N-1:

a) Calculate: $\hat{p}_{K,i}$ and $\hat{q}_{K,i}$, for each i .

b) Calculate: $Log\ P_K = \sum_{i=1}^{K} Log\ \hat{p}_{K,x_i} + \sum_{s=K+1}^{N} Log\ \hat{q}_{K,x_s}$.

(Logarithm is introduced to avoid too small probability values.)

3. *LogProbSample:= ($LogP_1$,…,$LogP_{N-1}$).*

4. *EstimateK:=* Position K with maximum value among the coordinates of *LogProbSample.*

5. Return *EstimateK.*

To find a change-interval for K we elaborate a resampling method based on the known Bootstrap, but with certain modifications in the choice of the elements of each sample of the simulations, in order to fit the method to our problem. The original sample is divided in two homogenous parts, such that the order of the elements of the new samples is important, since, by the linear arrangement, we must not mix *all elements* in a random way. The

generated samples must keep the particularity of having two homogenous parts. The process to follow is explained below:

Algorithm 2 (Calculation of a 90% level Change-Interval):

1. Introduce the sample X:=(x_1,x_2,...,x_N). N:= Size (X).
2. FOR K=1 until N-1

a) Calculate a weight for each K:

$$W_K = \left(\prod_{i=1}^{K} 10\, \hat{p}_{K,x_i} \right) \left(\prod_{s=K+1}^{N} 10\, \hat{q}_{K,x_s} \right)$$

b) Normalize the weights (and denote them by WN_K).

c) FOR L=1 until m (we generate m samples for each K):

c1) Generate K random numbers $\{u_{1,\ldots,}u_K\}$ of a discrete uniform distribution $U[1, K]$, and N-K random numbers $\{u_{K+1,\ldots,}u_N\}$ of distribution $U[K+1, N]$.

c2) We generate each sample with two homogenous zones, selecting the elements of the original sample according to the random positions obtained in c1) for both zones: X_L:=(x_{u_1}, …,x_{u_K}, $x_{u_{K+1}}$,..., x_{u_N}).

c3) We apply Algorithm 1 to the sample X_L, to obtain an estimate K_L for the change-point.

d) From the obtained values K_1,…, K_m, (m large enough) we calculate a distribution d_k of the change-point, for each fixed K.

3. We combine all these new distributions to obtain a unique distribution $\sum_K WN_K d_K$ for K, for which we calculate the 90%-level change-interval, with percentile 5 and percentile 95 as extremes.

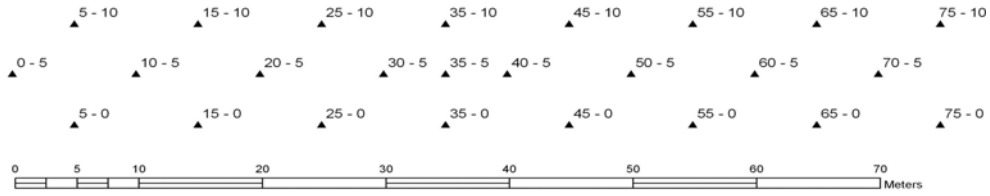
If we have a small amount of data, we can increase the number of data in the following way: we uniformly divide each quadrat of the linear transect into 100 small quadrats along a straight line. In each small quadrat the species will be present (value 1) or not (value 0). Let w be the number of small quadrats where the species is present. Then, using the statistical software "R" we generate randomly w values of a discrete uniform distribution between 1 and 100. These w values will indicate the positions of the small quadrats with one plant, and the remaining 100-w will be the small quadrats with no plant. Therefore, in these 100 small quadrats we represent a w % presence of the species. We denote this new data vector by S. Now, however, we may have too many data for a reasonable run time for the calculations. To reduce them, we sum the values of each 10 consecutive quadrats, and denote this new data vector by Z. We carry out this in the following:

Algorithm 3:

1. Introduce the original sample $X = (x_1, x_2, ..., x_N)$. $N :=$ Size(X).
2. We increase its size to $100N$, then X changes to a vector $S = (s_1, s_2, ..., s_{100 \cdot N})$ of 1s and 0s, and the frequency of 1s, uniformly placed in a random way between the positions $(i-1) \cdot 100 + 1, i \cdot 100$ in S, will be x_i.
3. We sum every 10 values of S obtaining a sample vector $Z = (z_1, z_2, ..., z_{10N})$, to which we apply Algorithm 1 and Algorithm 2 (with $m = 100$), obtaining the estimate for *K* and the 90%-level change-interval.

4.4 The linea

Consider the species *Rubus fruticosus* with data of 2006, taken from the following area, and described by Cartesian coordinates "*X*-*Y*":



Below, in the first three columns we present the data for the species at each location with coordinates *X* and *Y*. We know that the distribution at the centre of this area is different from those observed at the extremes. The change of distribution is observed around X=25 on the left, and around X=55 on the right. In order to estimate this change-point we distribute the data along a straight line in the way shown in the 5th and 6th columns. The new data set is given in the last column:

						Data taken from
X	Y	RUB_FRU_06	K (new order)	X	Y	RUB_FRU
0	5	0,00	1	0	5	0
5	0	0,00	2	75	0	0
5	10	0,00	3	75	10	0
10	5	0,00	4	5	10	0
15	0	0,00	5	5	0	0
15	10	0,00	6	70	5	0
20	5	0,00	7	10	5	0
25	0	5,00	8	65	10	0
25	10	25,00	9	65	0	0
30	5	15,00	10	15	0	0
35	0	40,00	11	15	10	0
35	5	35,00	12	60	5	0
35	10	0,00	13	20	5	0
40	5	45,00	14	55	10	0
45	0	5,00	15	55	0	0
45	10	25,00	16	25	0	5
50	5	25,00	17	25	10	25
55	0	0,00	18	50	5	25
55	10	0,00	19	30	5	15
60	5	0,00	20	45	0	5
65	0	0,00	21	45	10	25
65	10	0,00	22	35	10	0
70	5	0,00	23	35	0	40
75	0	0,00	24	35	5	35
75	10	0,00	25	40	5	45

Now, from the data of the last column, we want to detect the change of distribution. i.e. the *X*- coordinate of the change-point. To this end we apply Algorithm 3 to the last data column as "original sample", obtaining a vector *Z* of the following 250 data:

[1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

[39] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

[77] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

[115] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0

[153] 1 0 3 0 0 0 0 0 3 1 2 3 3 1 4 2 2 4 3 3 2 3 2 2 4 3 1 2 1 4 3 3 0 0 1 1 1 1

[191] 0 0 0 0 0 0 1 1 1 2 1 3 2 3 6 2 3 2 0 3 0 0 0 0 0 0 0 0 0 0 6 3 3 3 5 5 4 3

[229] 3 5 3 4 3 3 3 5 2 5 2 5 5 3 7 5 4 4 4 4 5 4

4.5 Results

With these 250 data, algorithms 3, 1 and 2 (with m=100) provide an estimated *K* equal to 150, and 90 % level change interval [149,160]. These results would correspond in the large data (*S*) to 1500 for *K* and [1490,1600] for the C.I., which in terms of the original quadrats would be 15 for *K*, and [15,16] for the C.I., with coordinates 55-0 for the estimated change-point, and 55-0;25-0 for the C.I.

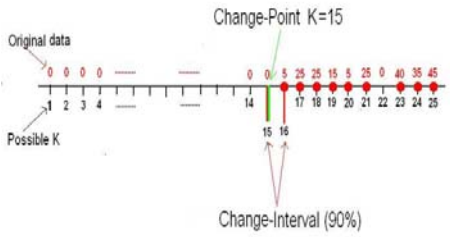


Figure 3: Estimated *K* and C. I., in terms of the original data

5. CONCLUSIONS

Based on the data of a plant species, bramble (*Rubus fruticosus* L.) collected in an experimental forest gap, we have shown how a bootstrap method can be applied for the estimation of the changes in plant densities implied by human intervention. At this initial stage of our study we investigated a relatively small data set concerning a single species, in a real situation there may be about 100 plant species, and different species usually to respond differently to environmental changes.

Once we have estimated where the densities of different plant species change, we will be able to investigate whether these plant species change in the same zone, or as a response to a changed environment, a special "plant community" has been formed. However, this may be the topic of further studies.

ACKNOWLEDGEMENTS

The authors wish to thank the Ministry of Education and Science of Spain for the financial support of the project TIN2007-67418-C03-02, and OTKA K62000, which has partially supported this work. TS and JG are grantees of the János Bolyai Scholarship.

REFERENCES

Becker, B., Lawrence J., Belisle P., Wolfson, D.B., Platt W.J. Bayesian change point analyses in ecology. *New Phytologist*. **174,** 456-467. 2007

Brodsky, B.E. and Darkhovsky, B.S., *Nonparametric Methods in Change-point Problems*. Kluwer Academic Publishers, The Netherlands, 1993.
Csörgö, M. and Horváth, L., *Limit Theorems in Change-point Analysis*. Wiley, Chichester, 1997.
Efron, B. and Tibshirani, R. *An Introduction to the Bootstrap*. Chap- man & Hall, New York, 184-188, 1993.
Ferger, D. Asymptotic distribution theory of change-point estimators and confidence intervals based on bootstrap approximation. *Mathematical Methods of Statistics, Vol. 3, No. 4*, 362-378, 1993.
Gálhidy, L., Mihók, B., Hagyó, A., Rajkai, K. and Standovár, T. 2006. Effects of gap size and associated changes in light and soil moisture on the understorey vegetation of a temperate deciduous forest. *Plant Ecology* 183:133-145.
Greig-Smith P. 1983. *Quantitative Plant Ecology*, 3rd edition. Blackwell, Oxford.
Jongman R. H. G., C. J. F. ter Braak, and O. F. R. van Tongeren. (eds.) 1987. Data Analysis in Community and Landscape Ecology. Pudoc, Wageningen
Matthews, J.D. 1991 *Silvicultural Systems*. Calderon Press, Oxford.
Mihók B., Gálhidy L., Kelemen K., and Standovár T. 2005. Study of Gap-phase Regeneration in a Managed Beech Forest: Relations between Tree Regeneration and Light, Substrate Features and Cover of Ground Vegetation. *Acta Silv. Lign. Hung*., 1:25-38.
Podani, J. 2000. *Introduction to the Exploration of Multivariate Biological Data.* Backhuys Publishers, Leiden.
Reed,W.J reconstruction ther history of forest fire frequency: Identifying hazard rate change points the Baysian information criterion. *Can. Jour. Stat***. 28**, 353-365. 2000
Schleip, C. Menzel, A., Estrella, N., Dose, V. The use of Bayesian analysis to detect recent change in phenological events throughout of years. *Agri. Forest Meteor*. **141,** 179-191 . 2006.

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Multivariate Spatio-Temporal Clustering (MSTC) as a Data Mining Tool for Environmental Applications

Forrest M. Hoffman [a], **William W. Hargrove**[b], **Richard T. Mills**[a], **Salil Mahajan**[c], **David J. Erickson**[a], **and Robert J. Oglesby**[d]

[a]*Oak Ridge National Laboratory (ORNL), Computer Science & Mathematics Division, Oak Ridge, Tennessee, USA (forrest@climatemodeling.org, rmills@climatemodeling.org, ericksondj@ornl.gov)*

[b]*USDA Forest Service, Southern Research Station, Asheville, North Carolina, USA (hnw@geobabble.org)*

[c]*Texas A&M University, College Station, Texas, USA (salilmahajan@neo.tamu.edu)*

[d]*University of Nebraska, Lincoln, Nebraska, USA (roglesby2@unlnotes.unl.edu)*

Abstract: The authors have applied multivariate cluster analysis to a variety of environmental science domains, including ecological regionalization; environmental monitoring network design; analysis of satellite-, airborne-, and ground-based remote sensing, and climate model-model and model-measurement intercomparison. The clustering methodology employs a k-means statistical clustering algorithm that has been implemented in a highly scalable, parallel high performance computing (HPC) application. Because of its efficiency and use of HPC platforms, the clustering code may be applied as a data mining tool to analyze and compare very large data sets of high dimensionality, such as very long or high frequency/resolution time series measurements or model output. The method was originally applied across geographic space and called Multivariate Geographic Clustering (MGC). Now applied across space and through time, the environmental data mining method is called Multivariate Spatio-Temporal Clustering (MSTC). Described here are the clustering algorithm, recent code improvements that significantly reduce the time-to-solution, and a new parallel principal components analysis (PCA) tool that can analyze very large data sets. Finally, a sampling of the authors' applications of MGC and MSTC to problems in the environmental sciences are presented.

Keywords: cluster analysis; parallel computing; geospatial data; ecoregions; general circulation models

1 INTRODUCTION

A multivariate statistical cluster analysis technique based on an iterative k-means algorithm [Hartigan, 1975] has been applied by the authors to analyses in a wide variety of environmental science domains. The clustering methodology was first used by Hargrove and Hoffman to derive ecoregions in an objective and repeatable manner from map stacks of 9 and 25 synoptic geophysical characteristics for the conterminous United States. As shown in Figure 1, the Multivariate Geographic Clustering (MGC) procedure involves a transformation of map cells from geographic space (on the left) to points in the d-dimensional data space (on the right) formed by treating each of the factors (or maps) as an orthogonal axis. Every cell is represented by a single point, p, in data space, and points representing cells with similar characteristics will be near to each other in data space. The clustering task is to group nearby points together in an iterative fashion to create

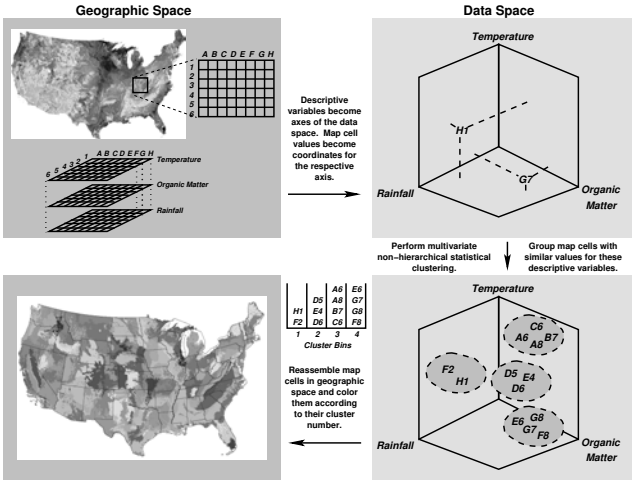


Figure 1: The Multivariate Geographic Clustering (MGC) procedure involves a transformation from geographic space (left) to data space (right), an iterative k-means cluster analysis, and a transformation back to geographic space.

the requested number of clusters, k, the centroids of which represent the mean conditions of the member cells. The cells are then projected back into geographic space and on a map are colored by their cluster assignment. Multivariate Spatio-Temporal Clustering (MSTC) employs the same approach except that multiple maps through time are combined, either as additional factors or axes in data space or as additional points in the same data space. That choice depends on whether the analysis is designed to group together points with similar histories or futures or to group points with similar conditions no matter when they occur to elucidate trajectories among clusters or states in the data or state space.

In order to cluster large ecological, climate, and remote sensing data sets, an efficient and scalable k-means algorithm that can make use of distributed memory parallel computers was created. Originally developed and implemented on a 128-node Beowulf-style parallel computer, called the Stone SouperComputer and constructed from surplus commodity desktop PCs [Hargrove et al., 2001], the high-performance parallel clustering algorithm [Hoffman and Hargrove, 1999] scales to thousands of processors on some of the largest supercomputers in the world. Described below are the clustering algorithm, recent code improvements that significantly reduce the time-to-solution, and a new parallel principal components analysis (PCA) tool that can analyze much larger data sets than any commercial package. Additionally, a sampling of the authors' applications of MGC and MSTC to problems in the environmental sciences are presented.

2 Cluster Analysis Algorithm

2.1 Overview

The clustering method consists of two parts: initial centroid determination, called seed finding, and iterative assignment of points to centroids until convergence is reached. Initial centroids are ordinarily established by sequentially examining each point in the data set and retaining a list of the k most widely separated points in data space, which become the seeds. This inherently serial process is difficult to parallelize; however, the data set can be divided equally among the m compute processes, each of which finds the best k candidate centroids, then a single node can find the "best of the best" k centroids from the $k \times m$ candidates. This parallel seed-finding scheme produces seeds that are not as widely dispersed as those produced by the serial scheme, but because the iterative portion of the clustering method runs quickly in parallel, starting with lower quality seeds that may require a few additional iterations does not significantly increase the time to solution. When the number of processes, m, is large, the cost (in time) for computing the

"best of the best" seeds on a single process can be large. As a result, an alternative scheme was implemented in which the upper half of the active processes sends its k candidate seeds to the lower half of the active processes, the upper half become inactive, and the lower half find the best k candidate seeds from the $2k$ each now possesses. This "folding" procedure repeats until only a single process is left, and its best k candidates become the seeds.

In the iterative portion of the method, each point is assigned to the cluster centroid to which it is closest, by simple Euclidean distance, in data space. After all the points are assigned to a cluster, new positions are calculated for each centroid as the mean value along every axis of the points assigned to that centroid. This procedure of assigning points to centroids and recomputing the centroid locations repeats until the number of points that change cluster assignment drops below a convergence threshold. Once this threshold is met, the final cluster assignments and centroid locations are saved. This algorithm is implemented with a traditional master/slave parallel architecture using the Message Passing Interface (MPI). The algorithm is nearly perfectly parallelizable and produces the same result whether run in serial or in parallel. In each iteration, the master process distributes the current centroids to the slave processes, assigns blocks of points to slaves for classification (*i.e.*, assignment to nearest centroid), collects those classifications, and recomputes new centroid locations based on cluster membership. This procedure repeats until convergence is attained. The block size is specified by the user and can be set to optimize code performance on various parallel systems with different numbers of processors, amounts of memory, and I/O characteristics.

2.2 Cluster Algorithm Improvements

The need to analyze and compare increasingly large model and observational data sets has motivated a number of recent software engineering efforts aimed at improving the performance of the clustering code. These include 1) reorganization of the parallel clustering code to improve readability and simplify debugging, 2) implementation of an acceleration technique that decreases the number of comparisons that must be performed as progress toward convergence is made, and 3) implementation of a method for intelligently handling centroids that lose all cluster membership in an iteration.

Acceleration Technique. Two modifications to the standard k-means algorithm, described by Phillips [2002a, b], can significantly reduce the time-to-solution without changing the clustering results. The first eliminates unnecessary point-to-centroid distance computations and comparisons based on the previous cluster assignment and the new inter-centroid distances. Ordinarily, each point p is compared against each centroid c_i resulting in $O(nk)$ comparisons and a running time of $O(ndk)$ for each iteration. Phillips [2002a] showed, using the triangle inequality, that if the distance between the former centroid assignment c_i and the next candidate centroid c_j is greater than or equal to twice the distance between the point p and the former centroid c_i, evaluating the distance between the point p and the candidate centroid c_j is unnecessary. The triangle inequality states that $d(c_i, c_j) \leq d(p, c_i) + d(p, c_j)$, where the function $d(a, b)$ gives the Euclidean distance between points a and b. As a result, $d(p, c_j) \geq d(c_i, c_j) - d(p, c_i)$. Therefore, if it is known that $d(c_i, c_j) \geq 2d(p, c_i)$, one can conclude that $d(p, c_j) \geq d(p, c_i)$ without having to calculate $d(p, c_j)$. The candidate centroid, c_j, is eliminated without ever computing its distance to the point p because the former centroid, c_i, is equally close or closer. The second modification of Phillips [2002a] further reduces evaluations by sorting inter-centroid distances so that new candidate centroids c_j are evaluated in order of their distance from the former centroid c_i. Once the critical distance $2d(p, c_i)$ is surpassed, no additional evaluations need to be performed; the nearest centroid is known from a previous evaluation. When the number of points n is much larger than the number of clusters k, the cost of sorting inter-centroid distances is negligible.

In order to implement this acceleration technique in the parallel clustering program, the code was modified so that the master process retains the former cluster assignments into the next iteration and distributes them to the slave processes when they receive a new block of points to assign to cluster centroids. In addition, a parallel sorting scheme was implemented in which, at the end

of each iteration and after the new centroid positions have been computed and distributed, the slave processes divide up the total number of centroids, k, and perform the necessary sorting of inter-centroid distances for their contiguous block of centroids. The k resulting ordered distance vectors are collected and distributed to all slave processes for use in the next iteration.

"Warping" Empty Clusters. Initial cluster centroids are chosen to be a subset of the points to avoid starting off the clustering process with centroids that have no members in the first iteration. This often prevents having centroids without any points assigned to them in all subsequent iterations, but experience has shown that centroids occasionally lose all cluster members in later iterations. We call such orphaned centroids empty clusters. A scheme has been developed to handle empty clusters by moving or "warping" them to the location in data space of the points that are the worst fit to their centroids.

This scheme is implemented by having all slave processes keep track of the farthest (or worst fitting) point for each cluster its subset of points, then these are passed to the master process. The master process keeps only the farthest point for each cluster it receives from all slave processes, called the "worst of the worst." At the end of an iteration, if any empty clusters are detected, the master process will sort the list of centroids by the distance to its farthest point. This list is read from the bottom since an ascending sort is performed, and each empty cluster is "warped" to the location of the point that has the next worst fit, called the "farthest of the far." When this occurs, the point is reassigned to the newly warped centroid and removed from the cluster to which it was assigned in the current iteration. Because some clusters may have only one member point, those points are not candidates for reassigning to empty clusters. In practice such single-member clusters will sort to the top of the list and are unlikely to be considered since the "farthest distance" will be zero for those clusters.

Theoretically, a maximum of $k/2$ empty clusters could be handled in this fashion. In practice, however, the presence of single-member clusters further reduces this number. If the list of candidate farthest points is exhausted, the remaining empty clusters are warped to the origin, $(0, 0, \ldots)$. In any iteration in which empty clusters are detected, the master process sets a "warp flag" that is broadcast to slave processes. This warp flag ensures that convergence is not reached in any iteration containing empty clusters, *i.e.*, at least one more iteration will be performed.

2.3 Parallel PCA Tool

Many data sets requiring analysis via MSTC are of high dimension: those incorporating hyperspectral imagery, for example, may involve on the order of hundreds of descriptive variables. Much of this information is redundant, and it is often possible to effectively describe the variability in the data using a significantly lower-dimensional representation. Working in a space of relatively low-dimension allows the data set to be clustered in less time than would be required using the original, high-dimensional representation, or, alternatively, allows a much larger data set to be clustered in the same amount of time. Additionally, clustering in the reduced space can sometimes lead to higher quality clusters due to noise suppression and other properties of the reduction.

One of the most common methods for unsupervised data reduction is principal components analysis (PCA), originally described by Pearson [1901] and Hotelling [1933], independently. PCA seeks an orthogonal linear transformation that projects the original data into a new coordinate system in which the first dimension (or *principal component*) captures the largest possible amount of variance, the second dimension is orthogonal to the first while capturing the largest possible amount of remaining variance, and so on. A reduced dimension representation that captures an arbitrary amount of the total variance of the original data set can be constructed by projecting the original data onto the first d' principal components that account for this total variance.

A great number of statistical software packages can perform PCA computations. Available packages are all designed for sequential computers, however, and are incapable of handling the very

large data sets we use with the MSTC tool, due to both wall-clock time constraints and memory constraints. We have recently developed a distributed memory, parallel tool for PCA computations. It has been built using the Parallel Linear Algebra Package (PLAPACK) [van de Geijn, 1997], an object-oriented framework for expressing parallel dense numerical linear algebra computations at a high level of abstraction. To calculate the PCA, we do not explicitly form the covariance matrix, opting instead for the more numerically robust approach of computing the singular value decomposition (SVD) of the scaled, standardized data matrix. We use the SVD routine provided by PLAPACK, which performs parallel matrix-vector operations to reduce the matrix to bidiagonal form. A sequential bottleneck exists because the singular values of the bidiagonal matrix are then computed sequentially; this portion of the calculation is small, however, and happens quickly. After the PCA is computed, the matrix-matrix multiplication required to project the original data onto the principal components space is performed in parallel. Our parallel tool has enabled us to calculate the PCA on far larger data sets than has previously been possible: We recently computed the PCA of a data set consisting of 47 841 280 points of 131 variables each (a roughly 48 GB file when stored as single precision floating-point values) in only 1 533 seconds using 40 compute nodes (using one compute core per node) on Jaguar, the Cray XT4 system at Oak Ridge National Laboratory.

3 Environmental Applications

Initial applications of the cluster analysis methodology to environmental problems were described by Hargrove and Hoffman [2004b]. A sampling of these, along with more recent follow-on analyses, are presented below.

3.1 Delineation of Ecoregions

Ecoregions are useful to ecologists as simplified generalizations of complex combinations of climatic, edaphic, and topographic conditions that may be associated with the assemblages of plant and animal species for which those conditions permit growth and reproduction. Historically, ecoregions were delineated based on human expertise, and such subjective regionalizations are not reproducible by others since the methods and data employed are not available. An intermediate step toward repeatability was taken by Holdridge [1947] and Köppen and Geiger [1928, 1930-1939], who explicitly formulated and codified rules for discriminating one region or zone from another. However, these rules were simple thresholds applied sequentially to a small number of characteristics and did not consider interactions among characteristics. Two of us (Hargrove and Hoffman) began experimenting with clustering methods for objectively and repeatably generating quantitative ecoregions, using many co-registered data layers from a Geographic Information System.

Qualitative ecoregions were originally developed for generic purposes. The same ecoregions used for predicting the spread of an invasive plant would also be used to stratify sampling for a threatened animal species. Even Holdridge- and Köppen-type ecoregions are based on environmental factors deemed to be broadly important, like temperature and precipitation. This "one size fits all" philosophy decreases the utility of the ecoregion concept. Quantitative ecoregions, on the other hand, can be delineated from a custom-created combination of environmental characteristics believed to be relevant to the species or phenomena under investigation. For example, Figure 2 shows a set of 90 ecoregions in the continental United States developed from 30 factors chosen to stratify sampling from eddy-covariance flux towers. These ecoregions indicate areas within which the characteristics of CO_2 flux are expected to be about the same. Factors used include diurnal temperature differences, heat and cold degree day accumulations, precipitation sums, and numbers of wet days, all independently calculated for the growing and non-growing seasons. Also included are soil depth, nitrogen, and organic content of soil, a topographic relative wetness index, and remote sensing-derived values: Enhanced Vegetation Index (EVI), Fraction of Photosynthetically Active Radiation (FPAR), Gross Primary Productivity (GPP), and a Respiration Index (RI) from MODIS (the Moderate Resolution Imaging Spectroradiometer) [Hargrove and Hoffman, 2004a]. A generalized set of ecoregions is unlikely to be useful for such specific applications.

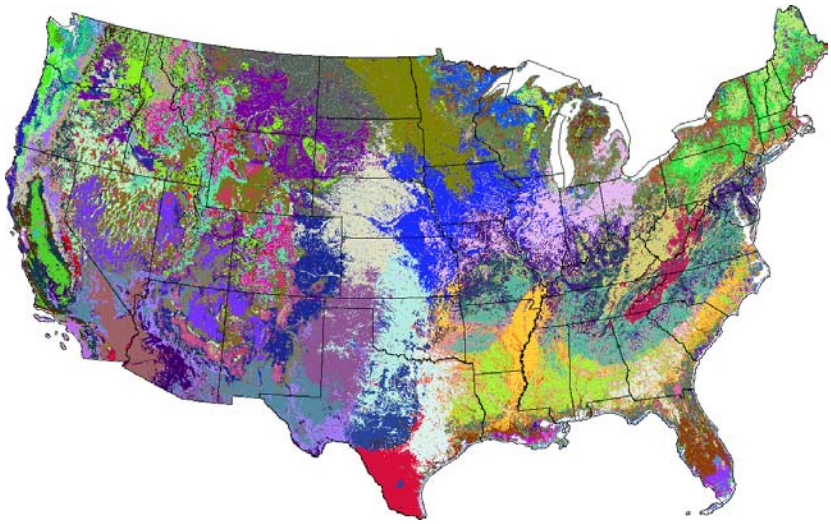


Figure 2: This map shows the 90 most-different ecoregions in the continental United States, colored randomly, custom developed for significance with respect to eddy-covariance flux tower sites.

3.2 Representativeness of Sampling Networks

The Euclidean distance from any region to the nearest training region indicates how well the mixture of conditions within that ecoregion represents the combination of conditions within the selected training area. The inverse of the same concept can be used to gage how well a network of sample locations represents the conditions found within a larger map that contains it. A network in this case can be a constellation of monitoring stations or locations where a set of samples are taken. By calculating the Euclidean distance from each ecoregion and the most similar of the sampling locations, we can map how well or poorly each ecoregion in the map is represented by the most similar station in the network. A map in which similar locations are colored white and dissimilar locations are colored black shows areas in the map that are well-represented and poorly represented (respectively) by this network. The sum of all Euclidean distances from each ecoregion to the most-similar network station inversely quantifies how well the network represents this area. Figure 3 shows the global representativeness of the international FluxNet network of eddy-covariance flux towers with respect to 14 characteristics that include climate and soil properties. Some incidental representativeness can be seen on continents not having any eddy-flux towers; however, some areas, like India and the Amazon, are very poorly represented by even the most similar station in the FluxNet network [Sundareshvar et al., 2007]. Such analyses can be used to design optimal sampling networks for any area before deployment [Schimel et al., 2007].

3.3 Model Diagnostics and Intercomparison

The MSTC technique has proved useful in evaluating the large time series output of global climate models. Such evaluations may be used to help diagnose model behavior, provide a basis for comparison of multiple ensemble simulations from one or more models, or extract patterns from model projections. MSTC was applied to the monthly time series output of temperature, precipitation, and soil moisture from five 99-year Business-As-Usual (BAU) transient simulations for years 2000–2098 from the Parallel Climate Model (PCM) [Hoffman et al., 2005]. In this case, the multivariate clusters form climate regimes in a climate phase space that vary seasonally and inter-annually. By treating these regimes as discrete climate states, the dynamic behavior of the climate system can be described as sequences of climate state occupancy. Cluster frequency plots showed that, over the course of these simulations, the conditions typical of Antarctic Winter (the coldest and driest climate state) shrank in land area while the conditions typical of desert regions (the hottest and driest climate state) grew significantly in land area.



Figure 3: This map shows the global representativeness of the FluxNet network of eddy-covariance flux tower sites (boxes) in levels of gray. Dark areas are poorly represented or sampled by the existing network while light and white areas are well represented.

3.4 Model-Observation Comparison

The same cluster analysis approach has been used to compare high frequency atmospheric observations with global climate model results [Mahajan et al., 2007]. The observational data—consisting of temperature and moisture at 48 levels in the atmosphere, wind speed at 62 levels, and surface pressure—were collected by the U.S. Department of Energy's Atmospheric Radiation Measurement (ARM) program at its Southern Great Plains (SGP) site centered near Lamont, Oklahoma, US. The model results were obtained from an ensemble member of a Community Climate System Model (CCSM) general circulation model (GCM) run under the IPCC SRES A2 scenario that produced 6-hourly output. A three-way process was employed to compare ARM observations with GCM output, where 1) CCSM output was projected onto states derived from ARM observations, 2) ARM observations were projected onto states derived from CCSM output, and 3) both ARM observations and CCSM output were projected onto states derived from the combination of the two data sets. Comparisons of twelve atmospheric states derived from the combination of ARM observations and CCSM output indicate that distinct singular states exist in each data set. That is, the model exhibited an atmospheric state that never appeared in the observations, while the observations contained atmospheric states never captured by the model. Comparing the populations of state frequencies in this manner points out model deficiencies and provides quantitative information to model developers about how models can be improved.

4 Conclusions

The MGC and MSTC methodologies are powerful techniques for detecting patterns in large data sets and generating useful products for ecological and climate analyses. Such data mining techniques are often the only tools that can be used to analyze very large data sets, like climate model results, without "hiding" patterns of potential interest within other statistics. The improved parallel implementation of the accelerated k-means clustering method with empty cluster "warping," combined with the PCA tool, can now be applied to data sets of unprecedented size. Future work will focus on applying these tools to change detection in satellite remote sensing data as an early-warning system for detecting forest threats.

Acknowledgments

This research was partially sponsored by the Climate Change Research Division (CCRD) of the Office of Biological and Environmental Research (OBER) within the U.S. Department of Energy's Office of Science (SC). This research was partially sponsored by the U.S. Department of Agriculture (USDA) Forest Service. This research used resources of the National Center for Computational Science (NCCS) at Oak Ridge National Laboratory (ORNL), which is managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

REFERENCES

Hargrove, W. W. and F. M. Hoffman. A Flux Atlas for Representativeness and Statistical Extrapolation of the AmeriFlux Network. Technical Memorandum ORNL/TM-2004/112, Oak Ridge National Laboratory, April 2004a.

Hargrove, W. W. and F. M. Hoffman. Potential of multivariate quantitative methods for delineation and visualization of ecoregions. *Environmental Management*, 34(5):s39–s60, 2004b. doi:10.1007/s00267-003-1084-0.

Hargrove, W. W., F. M. Hoffman, and T. Sterling. The Do-It-Yourself Supercomputer. *Scientific American*, 265(2):72–79, August 2001.

Hartigan, J. A. *Clustering Algorithms*. John Wiley & Sons, New York, 1975.

Hoffman, F. M. and W. W. Hargrove. Multivariate geographic clustering using a Beowulf-style parallel computer. In Arabnia, H. R., editor, *Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA '99)*, volume III, pages 1292–1298, Las Vegas, Nevada, June 1999. CSREA Press.

Hoffman, F. M., W. W. Hargrove, D. J. Erickson, and R. J. Oglesby. Using clustered climate regimes to analyze and compare predictions from fully coupled general circulation models. *Earth Interactions*, 9(10):1–27, August 2005. doi:10.1175/EI110.1.

Holdridge, L. R. Determination of world plant fomrations from simple climatic data. *Science*, 105:367–368, April 1947. doi:10.1126/science.105.2727.367.

Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441, 1933.

Köppen, W. P. and R. Geiger. Klimakarte der Erde. *Gotha*, 1928.

Köppen, W. P. and R. Geiger. *Handbuch der Klimatologie*, volume 1–5. Berlin, 1930-1939.

Mahajan, S., F. M. Hoffman, W. W. Hargrove, S. W. Christensen, and R. T. Mills. A cluster analysis approach to comparing Atmospheric Radiation Measurement (ARM) observations with general circulation model (GCM) results. *Eos Trans. AGU*, 88(52), December 2007. Fall Meet. Suppl., Abstract A41A-0010.

Pearson, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.

Phillips, S. J. Acceleration of k-means and related clustering algorithms. In Mount, D. M. and Stein, C., editors, *ALENEX '02: Revised Papers from the 4th International Workshop on Algorithm Engineering and Experiments*, pages 166–177, London, UK, 2002a. Springer-Verlag.

Phillips, S. J. Reducing the computation time of isodata and k-means unsupervised classification algorithms. In *Geoscience and Remote Sensing Symposium, 2002 (IGARSS'02)*, volume 3, pages 1627–1629, June 2002b. doi:10.1109/IGARSS.2002.1026202.

Schimel, D., W. Hargrove, F. Hoffman, and J. McMahon. NEON: A hierarchically designed national ecological network. *Frontiers in Ecology and the Environment*, 5(2):59, March 2007.

Sundareshvar, P. V., R. Murtugudde, G. Srinivasan, S. Singh, K. J. Ramesh, R. Ramesh, S. B. Verma, D. Agarwal, D. Baldocchi, C. K. Baru, K. K. Baruah, G. R. Chowdhury, V. K. Dadhwal, C. B. S. Dutt, J. Fuentes, P. K. Gupta, W. W. Hargrove, M. Howard, C. S. Jha, S. Lal, W. K. Michener, A. P. Mitra, J. T. Morris, R. R. Myneni, M. Naja, R. Nemani, R. Purvaja, S. Raha, S. K. S. Vanan, M. Sharma, A. Subramaniam, R. Sukumar, R. R. Twilley, and P. R. Zimmerman. Environmental monitoring network for India. *Science*, 316:204–205, April 2007.

van de Geijn, R. A. *Using PLAPACK: Parallel Linear Algebra Package*. MIT Press, Cambridge, MA, 1997.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A Particle Swarm Optimization derivative applied to cluster analysis

J.L. Díaz, M. Herrera, J. Izquierdo, I. Montalvo, R. Pérez

Centro Multidisciplinar de Modelación de Fluidos, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain
(jldiaz,mahefe,jizquier,imontalvo,rperez@gmmf.upv.es)

Abstract: Modern machine learning and data analysis hinge on sophisticated search techniques. In general, exploration in high-dimensional and multi-modal spaces is needed. Some algorithms that imitate certain natural principles, the so-called evolutionary algorithms, have been used in different aspects of Environmental Science and have found numerous applications in Environmental related problems. In this paper we apply a derivative of PSO (Particle Swarm Optimization), recently introduced by the authors to partitional clustering of a real-world data set obtained from a Water Supply Company. The PSO derivative we consider here improves several typical features of this optimization technique. For one thing, PSO is adapted to consider mixed discrete-continuous optimization since the problem we tackle here involves the use of both continuous and discrete variables. For another, one of the main drawbacks associated with PSO comes from the fact that it is difficult to keep good levels of population diversity and to balance local and global searches. This formulation is able to find optimum or near-optimum solutions much more efficiently and with considerably less computational effort because of the richer population diversity it introduces. Finally, the cumbersome aspect, common to all metaheuristics, of choosing the right parameter values is tackled through self-adaptive dynamic parameter control.

Keywords: Partitional clustering; Optimization; Evolutionary algorithm; Particle Swarm Optimization.

1. INTRODUCTION

Modern machine learning and data analysis hinge on sophisticated search techniques. Computer systems that are able to extract information from large amounts of data, that is to say, to perform Data Mining tasks, like pattern recognition, classification, diagnosis, etc. and, in general, systems that are adaptive and show capacity to learn, fundamentally rely on effective and efficient search techniques. Any adaptive system needs some kind of search mechanism in order to explore a feature space describing all possible states of the system. Due to the characteristics of many feature spaces exploration in high-dimensional and multi-modal spaces is needed.

Classical methods of optimization involve the use of gradients or higher-order derivatives of the fitness function. But they are not well suited for many real world problems since they are not able to process inaccurate, noisy, discrete and complex data [Bonabeau et al., 1999; Kennedy and Eberhart, 2001]. Thus robust methods of optimization are often required to generate suitable results. Some algorithms that imitate certain natural principles, evolutionary algorithms like Genetic Algorithms, Ant Colony Optimization, Particle Swarm Optimization, Harmony Search, etc., have been used in different aspects of Environmental Science and have found numerous applications in Environmental related problems [Downing, 1998; López, 2001; Nishida et al., 2004; Vojinovic and Solomatine, 2005; Afshar and Mariño, 2006; Crowe et al., 2006; Katsifarakis and Petala, 2006; Valdés

and Barton, 2007; Karterakisa et al., 2007; Montalvo et al., 2008a, 2008b; Izquierdo et al., 2007, 2008a, 2008b].

One of the evolutionary algorithms that has shown great potential and good perspective for the solution of various optimization problems [Dong et al., 2005; Janson et al., 2008; Jin et al., 2007; Liao et al., 2007; Pan et al., 2007; Montalvo et al., 2008a, 2008b; Izquierdo et al., 2008b] is Particle Swarm Optimization (PSO). Swarm intelligence is a relatively new category of stochastic, population-based optimization algorithms that are closely related to evolutionary algorithms based on procedures that imitate natural evolution. Swarm intelligence algorithms draw inspiration from the collective behaviour and emergent intelligence that arise in socially organized populations.

In this paper we apply a derivative of PSO, recently introduced by the authors [Izquierdo et al., 2008b; Montalvo et al., 2008a, 2008b] to partitional clustering of a real-world data set obtained from a Water Supply Company. In addition, we endow this PSO derivative with a self-adaptive feature that manages to internally control its parameters.

Clustering analysis [Everitt, 1980] plays an important role in many fields and can be used both for preliminary and descriptive data analysis and unsupervised classification [Hastie et al., 2001], and to summarize common features of groups of elements, like identification of centroids or baricenters. Central to all of the goals of cluster analysis is the notion of *similarity*, in terms of proximity, between the individual objects being clustered (otherwise, *dissimilarity* is used to explain the difference). A clustering method attempts to group the objects based on the definition of similarity supplied to it. In the present paper we will work with clusters based in a mixed dissimilarity. For this reason we will use medoids like a representative grouping element (understanding medoid as the object of a cluster, whose average dissimilarity to all the objects in the cluster is minimal).

The database utilized consists in a record of requests, complains and claims (PQR's in Spanish), for the year 2006 of Calarcá Water Supply Company Multipropósito, S.A. ESP, located in a town of the coffee Colombian region. These records are combined with the information of the network model for this year. The results obtained in this work are important to aid the management of the network and the decision making of most relevant places to be considered in future rehabilitations.

The PSO derivative we consider here is adapted to consider mixed discrete-continuous optimization since the problem we tackle here involves the use of both continuous and discrete variables and will work jointly with statistical clustering criteria arranged to take these type of mixed measures. Also, this formulation is able to find optimum or near-optimum solutions much more efficiently and with considerably less computational effort because of the richer population diversity it introduces. Finally, the cumbersome aspect, common to all metaheuristics, of choosing the right parameter values is tackled through self-adaptive dynamic parameter control.

The remainder of this paper is organized as follows. Next section presents the rules for the manipulation of the particles in each iteration and explains how parameters are controlled. Also, the main features of the PSO derivative we consider here are introduced. Then, the database under consideration is concisely described and the main objectives in this study presented. Next, section 4 introduces the necessary statistical clustering criteria including the description of the fitness evaluation and the search space. Finally, the main results are reported. A conclusions section wraps up the paper.

2. PSO AND THE PROPOSED DERIVATIVE

A swarm consists of a set of particles moving within the search space, which is D-dimensional, each representing a potential solution of the problem. Each particle has a position vector, $X_i = (x_{i1}, \ldots, x_{iD})$, a velocity vector, $V_i = (v_{i1}, \ldots, v_{iD})$ and the position at which the best fitness was encountered by the particle, $Y_i = (y_{i1}, \ldots, y_{iD})$. In each cycle of the evolution the position of the best of the N particles in the swarm, Y^* is identified.

2.1 Manipulation of particles

In each generation, the velocity of each particle is updated by means of its velocity history, its best encountered position and the best position encountered by any particle:

$$V_i = \omega V_i + c_1\, rand(\,)\,(Y_i - X_i) + c_2\, rand(\,)\,(Y^* - X_i), \quad (1)$$

On each dimension, particle velocities are clamped to minimum and maximum velocities to control excessive roaming of particles outside the search space:

$$V_{\min} \le V_j \le V_{\max}. \quad (2)$$

Usually $V_{\min}$ is taken as $-V_{\max}$, which is a user defined parameter.

The position of each particle is also updated every generation. This is done by adding the velocity vector to the position vector,

$$X_i = X_i + V_i. \quad (3)$$

The parameters are as follows: ω is a factor of inertia suggested by Shi and Eberhart [1998] that controls the impact of the velocity history into the new velocity. Acceleration parameters c_1 and c_2 are typically two positive constants, called cognitive and social parameter, respectively. *rand*( ) generates two independent random numbers between 0 and 1, and are used to maintain the diversity of the population.

2.2 Manipulation of parameters

The role of the inertia, ω, in (1), is considered critical for the PSO algorithm's convergence behaviour. Although initially the inertia was constant it may vary from one cycle to the next. As it permits to balance out global and local searches, it was suggested to have it decrease linearly with time, usually in a way to first emphasize global search and then, with each cycle of the iteration, prioritize local search, [Shi and Eberhart, 1999]. A significant improvement in the performance of PSO with the decreasing inertia weight over the generations is achieved by using [Jin et al., 2007]

$$\omega = 0.5 + \frac{1}{2(\ln(k)+1)}, \quad (4)$$

where k is the iteration number. In the framework herein described this parameter is adaptively controlled by using (4).

However, the acceleration coefficients and the clamping velocity are neither set to a constant value, like in standard PSO, nor set as a time varying function, like in adaptive PSO variants [Ratnaweera and Halgamuge, 2004; Aramugan and Rao, 2008]. Instead they are incorporated to the own optimization problem. Each particle will be allowed to self-adaptively set its own parameters by using the same process used by PSO given by equations (1) and (3). To this end, these three parameters are considered as three new variables that are incorporated to position vectors X_i. In general, if D is the dimension of the problem and P is the number of self-adapting parameters, the new position vector for particle i will be:

$$X_i = (x_{i1}, \ldots, x_{iD}, x_{iD+1}, \ldots, x_{iD+P}). \quad (5)$$

It is clear that the first D variables correspond to the real position vector of the particle in the search space, while the last P account for its personal parameters. Obviously, these new variables do not enter the fitness function, but are manipulated by using the same mixed individual-social learning paradigm used in PSO.

Also, V_i and Y_i, giving the velocity and best so far position for particle i, increase their dimension, with corresponding meaning:

$$V_i = (v_{i1}, \ldots, v_{iD}, v_{iD+1}, \ldots, v_{iD+P}) \text{ and} \quad (6)$$

$$Y_i = (y_{i1}, \ldots, y_{iD}, y_{iD+1}, \ldots, y_{iD+P}). \quad (7)$$

This way, by using equations (1) and (3), each particle will be endowed additionally with the ability of adjusting its parameters by aiming to both the parameters it had when it got its best position in the past and the parameters of the leader, which managed to take this best particle to its privileged position. As a consequence, particles not only use their cognition of individual thinking and the social cooperation to improve their positions but also to improve the way they do it by accommodating themselves to the best known conditions, namely, their conditions when getting the best so far position and the leader's conditions.

Before providing a schematic representation of the proposed algorithm two more observations have to be made.

For one thing, the discussion so far considers the standard PSO algorithm, which is applicable to continuous systems and cannot be used for mixed discrete-continuous problems, like the one we consider here. To tackle discrete variables this algorithm takes integer parts of the flying velocity vector discrete components into account; hence the new discrete component velocities V_i are integer and consequently the new position vector discrete components will also be integer (since the initial position vectors were generated with integer values). According to this idea, velocity updating for discrete variables turns out to be:

$$V_i = fix(\omega V_i + c_1\, rand(\,)\,(Y_i - X_i) + c_2\, rand(\,)\,(Y^* - X_i)), \tag{8}$$

where $fix(\cdot)$ implies that we only take the integer part of the result.

For another, in [Montalvo et al., 2007b], PSO was endowed with a re-generation-on-collision formulation, which further improves the performance of standard discrete PSO. The random regeneration of the many birds that tended to collide with the best birds was shown to avoid premature convergence, as it prevented clone populations from dominating the search. The inclusion of this procedure into the discrete PSO produces greatly increased diversity, improved convergence characteristics and higher quality of the final solutions. The modified algorithm can be given by the following pseudo-code, with k as iteration number.

---

1) $k = 0$
2) Generate a random population of M particles: $\{X_i(k)\}_{i=1}^{M}$, according to (5)
3) Evaluate the fitness of the particles (only the first D variables enter the fitness function)
4) Record the local best locations $\{Y_i(k)\}_{i=1}^{M}$; according to (7) the values of the corresponding parameters are also recorded
5) Record the global best location, $Y^*(k)$, and the list of the m best particles to check collisions (including their corresponding parameters)
6) While (not termination-condition) do
   a) Determine the inertia parameter $\omega(k)$, according to (4)
   b) Begin cycle from 1 to number of particles M
      Start
         (1) Calculate new velocity, $V_i(k+1)$, for particle i according to (1), and take its integer part (for discrete optimization) for the first D variables, according to (8)
         (2) Update position, $X_i(k+1)$, of particle i according to (3)
         (3) Calculate fitness function for particle i and update Y_i
         (4) If particle i has better fitness value than the fitness value of the best particle in history, then set particle i as the new best particle in history
         (5) If particle i is not currently the best particle but coincides with the best, then re-generate particle i randomly (including its parameters)
      End
   c) $k = k + 1$
7) Show the solution given by the best particle

In this study, a population size of M = 100 particles has been used. Also, among the different termination conditions that may be stated, a condition stopping the process if there is no improvement after a pre-fixed number of iterations has been considered.

The performance of the approach here introduced can be observed from the results reported in the next section.

3. THE DATABASE

The database utilized consists in a record of requests, complains and claims (PQRs in Spanish) issued in Calarcá Water Supply Company Multipropósito, S.A. ESP, for the year 2006. The municipality of Calarcá (Colombia) is located in the Andean area, it has a land area of 21,923 ha; 244 ha belongs to urban zones and 21,679 ha to rural sectors. Population is about 73,500 inhabitants.

These PQRs are reports of the users both in principal and domiciliary network sections. The PQRs indicated the type or description of damages, their locations, relevant technical concepts and the solutions. There were 846 records registered in that year.

Every record was located in a chart of the water network according to the address referenced in the PQR record; addresses were only in terms of street names and numbers, so a hard work had to be done for obtaining the UTM coordinates of every problem reported.

The main pre-processing task involved the selection of relevant and non relevant fields in the database. Breakage dates and times were excluded since their occurrence was deemed to depend strongly on the physical and working conditions of the network. Also, the names of personnel on duty, the repair date and the theoretical roughness of the pipe were neglected. Decision about what fields to include were made based on hydraulic criteria. Since rehabilitation was the main objective, geographical locations suggesting causes and occurrence of water loses were assessed of paramount importance. As a consequence, the information used included: pipe identification, to assess if it was subjected to a high or low number of faults; upstream and downstream node identification, to evaluate concurrence of faulty pipes on the same node pointing to pressure or demand problems at the node; type of reported breakage, either domiciliary or on the main network; pipe diameter; pipe length; pipe material and magnitude of the leak. In addition, data obtained from the mathematical model of the network were included in the database. Specifically, information related with demand characteristics and patterns. However, pressure data were not included, since pressure, being a decisive agent of water loses and breakages, would have blurred all the other specifically sought causes of the problem under consideration, more connected with materials, lengths, diameters, demand patterns, etc., in close relationship with rehabilitation purposes. Finally, also the UTM coordinates of the fault points were included.

As a matter of fact, typical pre-processing tasks for identifying outliers, missing values, etc., were performed. As a consequence, some records were modified, withdrawn, completed, etc.

4. STATISTICAL MEASURES AND FITNESS EVALUATION

4.1 Introduction

Clustering is the grouping of similar objects [Everitt, 1980]. An object can be described by a set of measurements or by its relation to other objects. The goals of cluster analysis are varied and include wide activities such looking for “natural” groups, hypothesis generation etc. Central to all of the goals of cluster analysis is the notion of similarity, in terms of proximity, between the individual objects being clustered (otherwise, dissimilarity is used to explain the difference). A clustering method attempts to group the objects based on the definition of similarity or dissimilarity supplied to it.

4.2 Dissimilarities

The dissimilarity between two objects measures how different they are [Hastie et al., 2001]. It has to be noted here that, although usual metrics can be used, they must not necessarily verify the triangle inequality.

The computation of the dissimilarity between two objects depends on the type of the original variables. Many data sets contain variables of different types. The next method solves the computation of the dissimilarity in a general form, considering that the data set contains *p* variables:

$$d(i,j) = \frac{\sum_{f=1}^{p} \partial_{ij}^{(f)} d_{ij}^{(f)}}{\sum_{f=1}^{p} \partial_{ij}^{(f)}} \quad \in [0,1] \tag{9}$$

With $\partial_{ij}^{(f)} = 0$ if x_{if} or x_{jf} is missing, or if $x_{if} = x_{jf} = 0$ and f is an asymmetric binary variable. Otherwise, $\partial_{ij}^{(f)} = 1$. $d_{ij}^{(f)}$ is the contribution of variable *f*, which depends on its type:

1. If *f* is binary or nominal, $d_{ij}^{(f)} = 0$ if $x_{if} = x_{jf}$ and $d_{ij}^{(f)} = 1$ otherwise.
2. If *f* is interval-scaled, $d_{ij}^{(f)} = \frac{|x_{if} - x_{jf}|}{\max_h(x_{hf}) - \min_h(x_{hf})}$.
3. For ordinal and ratio-scaled variables, ranks r_{if} and $z_{if} = \frac{r_{if} - 1}{M_f - 1}$ are computed, and then z_{if} is treated as interval-scaled.

4.3 Clustering algorithms

The goal of cluster analysis is to partition the observations into groups so that the pair-wise dissimilarities between those assigned to the same cluster tend to be smaller than those in different clusters. Among others, clustering algorithms can be classified in two categories: partitioning and hierarchical algorithms. Most partitional clustering algorithms assume an a priori number of clusters, *c*, and a partition of the data set into *c* clusters. To get the correct partition, an objective function must be formulated that measures how good a partition is with respect to the data set. Hierarchical clustering algorithms transform a proximity data set into a tree-like structure. The main drawbacks of these algorithms are its high computational cost and that always suffer from the problem of not knowing where to cut the generated tree.

In real-life problems very large data sets containing variables of several types are typically found. This requires for a clustering algorithm to be scalable and capable of handling different attribute types. Classical methods are not the answer: for example, PAM (Partitioning Around Medoids) algorithm [Kaufman & Rousseeuw, 1990] can handle various attribute types but is not efficient with large data sets. *k*-means algorithms [Hartigan & Wong, 1979; Likas et al., 2003] can handle large data sets but deal with only data sets formed from interval-scaled variables. CLARA (Clustering Large Applications) algorithm [Kaufman & Rousseeuw, 1990] is a combination of sampling approach and the PAM algorithm. Instead of finding medoids, each of which is the most located object in a cluster for the entire data set, CLARA draws a sample from the data set and uses the PAM algorithm to select an optimal set of medoids from the sample [Wei et al., 2003]. To alleviate sampling bias, CLARA repeats the sampling and clustering process multiple times and selects the best set of medoids as the final clustering. The objective function is the next:

$$Cost(M,D) = \frac{\sum_{i=1}^{n} d(O_i, rep(M, O_i))}{n} \tag{10}$$

where *D* is the data set to be clustered, *n* the number of objects in *D*, *M* is a set of selected medoids, $d(O_i,O_j)$ is the dissimilarity between objects O_i and O_j, and $rep(M,O_i)$ returns the medoid in *M* which is closest to O_i.

4.4 Statistical clustering criteria

Several clustering criteria have been proposed and three of them are based on the fundamental matrix equation: $T = W + B$, where W and B are the within-cluster and between-cluster variation, respectively. T is, then, the total scatter matrix. For univariate data the above expression reduces to usual ANOVA decomposition. Under this point of view, the ideal form of T is a matrix built with a small W and a large B, so that the distances within the clusters are small compared with distances between clusters medoids. Then, an intuitive procedure for choosing clusters is to minimize the "size" of W and/or maximize B.

The statistical criteria to measure the adequacy of the partition and define the optimization problem to solve the clustering paradigm via heuristics are based in W and B. Friedman and Rubin [1967] proposed minimizing the W trace. Another idea is to minimize $det(W)$ or maximize $trace(BW^{-1})$. More recently, McGregor et al. [2004] have developed new methodologies for validation results based in W too. Barbará et al. [2002] have worked with entropy based measures for categorical data clustering.

5. RESULTS

The PSO algorithm was run several times and results were almost identical. A population size of only 30 particles was used. Maximum and minimum velocities were established as:

- Maximum velocity for discrete variables = 50% of variable range
- Minimum velocity = - Maximum velocity

The termination condition stopped the process if after 20 iterations no improvement in the solution had been obtained. Results were obtained in a mean value of only 40 iterations.

Clusters were made using PSO and considering different possibilities: 2, 3 and 4 clusters. The search space was multi-dimensional and all dimensions were not in correspondence with the same type of variable. Clusters were analysed and some conclusions can rapidly be drawn:

- Two main different groups were identified; the first one was represented mainly by concrete and PVC pipes, while most of the pipes were made of PVC in the second one.
- Pipes length was established as short, medium and long, based on certain mean values and the range of lengths determined for all the pipes in the database. In the first group, most of the pipes were long, and short pipes were less significative. In the second group the amount of medium and long pipes was almost the same and short pipes were more significative.

In correspondence with cluster analysis, it was shown that problems were concentrated on medium size pipes made of concrete and on large size pipes made of PVC. Also some problems on short pipes made of PVC were detected. Either the rest of materials or medium PVC pipes did not have relevant influence on the problems detected.

6. CONCLUSIONS

Data mining analysis helped to discover where most of the problems in a real water distribution network were concentrated. Cluster analysis was carried out using the PSO algorithm. Results can be used as a strategic plan for network rehabilitation, considering that attention should go first where problems seems to appear more frequently.

Richer results could be obtained incorporating some new fields in the database. In that case, it would be necessary to use a bigger population size for solving the problem, but with no added conceptual difficulty.

PSO algorithm was compared to partitional clustering performed by other algorithms that work with various attribute types (such as PAM and CLARA). There were not significant differences on what was obtained, while PSO efficiency was superior. Thus those results

show the PSO's ability in cluster building. Bigger sizes of the database would enhance PSO superiority in efficiency terms (time of execution and best optimization solution), avoiding the need of adapting to new types of data, as happened with PAM, *k*-means and CLARA. However, PSO clustering should also be compared with CLARANS (CLARA based on Randomized Search) and bagged clustering procedures, based on the bootstrap re-sampling.

Searching for knowledge in databases is really necessary for the water supply sector. Incorporating a tool for clustering analysis to the pool of software packages related to water distribution networks could be very useful. However, it should be also important to incorporate some tools for analysing clustering results.

ACKNOWLEDGEMENTS

This work has been performed under the support of the projects Investigación Interdisciplinar nº 5706 (UPV) and DPI2004-04430 of the Dirección General de Investigación del Ministerio de Educación y Ciencia (Spain), including the support for I+D+i projects from the Consellería de Empresa, Universidad y Ciencia of the Generalitat Valenciana, and FEDER funds. Thanks also to the support of the Grant MAEC-AECI 0000202066 awarded to one of the authors by the Ministerio de Asuntos Exteriores y Cooperación of Spain.

A special gratitude must be acknowledged to Calarcá Water Supply Company Multipropósito, S.A. ESP for the collaboration with the data used in this work.

REFERENCES

Afshar, M.H., and Mariño, M.A., Application of an ant algorithm for layout optimization of tree networks, *Engineering Optimization*, 38(3), 353–369, 2006.

Arumugam, M.S., and Rao, M.V.C., On the improved performances of the particle swarm optimization algorithms with adaptive parameters, cross-over operators and root mean square (RMS) variants for computing optimal control of a class of hybrid systems, Applied Soft Computing 8 (2008) 324–336.

Barbará, D.; Couto, J.; Li, Y. (2002) "An entropy-based algorithm for categorical clustering" Proceedings of on Information and Knowledge Management (CIKM, 2002)

Bonabeau, E., Dorigo, M., and Théraulaz, G., From Natural to Artificial Swarm Intelligence, Oxford University Press, New York, 1999.

Crowe, A.M., McClean, C.J., and Cresser, M.S., An application of genetic algorithms to the robust estimation of soil organic and mineral fraction densities. *Environmental Modelling & Software*, 21, 1503-1507, 2006.

Dong, Y., Tang, B.X.J., and Wang, D., "An application of swarm optimization to nonlinear programming," Computers & Mathematics with Applications 49 (11-12), pp. 1655–1668, 2005.

Downing, K., Using evolutionary computational techniques in environmental modelling. *Environmental Modelling & Software*, 13 (5-6), 519-528, 1998.

Everitt, B. (1980) "Cluster Analysis" *Biometrics*, Vol. 37, No. 2 (Jun., 1981), pp. 417-418

Hartigan, J. A. and Wong, M. A. (1979) "A K-means clustering algorithm" Applied Statistics 28, 100–108

Hastie, T; Tibshirani, R.; Friedman, J. (2001) "Elements of Statistical Learning: Data Mining, Inference and Prediction" Springer-Verlag, New York

Izquierdo, J., López, P.A., Martínez, F.J., Pérez, R. Fault Detection in Water Supply Systems by Hybrid (Theory and Data-Driven) Modelling. Mathematical and Computer Modeling, 46/3-4, pp. 341-350, 2007.

Izquierdo, J., Díaz, J.L., Pérez, R., López, P.A., and Mora, J.J., Knowledge Discovery in Environmental Data. In: Meire, P., Coenen, M., Lombardo, C., Robba, M., Sacile, R. (Eds.), Integrated Water Management, vol. 80, Springer, 2008a.

Izquierdo, J., Montalvo, I., Pérez, R., Fuertes, V.S. Design optimization of wastewater collection networks by PSO. *Computer & Mathematics with Applications*. doi:10.1016/j.camwa.2008.02.007, 2008b.

Janson, S., Merkle, D., and Middendorf, M., "Molecular docking with multi-objective Particle Swarm Optimization," Applied Soft Computing 8, pp. 666–675, 2008.

Jin, Y. X., Cheng, H. Z., Yan, J.I., and Zhang, L., "New discrete method for particle swarm optimization and its application in transmission network expansion planning," Electric Power Systems Research 77(3-4), pp. 227-233, 2007.

Karterakisa, S.M., Karatzasa, G.P., Nikolosb, I.K., and Papadopouloua, M.P., Application of linear programming and differential evolutionary optimization methodologies for the solution of coastal subsurface water management problems subject to environmental criteria. *Journal of Hydrology*, 342, 3-4, doi:10.1016/j.jhydrol.2007.05.027

Katsifarakis, K.L., and Petala, Z., Combining genetic algorithms and boundary elements to optimize coastal aquifers' management. *Journal of Hydrology*, 327(1-2), 200-207, 2006.

Kaufman, L. and Rousseeauw, P. J. (1990) "Finding gropus in data: an introduction to cluster analysis" Wiley, New York

Kennedy, J. and Eberhart, R.C., Swarm Intelligence, Morgan Kaufmann, 2001.

Liao, C.J., Tseng, C.T., and Luarn, P., "A discrete version of particle swarm optimization for flowshop scheduling problems," Computers and Operations Research, 34(10), pp. 3099-3111, 2007.

Likas, A.; Vlassis, N.; Vebeek, J. L. (2003) "The global K-means clustering algorithm" Pattern Recognition 36, pp. 451 – 461

López, P.A., Metodología para la calibración de modelos matemáticos de dispersión de contaminantes incluyendo regímenes no permanentes. Doctoral Dissertation. Polytechnic University of Valencia (Spain), 2001.

McGregor, A.; Hall, M.; Lorier, P.; Brunskill, J. (2004) "Flow Clustering Using Machine Learning Techniques" Lecture Notes in Computer Science (Passive and Active Network Measurements), vol. 3015/2004, Springer - Berlin

Montalvo, I., Izquierdo, J., Pérez, R., Tung, M.M. Particle Swarm Optimization applied to the design of water supply systems. *Computer & Mathematics with Applications*, doi:10.1016/j.camwa.2008.02.006, 2008a.

Montalvo, I., Izquierdo, J., Pérez, R., Iglesias, P.L. A diversity-enriched variant of discrete PSO applied to the design of Water Distribution Networks. *Engineering Optimization*. DOI:10.1080/03052150802010607, 2008b.

Nishida, W., Noguchi, M., Matsushita, H., and Solomatine, D.P., A Study on the Application of Genetic Algorithm to Calibration of Water Quality Model. *Ann. J. of Hydraulic Engineering* 48(2), 1321-1326, 2004.

Pan, Q.K., Tasgetiren, F., and Liang, Y.C., "A discrete particle swarm optimization algorithm for the no-wait flowshop scheduling problem," Computers and Operations Research, doi: 10.1016/j.cor.2006.12.030, 2007.

Ratnaweera, A., and Halgamuge, S.K., Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficient, IEEE Trans. Evol. Comput. vol. 8 (June (3)) (2004) 240–255.

Shi, Y. and Eberhart, R.C., "A modified particle swarm optimizer," in Proc. of the IEEE Congress on Evolutionary Computation, Piscataway, NJ, 1998, pp. 69-73.

Shi, Y. and Eberhart, R.C., Empirical study of particle swarm optimization, Proceedings of Congress on Evolutionary Computation, IEEE Service Center, Piscataway, NJ, 1999.

Valdés, J., and Barton, A., Multi-objective Evolutionary Optimization of Neural Networks for Virtual Reality Visual Data Mining: Application to Hydrochemistry. *Proceedings of 2007 IEEE International Joint Conference on Neural Networks*. Orlando, Florida, USA. August 12-17, 2007.

Vojinovic, Z. and Solomatine, D.P., Multi-criterial global evolutionary optimisation approach to rehabilitation of urban drainage systems. *Geophysical Research Abstracts*, Vol. 7, 10720. EGU General Assembly, Vienna, 2005.

Wei, C.; Lee, Y.; Hsu, C. (2003) "Empirical comparison of fast partitioning-based clustering algorithms for large data sets" Experts Systems with Applications 24, pp. 351 – 363

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Automatic generation of conceptual descriptions of classifications in Environmental Domains

Alejandra Perez-Bonilla [a], **Karina Gibert**[a] **and Darko Vrecko**[b]

[a]*Department of Statistics and Operations Research. Technical University of Catalonia, Edif. C5, Jordi Girona 1-3, 08034 Barcelona, Spain (alejandra.perez@upc.edu, karina.gibert@upc.edu)*

[b]*Department of Systems and Control. Jozef Stefan Institute, Jamova 39. Ljubljana. Slovenia (darko.vrecko@ijs.si)*

Abstract: In this paper the *Methodology of conceptual characterization by embedded conditioning CCEC*, oriented to the automatic generation of conceptual descriptions of classifications that can support later decision-making in Environmental Domains is applied to the interpretation of previously identified classes characterizing situations on a Waste Water Treatment Plant (WWTP). The particularity of the method is that it provides an interpretation of a partition previously obtained on an *ill-structured domain*, on the basis of a previous hierarchical clustering. The methodology uses some statistical tools (such as the *multiple boxplot*) together with artificial intelligent tools (as some machine learning methods), to learn the structure of the data; this allows extracting useful information (using the concept of *characterizing variable*) for the automatic generation of a set of useful concepts for later identification of classes. In this paper the usefulness of *CCEC* for building domain theories as models for supporting later decision-making is addressed and contrasted with interpretation provided by experts.

Keywords: Knowledge Discovery and Data Mining, Hierarchical clustering, class interpretation, Induction rules, Waste Water treatment plants.

1 INTRODUCTION

In automatic classification where the classes composing a certain domain are to be discovered, one of the most important required processes and one of the less standardized, is the interpretation of the classes (Gordon [1994]), closely related with *validation*, and critical in the later usefulness of the discovered knowledge. The interpretation of the classes, so important to understand the meaning of the obtained classification as well as the structure of the domain, used to be done in an artistic-like way. But this process becomes more and more complicated as the number of classes grows. This work is involved with the automatic generation of useful interpretations of classes in such a way that decisions about the action associated to a new object can be modelled and it is oriented to develop, in the long term, decision support system.

The presented proposal integrates different findings from a series of previous works; Gibert [1996], Pérez-Bonilla and Gibert [2007] proposed a single methodological tool which takes advantage of the hierarchical structure of the clustering to overcome some of the limitations observed in Gibert [1996], Gibert et al. [1998], Gibert and Pérez-Bonilla [2006]. This paper is organized as follows. After the introduction, basics concepts are presented in §2, on order to be understand the methodology which is presented in §3. WWTP as well as description of the specific data base that has been analyzed are presented in §4. Results of applying *CCEC* to the data described are given in §5. Finally in §6 the conclusions and the future work are addressed.

2 BASIC CONCEPTS

Four main concepts of methodology CCEC that are used in this work (basic notation is introduced in Figure 1(left)) are:

- Support (Sup): Given a rule $r : x_{ik} \in I_s^k \longrightarrow C$, where $I_s^k \subset r_k$(see Figure 1(left)), the support of r is the proportion of objects in $\mathcal{I}$ that satisfy the antecedent of the rule, Liu [2000]. $Sup(r) = \frac{card\{i \in \mathcal{C} \ : \ x_{ik} \in I_s^k\}}{n}$
- Relative covering $(CovR)$: Given a rule, the relative covering is the proportion of class C that satisfy the rule. $CovR(r) = \frac{card\{i \in C \ : \ x_{ik} \in I_s^k\}}{card\{C\}} * 100$
- Confidence $(Conf)$: Given a rule, the confidence of r is the proportion of objects in $\mathcal{I}$ that satisfy the antecedent of the rule and belong to C, Liu [2000]. $Conf(r) = \frac{card\{i \in C \ : \ x_{ik} \in I_s^k\}}{card\{ \ x_{ik} \in I_s^k\}}$
- *B*oxplot based discretization (BbD), see Pérez-Bonilla and Gibert [2007]), as an efficient way of transforming a numerical variable into a qualitative one in such a way that the resulting qualitative variable maximizes the association with the reference partition. Basically the cut points are determine by the minimum and maximum values that the numerical variable take in a very groups induced by the categorical one.
- The methodology *boxplot based induction rules (BbIR)* is presented in Pérez-Bonilla and Gibert [2007]. It is a method for inducing probabilistic rules ($r : x_{ik} \in I_s^k \xrightarrow{p_{sc}} C$, $p_{sc} \in [0, 1]$ is a degree of certainty of r) with a minimum number of attributes in the antecedent, based on the *BbD* of X_k.

The standard input of a clustering algorithm is a data matrix with the values of K variables $X_1 \dots X_K$ (numerical or not) observed over a set $\mathcal{I} = \{1, \dots n\}$ of individuals. Variables are in columns, while individuals in rows. Cells contain the value (x_{ik}), taken by individual $i \in \mathcal{I}$ for variable $X_k, (k = 1 : K)$. The set of values of X_k is named $\mathcal{D}^k = \{c_1^k, c_2^k, ..., c_s^k\}$ for categorical variables and $D^k = r_k$ for numerical ones, being r_k = [$minX_k$,$maxX_k$] the range of X_K. A partition in ξ classes of $\mathcal{I}$ is denoted by $\mathcal{P}_\xi = \{C_1, ..., C_\xi\}$, and $\tau = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4, ..., \mathcal{P}_n\}$ is an indexed hierarchy of $\mathcal{I}$. Finally, $\mathcal{P}_2 = \{C_1, C_2\}$ is a binary partition of $\mathcal{I}$. Usually, τ is the result of a *hierarchical clustering* over $\mathcal{I}$, and it can be represented in a graphical way as a *dendrogram* (or hierarchical tree, see figure 1, Pérez-Bonilla et al. [2007]).

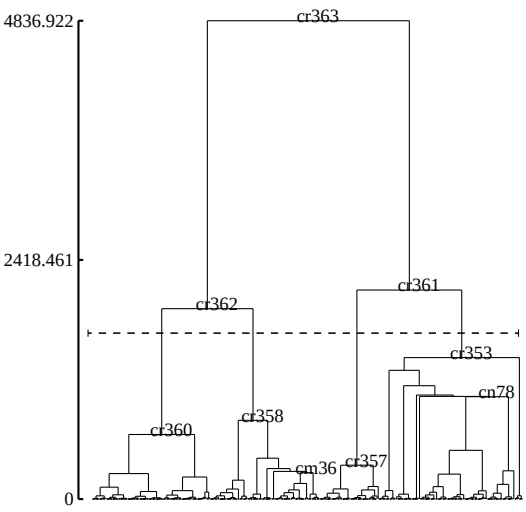


Figure 1: *Left:* Notation; *Right:* Hierarchical tree [$\tau_{Lj3,R2}^{EnW,G}$] (Pérez-Bonilla et al. [2007]) .

3 THE METHODOLOGY

CCEC takes advantage of the existence of τ, and uses the property of any binary hierarchical structure that $\mathcal{P}_{\xi+1}$ has the same classes of $\mathcal{P}_\xi$ except one, which splits in two subclasses in $\mathcal{P}_{\xi+1}$. Binary hierarchical structure will be used by *CCEC* to discover particularities of the final classes step by step also in hierarchical way. The *CCEC*, Pérez-Bonilla and Gibert [2007], allows generation of automatic conceptual interpretations of a given partition $\mathcal{P} \in \tau$. The steps to be followed are described below. The application of CCEC to the WWTP is illustrated in §5:

1. Cut the tree at highest level (make $\xi = 2$ and consider $\mathcal{P}_2 = \{C_1, C_2\}$).
2. Use *BbD* (Pérez-Bonilla and Gibert [2007]), to find (total or partial) characteristic values for numerical variables, Gibert et al. [1998].
3. Use *BbIR*, to induce a knowledge Base describing both classes.
4. For classes in $\mathcal{P}_2$, determine concepts A_1^{ξ,X_k} : "$[X_k \in I_s^k]$", $A_2^{\xi,X_k} : \neg A_1^{\xi,X_k}$ associated to C_1, C_2, by taking the intervals provided by a totally characteristic variable or the partial one with greater relative covering and $p_{sc} = 1$.
5. Go down one level in the tree, by making $\xi = \xi + 1$ and so considering $\mathcal{P}^{\xi+1}$. As said before $\mathcal{P}^{\xi+1}$ is *embedded* in $\mathcal{P}^\xi$ in such a way that there is a class of $\mathcal{P}^\xi$ splitting in two

new classes of $\mathcal{P}^{\xi+1}$, namely $C_i^{\xi+1}$ and $C_j^{\xi+1}$ and all other classes $C_q^{\xi+1}, q \neq i, j$, are common to both partitions and $C_q^{\xi+1} = C_q^{\xi} \ \forall q \neq i, j$.
Since in the previous step $C_i^{\xi+1} \cup C_j^{\xi+1}$ were conceptually separated from the rest, at this point it is only necessary to find the variables which separate (or distinguishes) $C_i^{\xi+1}$ from $C_j^{\xi+1}$, by repeating steps 2-4. Suppose $B_i^{\xi+1,X_k}$ and $B_j^{\xi+1}$ the concepts induced from $C_i^{\xi+1}$ and $C_j^{\xi+1}$, in the step $\xi + 1$.

6. Integrate the extracted knowledge of the iteration $\xi + 1$ with that of the iteration ξ, by determining the compound concepts finally associated to the elements of $\mathcal{P}_{\xi+1}$. The concepts for the classes of $\mathcal{P}_{\xi+1}$ will be: $A_q^{\xi+1,X_k} = A_q^{\xi,X_k}$, $A_i^{\xi+1,X_k} = \neg A_q^{\xi,X_k} \wedge B_i^{\xi+1,X_k}$ and $A_j^{\xi+1,X_k} = \neg A_q^{\xi,X_k} \wedge B_j^{\xi+1,X_k}$
7. Make $\xi = \xi + 1$, and return to the step 5 repeating until $\mathcal{P}_\xi = \mathcal{P}$.

4 CASE STUDY

A case study in this paper was the pilot plant, which is located in Domale-Kamnik wastewater treatment plant in Slovenia. A scheme of the pilot plant with sensors and actuators is shown in Figure 2. In the pilot plant the moving bed biofilm reactor (MBBR) technology is tested for the purpose of upgrading the whole plant for nitrification and denitrification. The pilot plant with the volume of 1125 m3 consists of two anoxic and two aerobic tanks that are filled with the plastic carriers on which the biomass develops, a fifth tank, which is a dead zone without plastic carriers and a settler. The total air flow to the both aerobic tanks can be on-line manipulated in such a way that oxygen concentration in the first aerobic tank is controlled at the desired value. The wastewater rich with nitrate is recycled with the constant flow rate from the fifth tank back to the first tank. The influent to the pilot plant is wastewater after mechanical treatment, which is pumped to the pilot plant. The inflow is kept constant to fix the hydraulic retention time. The influent flow rate can be adjusted manually to observe the plant performance at different hydraulic retention times.

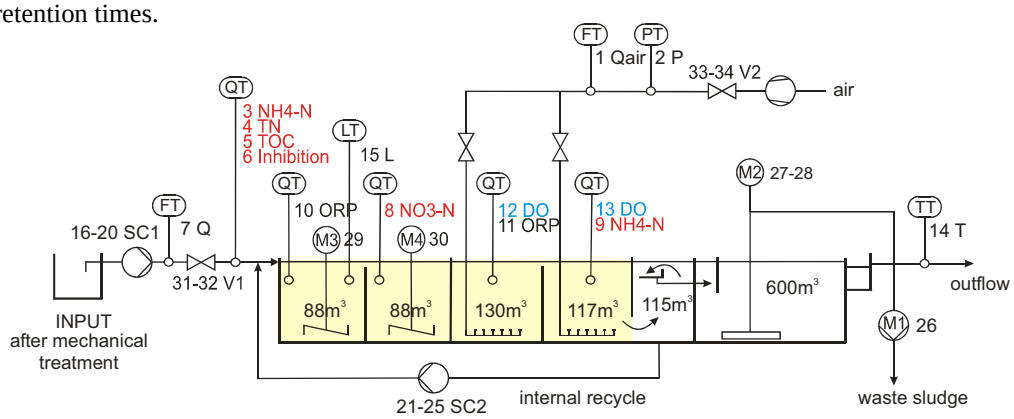


Figure 2: MBBR (Moving Bed Biofilm Reactor) pilot plant with sensors and actuators.

The database that was used in this study consists of 365 daily averaged observations that were taken from the 1st of June 2005 to the 31th of May 2006. Every observation includes measurements of the 16 variables that are relevant for the operation of the pilot plant. The variables that were measured are:

- NH4-infuent: ammonia concentration at the influent of the pilot plant(pp) (3 in the Fig. 2).
- Q-influent: wastewater influent flow rate of the pp (7 in the Fig. 2).
- TN-influent: concentration of the total nitrogen at the influent of the pp (4 in the Fig. 2).
- TOC-influent: total organic carbon concentration at the influent of the pp (5 in the Fig. 2).
- Nitritox-influent: measurement of the inhibition at the influent of the pp (6 in the Fig. 2).
- h-wastewater: height of the wastewater in the tank (no in the Fig. 2).
- O2-1aerobic: dissolved oxygen concentration in the 1st aerobic tank (3rd tank) (12-Fig. 2).
- Valve-air: openness of the air valve (between 0 - 100%), highly related with Q-air (V2 in the Fig. 2).
- Q-air: total air flow that is dosed in both aerobic tanks (1 in the Fig. 2).

- NH4-2aerobic: ammonia concentration in the second aerobic tank (9 in the Fig. 2).
- O2-2aerobic: dissolved oxygen concentration in the 2nd aerobic tank (4th tank)(13-Fig. 2).
- TN-effluent: concentration of the total nitrogen at the effluent of the pp (no in the Fig. 2).
- Temp-wastewater: temperature of the wastewater (14 in the Fig. 2).
- TOC-effluent: total organic carbon concentration at the effluent of the pp (no in the Fig. 2).
- Freq-rec: frequency of the internal recycle flow rate meter (no in the Fig. 2).
- FR1-DOTOK-20s (Hz): frequency of the motor that pums the wastewater into the pilot plant. This frequency is highly correlated with Q-inffluent.

The data base was clustered in a previous work by using clustering based on rules proposed by Gibert [1996] using the following Knowledge Base:

$KB = \{((and(>= (NH4-2aerobic)10.0)(> (TN-effluent)18.0)) \rightarrow Mmonia),$
$((and(< \quad (NH4-2aerobic)10.0)(> (TN-effluent)18.0)) \rightarrow Nitrogen)\}$

with 38 objects in class Mmonia, 80 objects in class Nitrogen and Residual class with 247 objects. The results are presented in Pérez-Bonilla et al. [2007]. The hierarchical tree of Figure 1(right) was produced and a $\mathcal{P}_4 = \{Cr_{353}, Cr_{357}, Cr_{358}, Cr_{360}\}$ is obtained. Figure 3 contains the *class panel graph*, Gibert et al. [2005] of the 16 variables regarding the partition $\mathcal{P}_4$ where the *multiple boxplot*, Tukey [1977] of variables for each class are displayed. As usual in hierarchical clustering, the final partition is the horizontal cut of the tree that maximizes the ratio between *heterogeneity* between classes with respect to *homogeneity* within classes, what guarantees the *distinguishability* of classes. The result is a 4-class partition $\mathcal{P}_4$.

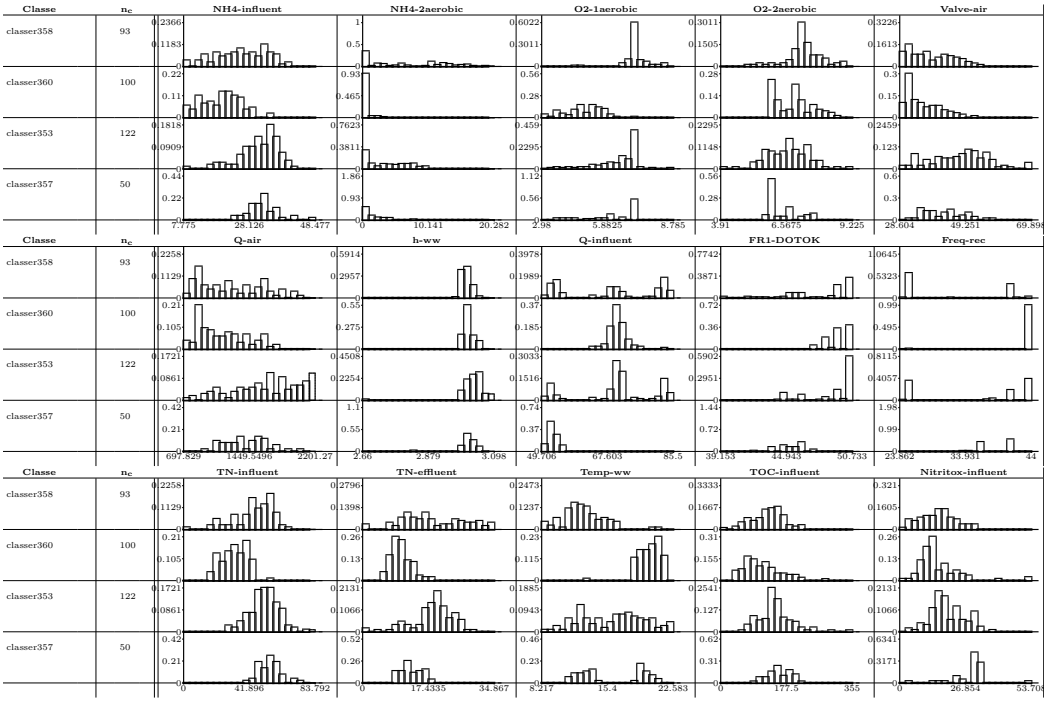


Figure 3: Class panel graph of $\mathcal{P}_4 = \{Cr_{353}, Cr_{357}, Cr_{358}, Cr_{360}\}$.

From the class panel graph, experts provided the following interpretation:

- Cr_{353}, represents the plant operation under the high load. In this case influent nitrogen concentrations are high and also influent flow rate is quite high as well. Even though the oxygen concentration in the aerobic tanks are high this can not decrease the effluent nitrogen concentrations. It means that when the plant is overloaded high effluent concentrations at the effluent of the plant can be expected.
- Cr_{357}, represents the situation when the influent flow rate is low, that is, when the hydraulic retention time of the plant is high. In this case we get quite low effluent nitrogen concentrations if of course oxygen concentration in the aerobic tank is high enough. It means when the influent flow rate to the plant is low the effluent concentrations of the plant can be obtained at the low level if the oxygen concentration in the aerobic tanks is high.

- Cr_{358}, explains the situation when the wastewater temperature is low. In this case nitrogen removal efficiency of the plant is rather low. This is so because microorganisms in the tanks don't work so intensively in cold conditions and therefore higher concentrations at the effluent of the plant can be expected.
- Cr_{360}, shows the situation when the wastewater temperature is high. In warmer conditions the microorganisms in the plant work faster, so the effluent nitrogen concentrations can be low even when the oxygen concentrations in the aerobic tanks are quite low.

5 RESULTS

Application of the methodology presented in §3 produced the following results.

The 2-class Partition. For the presented data the following 2-class partition $\mathcal{P}_2 = \{C_{361}, C_{362}\}$ is obtained. As stated in §2, I^k is built using *Boxplot based discretization (BbD)* for all variables. Next, *BbIR* is used to generate all the rules induced for $\mathcal{P}_2$, by taking the intervals provided by a totally characteristic variable or the partial one with greater relative covering and $p_{sc} = 1$. Here, TN-influent has the greatest relative covering (22,80%), see Table 1.

The following association concept can be done.

- $A^{2,TN-influent}_{Cr_{362}}$= "$x_{TN-influent,i} \in [0.0, 28.792)$" is associated with Cr_{362}
- $A^{2,TN-influent}_{Cr_{361}}$=$\neg A^{2,TN-influent}_{Cr_{362}}$= "$x_{TN-influent,i} \in [28.79, 83.79]$" associated with Cr_{361}

Or, in other words:

- Class Cr_{361}, "Not low Concentration of the total nitrogen at the influent".
- Class Cr_{362}, "Low Concentration of the total nitrogen at the influent".

The next step is to go one level down the tree.

Table 1: Summary of Knowledge Base for Cr_{361} and Cr_{361} from $P2^{EnW,G}_{Lj3,R2}$

Concep	Knowledge Base Cr_{361} (172 days) and Cr_{362} (193 days)	Cov	$CovR$
$A^{2,NH4-influent}_{C_{361}}$	$r^{NH4-influent}_{3,Cr_{361}}$: $x_{NH4-influent,i} \in (40.541, 48.477] \xrightarrow{1.0} Cr_{361}$	5	2.91%
$A^{2,O2-1aerobic}_{C_{361}}$	$r^{O2-1aerobic}_{3,Cr_{361}}$: $x_{O2-1aerobic,i} \in (8.371, 8.785] \xrightarrow{1.0} Cr_{361}$	1	0.58%
$A^{2,O2-2aerobic}_{C_{361}}$	$r^{O2-2aerobic}_{1,Cr_{361}}$: $x_{O2-2aerobic,i} \in [3.91, 4.94) \xrightarrow{1.0} Cr_{361}$	2	1.16%
$A^{2,Valve-air}_{C_{361}}$	$r^{Valve-air}_{3,Cr_{361}}$: $x_{Valve-air,i} \in (54.777, 69.898] \xrightarrow{1.0} Cr_{361}$	**28**	**16.28%**
$A^{2,Q-air}_{C_{361}}$	$r^{Q-air}_{3,Cr_{361}}$: $x_{Q-air,i} \in (2030.77, 2201.27] \xrightarrow{1.0} Cr_{361}$	27	15.70%
$A^{2,h-ww}_{C_{361}}$	$r^{h-ww}_{3,Cr_{361}}$: $x_{h-ww,i} \in (3.058, 3.098] \xrightarrow{1.0} Cr_{361}$	16	9.30%
$A^{2,Q-influent}_{C_{361}}$	$r^{Q-influent}_{1,Cr_{361}}$: $x_{Q-influent,i} \in [49.706, 50.99) \xrightarrow{1.0} Cr_{361}$	6	3.49%
$A^{2,Freq-rec}_{C_{361}}$	$r^{Freq-rec}_{3,Cr_{361}}$: $x_{Freq-rec,i} \in (43.97, 44.0] \xrightarrow{1.0} Cr_{361}$	3	1.74%
$A^{2,TN-influent}_{C_{361}}$	$r^{TN-influent}_{3,Cr_{361}}$: $x_{TN-influent,i} \in (65.25, 83.792] \xrightarrow{1.0} Cr_{361}$	9	5.23%
$A^{2,Temp-ww}_{C_{361}}$	$r^{Temp-ww}_{3,Cr_{361}}$: $x_{Temp-ww,i} \in (21.896, 22.583] \xrightarrow{1.0} Cr_{361}$	5	2.91%
$A^{2,TOC-influent}_{C_{361}}$	$r^{TOC-influent}_{3,Cr_{361}}$: $x_{TOC-influent,i} \in (290.212, 355.0] \xrightarrow{1.0} Cr_{361}$	20	11.63%
$A^{2,TOC-effluent}_{C_{361}}$	$r^{TOC-effluent}_{3,Cr_{361}}$: $x_{TOC-effluent,i} \in (44.053, 52.57] \xrightarrow{1.0} Cr_{361}$	10	5.81%
$A^{2,NH4-influent}_{C_{362}}$	$r^{NH4-influent}_{1,Cr_{362}}$: $x_{NH4-influent,i} \in [7.775, 7.972) \xrightarrow{1.0} Cr_{362}$	3	1.55%
$A^{2,NH4-2aerobic}_{C_{362}}$	$r^{NH4-2aerobic}_{3,Cr_{362}}$: $x_{NH4-2aerobic,i} \in (9.846, 20.282] \xrightarrow{1.0} Cr_{362}$	38	19.69%
$A^{2,O2-1aerobic}_{C_{362}}$	$r^{O2-1aerobic}_{1,Cr_{362}}$: $x_{O2-1aerobic,i} \in [2.98, 3.297) \xrightarrow{1.0} Cr_{362}$	4	2.07%
$A^{2,Valve-air}_{C_{362}}$	$r^{Valve-air}_{1,Cr_{362}}$: $x_{Valve-air,i} \in [28.604, 28.934) \xrightarrow{1.0} Cr_{362}$	5	2.59%
$A^{2,Q-air}_{C_{362}}$	$r^{Q-air}_{1,Cr_{362}}$: $x_{Q-air,i} \in [697.829, 739.819) \xrightarrow{1.0} Cr_{362}$	2	1.04%
$A^{2,Q-influent}_{C_{362}}$	$r^{Q-influent}_{3,Cr_{362}}$: $x_{Q-influent,i} \in (85.092, 85.5] \xrightarrow{1.0} Cr_{362}$	1	0.52%
$A^{2,FR1-DOTOK}_{C_{362}}$	$r^{FR1-DOTOK}_{1,Cr_{362}}$: $x_{FR1-DOTOK,i} \in [39.153, 42.276) \xrightarrow{1.0} Cr_{362}$	5	2.59%
$A^{2,TN-influent}_{C_{362}}$	$r^{TN-influent}_{1,Cr_{362}}$: $x_{TN-influent,i} \in [0.0, 28.792) \xrightarrow{1.0} Cr_{362}$	**44**	**22.80%**
$A^{2,TN-efluent}_{C_{362}}$	$r^{TN-effluent}_{3,Cr_{362}}$: $x_{TN-effluent,i} \in (28.933, 34.867] \xrightarrow{1.0} Cr_{362}$	14	7.25%
$A^{2,TOC-influent}_{C_{362}}$	$r^{TOC-influent}_{1,Cr_{362}}$: $x_{TOC-influent,i} \in [0.0, 63.22) \xrightarrow{1.0} Cr_{362}$	20	10.36%
$A^{2,Nitritox-influent}_{C_{362}}$	$r^{Nitritox-influent}_{1,Cr_{362}}$: $x_{Nitritox-influent,i} \in [0.0, 3.833) \xrightarrow{1.0} Cr_{362}$	4	2.07%
$A^{2,TOC-effluent}_{C_{362}}$	$r^{TOC-effluent}_{1,Cr_{362}}$: $x_{TOC-effluent,i} \in [0.0, 2.014) \xrightarrow{1.0} Cr_{362}$	1	0.52%

The 3-class partition. Take $\mathcal{P}_3 = \{Cr_{362}, Cr_{353}, Cr_{357}\}$ and first identify correspondences between the classes of $\mathcal{P}_2 = \{Cr_{361}, Cr_{362}\}$ and $\mathcal{P}_3 = \{Cr_{362}, Cr_{353}, Cr_{357}\}$. Class Cr_{361} splits into Cr_{353} (referred as $Cr_i^{\xi+1}$ in methodology), Cr_{357} (referred as $Cr_j^{\xi+1}$ in methodology) while Cr_{362} (Cr_{362} is referred as CR_q^{ξ} in methodology) remains in both $\mathcal{P}_2$ and $\mathcal{P}_3$. From the previous iteration it is already known that there is a common characteristic of both classes, Cr_{357} and Cr_{353}, (TN-influent is high) which distinguishes them from Cr_{362} (with low TN-influent). Thus, it is only required to find the separation between Cr_{353} and Cr_{357}. A similar procedure as was used in previous step for separating CR_{362} and Cr_{361} was used with all the variables regarding Cr_{353} and Cr_{357}. Here Q-influent has the greatest relative covering of 81,97%. Combining with the results of previous iteration of *CCEC* leads to the following interpretation of $\mathcal{P}_3$:

-Cr_{362} is such that, $A^{2,TN-influent}_{Cr_{362}}$= "$x_{TN-influent,i} \in [0.0, 28.792)$"
-Cr_{353} is such that, $\neg A^{2,TN-influent}_{Cr_{362}} \wedge B^{3,Q-influent}_{Cr_{353}}$
"$x_{TN-influent,i} \in [28.792, 83.792]$" $\wedge$ "$x_{Q-influent,i} \in (55.666, 85.092]$"
-Cr_{357} is such that, $\neg A^{2,TN-influent}_{Cr_{362}} \wedge \neg B^{3,Q-influent}_{Cr_{353}}$
"$x_{TN-influent,i} \in [28.792, 83.792]$" $\wedge$ "$x_{Q-influent,i} \in (49.706, 55.666]$"

Degree of certainty should be associated depending on the probabilities of the corresponding generated concepts.

The final partition. The process would continue separating the Cr_{360} and Cr_{358} of the partition $\mathcal{P}_4$ (see figure 1(right)), which are the subdivision of Cr_{362}. Here, a Temp-ww with a relative covering of 66,67%, is chosen. Similarly, the interpretation of $\mathcal{P}_4$, which is the final partition obtained with the Knowledge Base with certain rules for Cr_{360} and Cr_{358} is as follows:

– Cr_{353} is such that "$x_{TN-influent,i} \in [28.792, 83.792]$" $\wedge$ "$x_{Q-influent,i} \in (55.666, 85.092]$"
"Non low Concentration of the total nitrogen at the influent and high values of Wastewater influent flow rate".
– Cr_{357} is such that "$x_{TN-influent,i} \in [28.792, 83.792]$" $\wedge$ "$x_{Q-influent,i} \in [49.706, 55.666]$
"Non low Concentration of the total nitrogen at the influent and not high values of Wastewater influent flow rate".
– Cr_{358} is such that "$x_{TN-influent,i} \in [0.0, 28.792)$" $\wedge$ "$x_{Temp-ww,i} \in [8.472, 13.327)$"
"Low Concentration of the nitrogen and low values of water's temperature."
– Cr_{360} is such that "$x_{TN-influent,i} \in [0.0, 28.792)$" $\wedge$ "$x_{Temp-ww,i} \in [13.327, 21.896]$"
"Low Concentration of the nitrogen and non low values of temperature".

This set of concepts can, in fact, be considered as a domain model which can support later decision for the treatment that is applied to a new day, provided that a standard treatment is previously associated to every class by experts. In this association the possibility of easily interpreting the classes is critical as well as the knowledge from the experts for easily understanding the meaning of the classes. In this sense the proposed method provides simple and short rules which use to be easier to handle than those provided by other induction rules algorithms.

Evaluating Confidence and Support terms and comparing them with the class interpretation provided by the expert two modifications of the method are assessed in this work. First includes the best (the greatest relative covering) rule for every class in the description and seconds uses a closed-world assumption.

Best local concept: Here Valve-air is chosen for Cr_{361} and TN-influent for Cr_{362}, see Table 1, Q-influent is chosen for the Cr_{353} ($CovR = 81,97\%$) and FR1-DOTOK ($CovR = 16,00\%$) for Cr_{357} and Temp-ww is chosen for the Cr_{360} ($CovR = 29,00\%$) and Cr_{358} ($CovR = 66,67\%$). The final interpretation of 4 classes is the following:

– Cr_{353} is such that "$x_{Valve-air,i} \in (54.77, 69.898]$" $\wedge$ "$x_{Q-influent,i} \in (55.666, 85.092]$"
High percentages of the air valve openess and high values of Wastewater influent flow rate.
– Cr_{357} is such that "$x_{Valve-air,i} \in (54.77, 69.898]$" $\wedge$ "$x_{FR1-DOT,i} \in [42.276, 44.167)$"
High percentages of the air valve openess and low values of frequency of the influent flow rate meter.

– Cr_{358} is such that "$x_{TN-influent,i} \in [0.0, 28.792) \wedge$ "$x_{Temp-ww,i} \in [8.472, 13.327)$"
Low values of the total nitrogen concentration at the influent and low values of the wastewater temperature.
– Cr_{360} is such that "$x_{TN-influent,i} \in [0.0, 28.792) \wedge$ "$x_{Temp-ww,i} \in (20.928, 21.896]$"
Low values of the total nitrogen concentration at the influent and high values of the wastewater temperature.

Best local concept and partial Close-World Assumption: Includes the same variables as the first modification together with the complementary concepts of the other class except when the best variable for the two classes is the same in iteration. In this case the original concepts is kept (this criteria behaves better than CCEC, if does the comparison with the concept those expert).

- Cr_{353} is such that ("$x_{TN-influent,i} \in [28.792, 83.792]$" $\vee$ "$x_{Valve-air,i} \in (54.77, 69.898]$") $\wedge$
("$x_{Q-influent,i} \in (55.666, 85.092]$" $\vee$ "$x_{FR1-DOTOK,i} \in [44.167, 50.7]$")
"High-medium of the total nitrogen concentration at the influent or high percentages of the air valve openess and high values of Wastewater influent flow rate or high-medium values of the influent motor pump frequency"
- Cr_{357} is such that ("$x_{TN-influent,i} \in [28.792, 83.792]$" $\vee$ "$x_{Valve-air,i} \in (54.77, 69.898]$") $\wedge$
("$x_{Q-influent,i} \in [49.706, 55.666]$ " $\vee$ "$x_{FR1-DOTOK,i} \in [42.276, 44.167)$")
"High-medium of the total nitrogen concentration at the influent or high percentages of the air valve openess and low-medium values of Wastewater influent flow rate or low values of the influent motor pump frequency"
- Cr_{358} is such that "($x_{TN-influent,i} \in [0.0, 28.792) \vee$ " $x_{Valve-air,i} \in [28.604, 54.777]) \wedge$
"$x_{Temp-ww,i} \in [8.472, 13.327)$"
"Low of the total nitrogen concentration at the influent or low-medium percentages of the air valve openess and low values of the wastewater temperature."
- Cr_{360} is such that "$x_{TN-influent,i} \in [0.0, 28.792) \vee$ "$x_{Valve-air,i} \in [28.604, 54.777]) \wedge$
"$x_{Temp-ww,i} \in (20.928, 21.896]$"
"Low of the total nitrogen concentration at the influent or low-medium percentages of the air valve openess and high values of the wastewater temperature."

Three approaches generate different concepts for each class. Support and confidence factors together with average confidences and global support of concepts, are shows in Table 2.

Table 2: Comparison among the 3 proposals

	CCEC Best global rule		Best local rule		Best local rule+CWA global		partial	
Ruler	$Conf.$	$Supp.$	$Conf.$	$SupP.$	$Conf.$	$Supp.$	$Conf.$	$Supp.$
$\mathcal{R}_{Cr_{353}}$	45,00%	27,12%	100,0%	7,40%	39,48%	32,42 %	39,48%	33,42%
$\mathcal{R}_{Cr_{357}}$	46,94%	12,60%	100,0%	0,27%	47,52%	13,15%	47,52%	13,15%
$\mathcal{R}_{Cr_{358}}$	100,0%	1,64%	100,0%	1,64%	54,79%	16,99 %	54,87%	16,99%
$\mathcal{R}_{Cr_{360}}$	86,84%	9,04%	100,0%	0,82%	45,00%	27,12 %	82,86%	7,95%
Average	69,695%		100%		46,697%		56,18%	
Global Supp.		50,41%		10,13%		90,68%		71,51%

6 CONCLUSIONS AND FUTURE WORK

In this paper a methodology to generate automatic conceptual interpretations of a group of classes is presented. Concepts associated with classes are built taking advantage of hierarchical structure of the underlying clustering. The *Conceptual characterization by embedded conditioning CCEC*, Gibert and Pérez-Bonilla [2006], Pérez-Bonilla and Gibert [2007], is a quick and effective method that generates a conceptual model of the domain, which will be of great support to the later decision making based on a combination of *BbD* and an interactive combination of concepts upon hierarchical subdivisions of the domain. This is a preliminary proposal that has been applied with success to real data coming from a WWTP. Benefits of this proposal are specially interesting in the interpretation of partitions with a large number of classes. Automatic generation of interpretations

cover the important goal of KDD of describing the domain Fayyad [1996.]. However, in this proposal a direct connection between the generated concepts and the automatic rules generation allows direct construction of a decision model for the later class prediction. As a matter of a fact, automatic production of probabilistic or fuzzy classification rules regarding concepts provided by *CCEC* is direct, as discussed in Gibert and Pérez-Bonilla [2005]. By associating an appropriate characteristic to every class a model for how to operate the wastewater plant based on a concrete day upon a reduced number variables is obtained together with an estimation of the risk associated to that decision (which is related with the certainty of the rule). In this work three different criteria for deciding which variable is kept at every iteration are assessed. The one that gives the most similar interpretation to those expert is the Best local rule +partialCWA, which from a technical point of view also seems to represent the more equilibrated option with high values in confidence and support factors. In future work we will study how to propagate uncertainty from one iteration to the next needs to be analyzed in depth (here rules with $p_{sc} = 1$ is used). The idea is to use an approach which avoids the explicit construction of all the concepts to evaluate their relative coverage. Comparison of rules produced with *CCEC* and with other methods have been presented in Gibert et al. [2006]. Finally a modification of the methodology is required to guarantee that important variables at the output are also included in the final description. In long term proposal could be extended to keep more the one variable per iteration.

ACKNOWLEDGMENTS

We would like to thank the staff of the Domale-Kamnik WWTP for providing us with the data from the pilot plant and the Proyect TIN 2004-01368.

REFERENCES

Fayyad, U. *From Data Mining to Knowledge Discovery: An overview*. 1996.

Gibert, K. The use of symbolic information in automation of statistical treatment for ill-structured domains. *AI Communications*, 9(1):36–37, 1996.

Gibert, K., T. Aluja, and U. Cortés. Knowledge Discovery with Clustering Based on Rules. Interpreting results. LNAI v. 1510, pages 83–92. Springer, 1998.

Gibert, K., R. Nonell, J. M. Velarde, et.al. Knowledge discovery with clustering: impact of metrics and reporting phase by using klass. *Neural Network World*, 4/05:319–326, 2005.

Gibert, K. and A. Pérez-Bonilla. Taking advantage of the hierarchical structure of a clustering for automatic generation of classification interpretations. In *Fuzzy sets in learning and data mining.* , pages 524–529, Barcelona, España, septiembre 2005. EUSFLAT.

Gibert, K. and A. Pérez-Bonilla. Towards automatic generation of interpretation as a tool for modelling decisions. In *Proceedings of III International Conference on Modeling Decisions for Artificial Intelligence*, pages 515–524, Tarragona, Abril 2006. MDAI.

Gibert, K., A. Pérez-Bonilla, and G. Rodriguez. A Comparative Analysis of different techniques for class interpretation supporting tools. In *In Procs. Nov Congrs Catal d'Intelligncia Artificial.*, pp. 37–46,IOS press. Perpignan, France, noviembre 2006. CCIA06.

Gordon, A. D. Identifying genuine clusters in a classification. *Computational Statistics and Data Analysis*, pages V.18: 561–581, 1994.

Liu, B. Hsu, W. C. S. M. Y. Analyzing the subjective interestigness of association rules. *IEEE Intelligent Systems*, pages 47–55, 2000.

Pérez-Bonilla, A. and K. Gibert. Towards automatic generation of conceptual interpretation of clustering. In *Progress in Pattern recognition, Image analysis and Application.*, volume 4756 of *Lecture Notes in Computer Science*, pp. 653–663, Springer-Verlag. Valparaiso-Chile, 2007.

Pérez-Bonilla, A., K. Gibert, and D. Vrecko. Domale-kamnik wastewater treatment plant (ljubljana - slovenia). clustering and induction knowledge base. Research DR 2007/09, Dep. Estadística e Investigación Operativa. UPC, Barcelona, España, Diciembre 2007.

Tukey, J. *Exploratory Data Analysis*. Addison-Wesley, 1977.

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A neural network approach for selecting the most relevant variables for foaming in anaerobic digestion

Jordi Dalmau[a], Joaquim Comas[a], Ignasi Rodríguez-Roda[a], Eric Latrille[b] and Jean-Philippe Steyer[b]

[a]*Laboratory of Chemical and Environmental Engineering (LEQUIA), University of Girona, Faculty of Sciences, Campus Montilivi s/n, 17071 Girona, Catalonia, Spain. (E-mail: jordi@lequia.udg.cat; quim@lequia.udg.cat; ignasi@lequia.udg.cat)*

[b]*INRA, UR050, Laboratoire de Biotechnologie de l'Environnement, Avenue des Étangs, 11100 Narbonne, France. (E-mail: latrille@supagro.inra.fr;* steyer@supagro.inra.fr*)*

Abstract: Activated sludge processes are complex biological systems in which organic matter and nutrients (nitrogen and phosphorous) are removed from wastewater. One of the most common alternatives for the treatment of waste from activated sludge systems is Anaerobic Digestion (AD). AD is even a more complex biological process. One of the most important problems which can appear in anaerobic digesters is foaming. In the literature there is not a big agreement on the foaming causes. Therefore, the aim of the paper is to apply a methodology based on artificial neural networks to determine the most relevant variables for foaming diagnosis in anaerobic digestion. To perform the study real data from a pilot plant located in the LBE in Narbonne were used. Results show inflow rate, total organic carbon and carbon dioxide percentage among the relevant variables which are, according to the literature, some of the factors which influence the presence of foaming. This methodology can be valuable when selecting probes to monitor AD processes or to select the most relevant variables to diagnose the state of the process when foaming is involved.

Keywords: Anaerobic digestion; artificial neural networks; data mining; diagnosis; foaming.

1. INTRODUCTION

1.1 Activated sludge systems

Activated sludge processes are complex biological systems in which organic matter and nutrients (nitrogen and phosphorous) are removed from wastewater. The system consists of an aeration tank where oxygen is selectively supplied and it is used by the microbial consortia (i.e. biomass and/or sludge) to grow and reproduce by consuming the substrate (i.e. pollutants) present in the wastewater.

The system also includes a secondary settler in which the treated water is separated from the biomass. From the bottom of the clarifier a fraction of the activated sludge is returned to the reactor in order to maintain the biomass constant in the reactor. To prevent overgrowth of the biomass in the system, a small fraction of the sludge is wasted from the system. This fraction represents a significant cost for the activated sludge process, since further treatment is required.

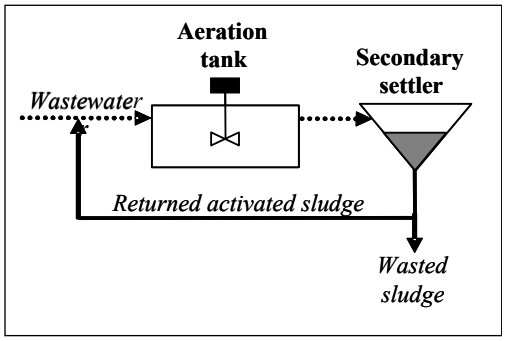


Figure 1. Activated Sludge System.

1.2 Anaerobic Digestion (AD)

The most common alternative for sludge treatment is AD. In this process the organic matter (e.g. sludge coming from activated sludge treatment) is biologically degraded in a digester in absence of oxygen. AD processes advantages are numerous since they provide a treatment for high loaded influents, low sludge production and production of energy in form of methane. AD processes are, like in the activated sludge, very complex biological systems since a huge amount of species are involved in the process.

1.3 Foaming

Within this complexity some bacteria can see its own growth promoted by certain conditions which can cause imbalances in the digester in the form of a thick foam blanket. This foaming can have many effects as: fouling of gas collection pipes, probes failure, problems with mixing devices, lose of digester efficiency… (Pagilla et al. [1997])

There is not yet a complete agreement on the parameters that favours conditions for foaming forming bacteria. Some authors state that a proper control of the feeding will prevent excessive foaming to appear (Massart et al. [2006]; Schaffer et al. [2006]). Others state that pre-treatment of the feeding is necessary to avoid foaming appearing (Barjenbruch [2003]; Elliott [2007]). Besides, some claim that the presences of some filamentous bacteria in the activated sludge system are the cause for foaming problems in the anaerobic digester (Pagilla et al. [1997]; Westlund et al. [1998]).

1.4 Artificial Neural Networks (ANN)

When dealing with complex processes which their dynamics are not well known, ANNs are a useful tool to obtain an approximated model. Hence, in the present work they will be used to select the relevant variables and, at the same time, to finally obtain a black-box model for foaming diagnosis based on the most relevant variables. When dimensionality reduction is involved other techniques such as principal component analysis are easy to apply and they do not require a target data set as ANN does, nevertheless they run the risk of being suboptimal since they take no account of the target data (Bishop, [1994]).

The objective of this paper is to select the most relevant variables to control or prevent foaming in AD processes. Likewise, this work will highlight the variables most related to foaming in anaerobic digesters. Firstly, the used methodology will be presented followed by the results presenting the most relevant variables with some discussion to finish with conclusions.

2. METHODOLOGY

The methodology is a wrapper approach with a hill-climbing elimination strategy (Kohavi and John, [1997]). The same methodology was used in Dalmau et al. [2007] in which was applied in order to find the most relevant variables for acidogenic states in anaerobic digestion. In the current work, the same approach is applied to foaming in AD which is, as commented above, a more changeling issue since there is no full agreement on its causes.

Figure 2 depicts the procedure that starts with the ten times training of the reference ANN with all the variables. Its average Root Square Mean Error (RSME) is calculated and stored as the reference error. Next, one input variable is removed and a new ANN (ANN1 in figure 1) is trained ten times without it. This last step is repeated for each input variable ending up with *n* ANNs 1 one for each removed input variable with their related average RSME 1. Whenever a relevant variable is removed, the average RSME 1 of the related ANN 1 will increase with respect to the average reference error. On the other hand, whenever a non-relevant variable is removed the RSME 1 of the related ANN1 will decrease. Therefore, the variables which RSME 1 is higher than the reference error are selected as relevant variables.

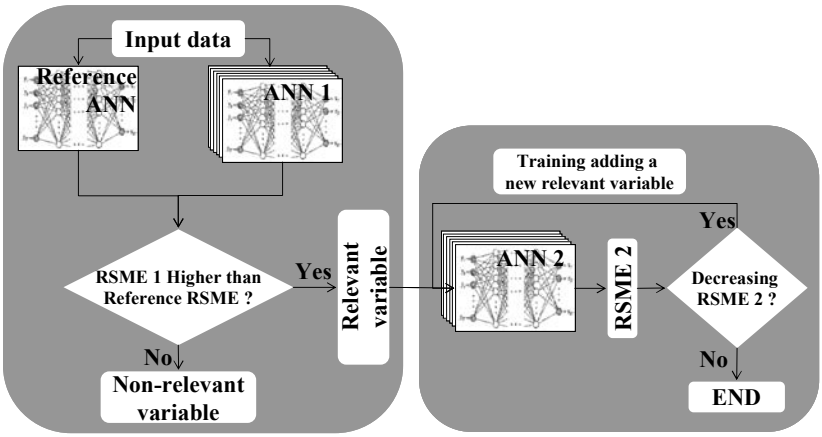


Figure 2. Methodology schema, based on Dalmau et al. [2007].

Among relevant variables the one with the higher RSME 1 is selected first and a new ANN (ANN 2 this time) is trained ten times again using it as the only input. If the related average RSME (RSME 2) is higher than the average reference error no improvement is found, so the variable with the second higher average RSME 1 is selected and a new ANN 2 is trained (ten times as well) with both variables, and again, its average RSME 2 is compared with the reference. This iterative process is repeated until an average RSME 2 lower than the average reference RSME is obtained.

2.1 Artificial Neural Network toolbox

A home made neural network toolbox for static models for use in MATLAB 5.3 or higher was used in this study. In all the ANN architectures two layers were used: a hidden layer of neurones with sigmoid transfer functions and an output layer with linear transfer functions for outputs. The initialisation method was performed using the Nguyen-Widrow algorithm option (Nguyen [1990]) was selected as initialisation method. This algorithm initialises the weights with random values, later selecting their probability distributions to make all neurones active for the expected data ranges. Automatic data scaling and weights conversion are included as well. Bayesian regularisation approach is used to prevent over-

fitting. Finally, training process is based on a Levenberg-Marquardt algorithm (Marquardt [1963] and Levenberg [1944]).

3. RESULTS

3.1 Data

Experimental data used were obtained from a pilot plant from LBE of the INRA, France. Among all variables a first selection was done based on the common variables which are available in real plants. Some others were not selected for instance, temperatures since it is usually constant so it will be difficult to extract information from its profile. Table 1 presents the input variables involved in this study. As output, foaming appearance in the digester was used, based on the heuristic knowledge provided by the experts. It was noticed that when foaming appeared in the digester high variations of the gas flow rate and pressure coincided due to the slug release of gas bubbles trapped inside the foam. In order to have a suitable foaming index between 0 and 1 a fuzzy system was used. It is important to point that even though foaming can be determined likewise, this is an approach to study variables influence or relation.

Table 1. Input variables.

Inflow rate	**Volatile fatty acids concentration in digester**	**Total organic carbon in digester**	**pH in the inflow rate**	**pH in the digester**	**Carbon dioxide percentage in the gas phase**	**Methane percentage in the gas phase**
qIn	vfaDig	tocsDig	phIn	phDig	co2Gas	ch4Gas

Besides, the objective of the study is to determine the most relevant variables with the purpose of either saving in instrumentation or know which variables are most important to control to prevent foaming in the anaerobic digester.

3.2 Reference ANN and ANNs 1

The average RMSE of the reference ANN and the seven ANNs 1 are presented in Table 2.

Table 2. Reference ANN and ANNs1 average RSME

Reference	**qIn**	**vfaDig**	**tocsDig**	**phIn**	**phDig**	**co2Gas**	**ch4Gas**
0.11700	0.11934	0.11377	0.12764	0.12654	0.11328	0.11796	0.12179

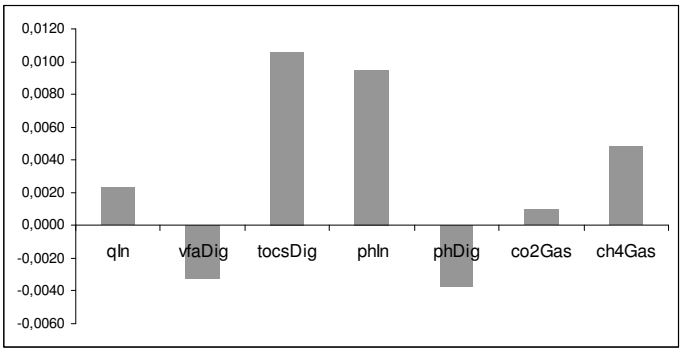


Figure 3. Average RSME minus average reference RSME for each variable.

Figure 3 shows the differences between RSME of each variable and the reference RSME. As depicted 5 variables have RSME higher than the reference. Therefore, the selected as relevant variables were qIn; tocsDig; phIn; co2Gas and ch4Gas.

3.3 ANNs 2

To train ANNs 2 first tocsDig was selected. Being its average RSME 2 higher than the average reference RSME the second most relevant variable was added and ANN 2 was trained again. Table 3 summarizes the average RSME 2 for each ANN.

Table 3. Average RSME 2 for the relevant variables.

	tocsDig	**tocsDig; phIn**	**tocsDig; phIn;ch4Gas**	**tocsDig;phIn; ch4Gas;qIn**	**tocsDig;phIn; ch4Gas;qIn; co2Gas**
RSME 2	0.12666	0.12256	0.12141	0.12138	0.11944

In all cases RSME 2 were higher than the reference error. However, a test t performed comparing the averages between each RSME 2 and the RSME of the reference ANN revealed that in the last ANN 2 (last column in Table 3), the average RSME 2 was not significantly different than the reference RSME. This means that even though the average RSME 2 is not lower than the reference, by removing two variables from the whole input, similar results can be obtained. This can be important in terms of saving sensors and in terms of monitoring relevance.

The relevance of gas-related variables (i.e. co2Gas and ch4Gas) can be due to the approach taken to determine foaming (i.e. the fuzzy system used). So, taking a look to the other variables, precisely tocsDig and qIn, the results can be related to some statements present in the literature. In Massart et al. [2006] it is stated that inconsistent feeding in the digester is one of the causes for foaming. Feeding is related to Organic Loading Rate (OLR), which is related to the inflow rate (qIn) and the amount of sludge feed to the digester (Metcalf [2003]) related at the same time to the organic matter present in the digester (tocsDig). Finally, as shown, addition of co2Gas to the relevant variables reduces RSME 2. According to Zhao [2004] high CO_2 production is representative of poor digestion and may lead to foaming.

4. CONCLUSION

The presented methodology has been helpful to select the relevant variables related to foaming in anaerobic digesters. From seven variables two have shown to be non-relevant associated to foaming. This is indeed an indication of the complexity of this microbiology related problem and it shows that a lot of variables are more or less related to it. Nevertheless, when diagnosing or controlling the conditions to prevent foaming the selected variables are helpful indicators of what it is worth to control to avoid excessive foaming.

ACKNOWLEDGEMENTS

Authors wish to thank Dr Cristian Trelea from the Institut National Agronomique Paris-Grignon for developing and providing us with the Neural Network toolbox. The work was also partially supported by the Spanish Ministry of Education and Sciences project DPI2006-15707-C02-01.

REFERENCES

Barjenbruch M. and O. Kopplow, Enzymatic, mechanical and thermal pre-treatment of surplus sludge, *Advances in Environmental Research, 7,* 715-720, 2003.

Bishop C. M., Neural networks and their applications, *Review of Scientific Instruments 65* (6), 1803-1832, 1994.

Dalmau J., J. Comas, I. Rodríguez-Roda, E. Latrille and J.P. Seteyer, Using artificial neural networks for selecting relevant information in anaerobic digestion, Paper presented at 11th IWA Specialitst Conference on Anaerobic Digestion (AD11), Brisbane, Australia, September 23-27, 2007.

Elliot A. and T. Mahmood, Pretreatment technologies for advancing anaerobic digestion of pulp and paper biotreatment residues. *Water Research 41*, 4273-4286, 2007.

Kohavi R. and G. H. John, Wrappers for feature subset selection, *Artificial Intelligence 97*, 273-324, 1997.

Levenberg K., A method for the solution of certain non-linear problems in least squares, *Quarterly Journal of Applied Mathematics 2*(2), 164–168, 1944.

Marquardt D.W., An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society of Industrial and Applied Mathematics 11*(2), 431–441, 1963.

Massart N., R. Bates, B. Corning and G. Neun, Design and operational considerations to avoid excessive anaerobic digester foaming, Paper presented at the Annual Water Environment Federation Technical Exhibition and Conference, Dallas, USA, October 21-25, 2006.

Metcalf and Eddy, *Wastewater engineering: Treatment and reuse*, McGraw-Hill, New York, USA, 2003.

Nguyen D. and B. Widrow, Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights, Paper presented at International Joint Conference on Neural Networks, San Diego, California, USA, June 17-21, 3, 21-26, 1990.
Pagilla K. R., K. C. Craney and W. H. Kido, Causes and effects of foaming in anaerobic sludge digesters, *Water Science and Technology*, 36(6-7), 463-470, 1997.

Pagilla K. R., Kraney K.C. and W. H. Kido, Causes and effects of foaming in anaerobic sludge digesters, *Water Science and Technology,* 36(6-7), 463-470, 1997.

Schafer P., W. Uhte and G. Newman, Improved anaerobic digestion feed systems and concepts, *Proceedings of the Annual Water Environment Federation Technical Exhibition and Conference,* Dallas, USA, 2006.

Westlund A. D., E. Hagland and M. Rothman, Foaming in anaerobic digesters caused by *Microthrix parvicella, Water Science and Technology,* 37(4-5), 51-55, 1998.

Zhao H. and T. Viraraghavan, Analisys of the performance of an anaerobic digestion system at the Regina wastewater treatment plant, *Bioresource Technology, 95,* 301-307, 2004.

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Trajectories' Mining Between Subprocesses in a Wastewater Treatment Plant

K. Gibert [a], **G. Rodríguez Silva**[a] **and I. Rodríguez Roda** [b]

[a]*Universitat Politècnica de Catalunya, Campus Nord Ed.C5, Jordi Girona 1-3 , Barcelona (karina.gibert@upc.edu, gustavo.rodriguez-silva@upc.edu)*

[b]*Universitat de Girona, Campus Montilivi, Girona (ignasi.rodriguezroda@udg.edu)*

Abstract: In this work the Trajectories' Mining between subprocesses is presented in regards to the dynamics of a Wastewater Treatment Plant (WWTP). This kind of Environmental Systems involves a high complexity inherent to its own characteristics. Intelligent Environmental Decision Support Systems (IEDSS) can improve the decision making process, assisting decision makers in the evaluation of alternatives and improving management and control of Environmental Systems in general and for this particular application of WWTP. Our line of work is based on the development of methodologies of Artificial Intelligence and Statistics to solve problems of Knowledge Discovery of Data as is shown by Fayyad [1996]. In this work, knowledge extraction is approached with a methodology named Clustering Based on Rules by States (ClBRxE) formally presented by Gibert and Rodríguez [2007]. Basically, consists of analyzing a process that can be divided in $\mathcal{S} = \{e_1 \dots e_E\}$ states or subprocesses. After dividing variables based on the subprocesses to which they refer, knowledge discovered is integrated from each subprocess into a unique model of global operation of the phenomenon. Once the data of the WWTP has been analyzed, it was empirically verified that probability of a transition between two consecutive states depends on not only on the previous state but on the whole sequence of previous states.

Keywords: clustering; rules; states; trajectories; wastewater

1 INTRODUCTION

To correctly treat wastewater different operations and unique processes are required. A mixture of physics, chemical and biological agents is needed to form the diagram of the process of each wastewater station. The global process always follows a logical sequence of treatment divided in different stages that can varied according to the structure and objectives of the plant as is shown in Metcalf and Eddy [2003]. The Data presented in this work has been previously clustered and interpreted by Gibert and Roda [2000]. In our current work knowledge discovery is approached with Clustering Based on Rules by States (ClBRxE). Once the data has been analyzed, probability of transition between states is verified to model global operation of the phenomenon.

2 APPLICATION DOMAIN AND PREVIOUS WORK

A sample of 396 observations taken from September the first of 1995 to September the 30th of 1996 from a Catalan WWTP is used. The plant is described daily with measures taken in the following stages of the depuration process: Input(E), Primary Settler(D), Bioreactor (B) and Output (S). In the database 4 different stages of the depuration process are identified. A selection of 25 variables was considered the most relevant by the opinion of an expert and indicated which variables correspond to each stage of the depuration process as is shown in table 1. This data has been previously clustered using Clustering Based on Rules *(ClBR)* with a knowledge base that collects the legal limits of certain physics and biological parameters that classify the quality of wastewater at the plant's exit. The rules base $\mathcal{R}$ is in figure 1, Right.

Table 1: Variables used in the Clustering.

INPUT (**E**)	SETTLER (**D**)	BIOREACTOR (**B**)	OUTPUT (**S**)
1.Inflow (Q-E)	1.Hydrogen Potential (pH-D)	1.Inflow (QB-B)	1.Hydrogen Potential (pH-S)
2.Iron pre-treatment (FE-E)	2.Suspended Solids (SS-D)	2.Recirculated Inflow (QR-G)	2.Suspended Solids (SS-S)
3.Hydrogen Potential (pH-E)	3.Volatile Suspended Solids (SSV-D)	3.Purge Inflow (QP-G)	3.Volatile Suspended Solids (SSV-S)
4.Suspended Solids (SS-E)	4.Chemical Organic Matter (DQO-D)	4.Aireation (QA-G)	4.Chemical Organic Matter (DQO-S)
5.Volatile Suspended Solids (SSV-E)	5.Biodegradable Organic Matter (DBO-D)	5.Index 30 at the Biological Reactor (V30-B)	5.Biodegradable Organic Matter (DBO-S)
6.Chemical Organic Matter (DQO-E)		6.Mixed Liquor Suspended Solids (MLSS-B)	
7.Biodegradable Organic Matter (DBO-E)		7.Mixed Liquor Volatile Suspended Solids (MLVSS-B)	
		8.Mean Cell Residence Time (MCRT-B)	

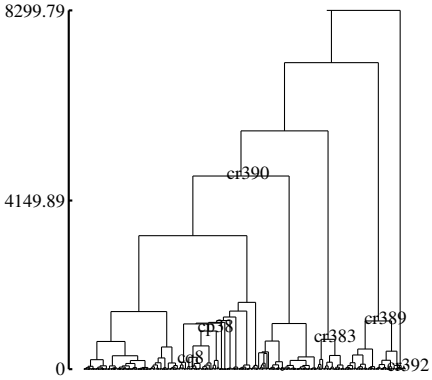
ClBR is implemented on *KLASS* software as is shown by Gibert and Roda [2000], this can take advantage of a prior knowledge base to bias classes construction to improve interpretability. Details on this technique and comparison with other clustering techniques are shown in Gibert [1996]. As usual in hierarchical clustering, the final partition is the horizontal cut of the dendrogram $\tau^{En,G}_{Gi2,R1}$ that maximizes the ratio between *heterogeneity* between classes with respect to *homogeneity* within classes, what guaranties the *distinguishability* between classes. Following this classical criterion and validated by the experts, a partition in four clusters was performed $P_4 = \{c383, c389, c390, c392\}$. Experts of this environmental process associate concepts to each resulting class, where statistics from each class and Class Panel Graph, figure 2 helped us in the interpretation as is shown by Gibert and Roda [2000].

Class C_{392}: This class contains many variables with specific values of this class, most of them at low values. Verifying the corresponding dates assigned to C_{392}, allowed the experts identify this class as the storm days. To face the storm, the decision is to protect the system minimizing inflow (Q-E), closing input valves and maintaining bioreactor microorganism's trough minimizing purged flow (QP-G). This produces an increment of biomass in the reactor and justifies the high values of (MCRT-B).

Class C_{389}: This class has input inflow, recirculated inflow and purge inflow at non-low values, cellular age is not high, on the contrary of C_{392}. These days the plant works very well even reduces ammonium (NH4-D y NKT-D lows) in spite of not being a plant specifically designed for this. In addition, the reactor is working a full yield (QB-B high) and the degradation is good (MCRT-B low), as a result the water is so clean. The experts identify this class with a profile of excellent working of the plant.

Class C_{383}: In this class, the input inflow, bioreactor inflow, recirculated inflow and purge inflow are lows, cellular age is non-high, ammonium reduction is lower, output flow is no so clean (DBO-S non-low). Contains days where water is very dirty (SS-E, SSV-E, DQO-E and DBO-E highs) as in organic material as in suspended solids. According to the experts exists a shock of load (organic material) of solids in the input of the process, as indicates the volatile suspended solids inside bioreactor are very lows (MLVSS-B).

Class C_{390}: The assigned days in this class present some punctual problem, the inflow is not so dirty (as indicated by the low or medium values of SS-E, SSV-E, DQO-E and DBO-E), but suspended solids in biological reactor are higher than expected (MLVSS-B), indicating a non-optimal depuration.



$\mathcal{R} = \{r_1$: If $(SS - S) > 20$ &(DBO-S)$> 35 \longrightarrow$S
r_2 : If $(SS - S) > 20$ &(DBO-S)$< 35 \longrightarrow$P
r_3 : If $(SS - S) < 20$ &(DBO-S)$> 35 \longrightarrow$Q$\}$

S: states for (abnormal operation in general)
P: states for (failure in suspended solids treatment)
Q: states for (failure in organic matter treatment)

Figure 1: Left: General tree of the Clustering Based on Rules. $[\tau^{En,G}_{Gi2,R1}]$. Right: Rules Base $\mathcal{R}$

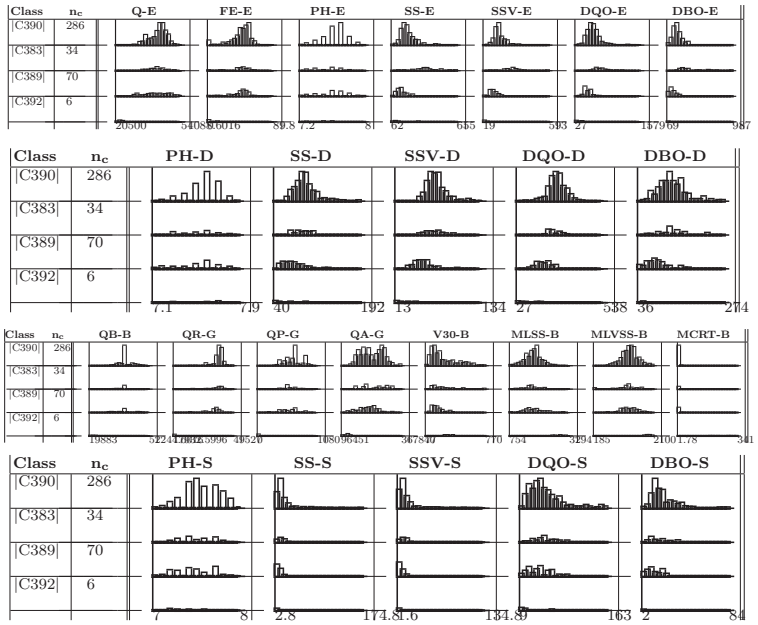


Figure 2: Class Panel Graph of all variables versus Global Clustering in 4 classes [$\tau^{En,G}_{Gi2,R1}$]

3 CLUSTERING BASED ON RULES BY STATES (CLBRXE)

Given an environmental domain in which a process is taking place in such a way that it can be divided in $\mathcal{S} = \{e_1 \dots e_E\}$ states or subprocesses, with $\mathcal{I} = \{i_1 \dots i_n\}$ observations described by $X_1 \dots X_K$ variables and given an apriori knowledge base $\mathcal{R}$, which can be partial, containing logic rules as described by Gibert [1996], our proposal is:

1. Calculate a rules-induced partition on $\mathcal{I}$ by evaluating $\mathcal{R}$ over $X_1 \dots X_K$ ($\mathcal{P}_\mathcal{R}$).
2. Divide the variables based on the subprocess to which they refer (let $X^e_1, \dots X^e_{Ke}$ be the set of variables referring to the subprocess e).
3. For each subprocess $\mathcal{S} = \{e_1 \dots e_E\}$:
   a Do $\mathcal{I}De = \{i \in \mathcal{I} : x^e_{i1} = x^e_{i2} = \dots = x^e_{iKe} = *\}$, where * indicates missing value.
   b Do a *ClBR*, as described by Gibert [1996], over $\mathcal{I} \setminus \mathcal{I}De$ with variables referring to state e $X^e_1, \dots, X^e_{1Ke}$ but using $\mathcal{P}_\mathcal{R}$ as a rules-induced partition for all the states of the process. Elements of $\mathcal{I}De$ can not be used in state e since they miss useful information.
   c Analyze the resulting dendrogram τ^e, perform an horizontal cut to obtain $\mathcal{P}^*_e$ and associate concepts to each resulting class $\mathcal{C} \in \mathcal{P}^*_e$.
   d Construct $\mathcal{P}_e = \mathcal{P}^*_e \cup \{\mathcal{I}De\}$
4. Construct cross tables $\mathcal{P}^e \times \mathcal{P}^{e+1}$, applying required delay, if any, between observations and analyze most probable trajectories of individual through classes of sequential states. Crossing can be done at different stages of the process by considering that the individual that is a certain stage delay some time till the next stage. Thus, cell C, C' of the cross table contains the number of elements that $x_{(i-d)K_e} = C$ and $x_{iK_{e+1}} = C'$ instead of $x_{iKe} = C \wedge x_{iK_{e+1}} = C'$ where d is the delay of the individual from state e to $e+1$. Delays between steps are to be provided by experts. This information allows wider analysis of the dynamics of the process and it is being studied if this approach can give richer models than the instantaneous conjoint global analysis of all phases performed previously.
5. Identify trajectories related to informations flow between the $\mathcal{S} = \{e_1 \dots e_E\}$ subprocess.
6. Select the most typical trajectories and interpret.

4 KNOWLEDGE DISCOVERY IN WWTP WITH CLBRXE

4.1 ClBR Local to Each State In our methodology $\mathcal{R}$ is evaluated on $\mathcal{I}$ and a partition is induced $\mathcal{P}_{\mathcal{R}} = \{S, P, Q, Residual\}$ S contains 11 elements of $\mathcal{I}$ that satisfy r_1, P contains 40 elements of $\mathcal{I}$ that satisfy r_2, Q contains 10 elements of $\mathcal{I}$ that satisfy r_3 and Residual contains 335 elements of $\mathcal{I}$ that do not satisfy any rule or many at the same time which are contradictory. Each stage is treated separately then the relationship between subprocesses is analyzed. The number of classes is identified a posteriori for each subprocess upon the resulting dendrogram.

Input Subprocess. 7 variables are available: Q-E,FE-E,PH-E,SS-E,SSV-E,DQO-E,DBO-E. All prototypes will be calculated for each rules-induced partition in the previous step. We will manually construct an extended residual class that includes those prototypes in addition to residual objects. Finally, a clustering without rules of the extended residual class is performed with the following criteria: Normalized Euclidean Squared Metric, Linkage Wards Criteria, since all the variables are numerical in this application, which is equivalent to perform a *ClBR* of the Input variables using a knowledge base that includes variables from another state. This is a technical solution to be used in the meanwhile, statistical software *KLASS* is modified to include this case. The tree τ_E of clustering process is shown in figure.

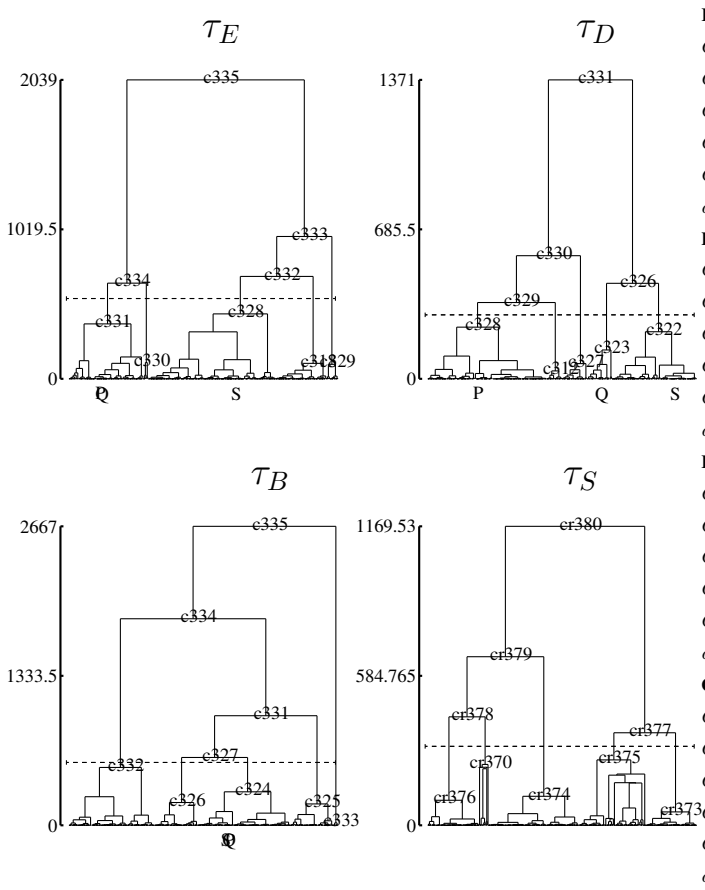
Settler Subprocess. 5 variables are available: PH-D, SS-D, SSV-D, DQO-D, DBO-D. As in the Input stage,in Settler is proceed on the same way. The tree τ_D of clustering process is shown in figure3.

Bioreactor Subprocess. 8 variables are available: QB-B, QR-G, QP-G, QA-G,V30-B, MLSS-B, MLVSS-B, MCRT-B. As in the Input stage,in Bioreactor is proceed on the same way. The tree τ_B of clustering process is shown in figure 3.

Output Subprocess. 5 variables are available: PH-S, SS-S, SSV-S, DQO-S, DBO-S. The rules provided by the expert affected the variables that are used in this stage, therefore we directly realized *ClBR* with the following criteria: Normalized Euclidean Squared Metric, Linkage Wards Criteria. The tree τ_S of clustering process is shown in figure 3.

Analyzing dendrograms τ_E, τ_D, τ_B and τ_S, the corresponding partitions of 5 classes and Class Panel Graphs were performed, see figure 4. With these elements some concepts can be related to each class as are shown in the right side of figure 3.

4.2 Trajectories Analysis Once the typical situations in each stage are identified, we proceed to study the transit of the wastewater from one stage to the next. The starting point will be the crossing between the partitions of two consecutive stages. On this specific application daily averages were used while wastewater keeps only few hours at each stage. According to this is not required a delay between stages before performing the crossing. To study transitions between states, frequencies of one stage conditioned by the previous stage are shown on table 2, for example $\mathcal{P}^S \mid \mathcal{P}^B$ shows that poorest quality water hardly ever comes out of the plant when the Bioreactor is functioning correctly ($P_{C324_B, C370_S} = 0.0632$). In figure 5, Transition Diagram derived from tables of conditional frequencies of subsequent states by ($\mathcal{P}^D \mid \mathcal{P}^E$), ($\mathcal{P}^B \mid \mathcal{P}^D$) and ($\mathcal{P}^S \mid \mathcal{P}^B$) is built. Classes identified for each subprocess are aligned from highest quality water at the top to most polluted water at the bottom. Transition probability is visualized with the thickness of arrows. Transitions between two consecutive states, for instance, from Input Subprocess (E) to Settler Subprocess can be arranged in a transition matrix, which provides transition probabilities between any pair of states. For example, P_{ED} provides probabilities from Input Subprocess to Settler Subprocess. The sequence of classes of wastewater through the states of the process is denoted as *trajectory*: $T : \{C_E, C_D, C_B, C_S\}$, where $\mathcal{C}_E \in \mathcal{P}_E$, $\mathcal{C}_D \in \mathcal{P}_D$, $\mathcal{C}_B \in \mathcal{P}_B$ and $\mathcal{C}_S \in \mathcal{P}_S$. For instance, for $T_1 : \{C331_E, C328_D, C324_B, C375_S\}$ water flows from $C331_E$ at the Input to $C328_D$ at the Primary Settler to $C324_B$ at the Bioreactor and $C375_S$ at the Output, what means that wastewater has normal levels of pollutants at the input, still keeps average levels of solids in the primary settler, underload wastewater with average waste purged flow rate in the bioreactor and exits with average quality. The 12-IX-1995 was one of the days that wastewater followed this path. Finally, a probability should be associated to every trajectory to identify the most typical patterns and characteristical dynamics observed in WWTP.



Input Subprocess $\mathcal{P}^E$:
$C329_E$ low addition of FE-E
$C318_E$ underflow WW(Q-E)
$C328_E$ underload WW and high inflow rate
$C331_E$ WW at normal level
$C330_E$ overload WW and high inflow rate
d_1 unknown

Primary Settler Subprocess $\mathcal{P}^D$:
$C327_D$ low level of SS-D
$C312_D$ low pH-D
$C328_D$ average levels of SS-D, SSV-D
$C322_D$ high levels of SS-D, DQO-D
$C323_D$ high levels of SSV-D, DBO-D
d_2 unknown

Bioreactor Subprocess $\mathcal{P}^B$:
$C326_B$ underload WW, high aeration
$C332_B$ underload WW, low aeration
$C324_B$ underload WW, average waste purged
$C325_B$ overload WW, high aeration
$C333_B$ bioreactor closed
d_3 unknown

Output Subprocess $\mathcal{P}^S$:
$C373_S$ high quality water
$C374_S$ high quality, high pH-S
$C375_S$ average quality water
$C376_S$ poor quality, high DQO-S
$C370_S$ poorest quality water
d_4 unknown

Figure 3: Left: τ_E, τ_D, τ_B and τ_S partition in 5 Classes. Right:Concepts related to each class

4.3 Trajectories Probability Induction Association of probability to the trajectories can be approached from different perspectives and in this paper two of them are presented.

Markov Process Asumption: The probability of a given trajectory would be the product of corresponding transition probabilities if wastewater treatment was a Markov process and satisfies the property *"for any given time instant (say t_n) the future behaviour, for instance the value of $X_{t_{n+1}}$, is totally independent of its history, i.e., the values of $X_{t_{n-1}}$,$X_{t_{n-2}}$ and so on. It only depends on the state occupied at the current time instant t_n, given by the value of X_{t_n}."*, Holger [2002]. Consider trajectory $T : (C_E, C_D, C_B, C_S)$. Assuming that a WWTP is a Markov Process:

$$p(T) = P_E \cdot P_{ED} \cdot P_{DB} \cdot P_{BS} \tag{1}$$

Frequentist Non Parametric Assumption: In this approach, probabilities at a given stage will be conditioned by the whole path previously followed by the water. $P(T)$ is then calculated as:

$$p'(T) = p(_{C_E}) \cdot p(_{C_D|C_E}) \cdot p(_{C_B|C_D \wedge C_E}) \cdot p(_{C_S|C_B \wedge C_D \wedge C_E}) \tag{2}$$

This is equivalent to direct frequentist estimation of the trajectories along the whole sample.

As the WWTP process really was a Markov process and satisfies the non-memory property announced above, both approaches would produce same results. However, in this application it is empirically verified that WWTP is not a Markov process and results differ on both approaches. As an example consider $T_1 : (C331_E, C328_D, C324_B, C375_S)$

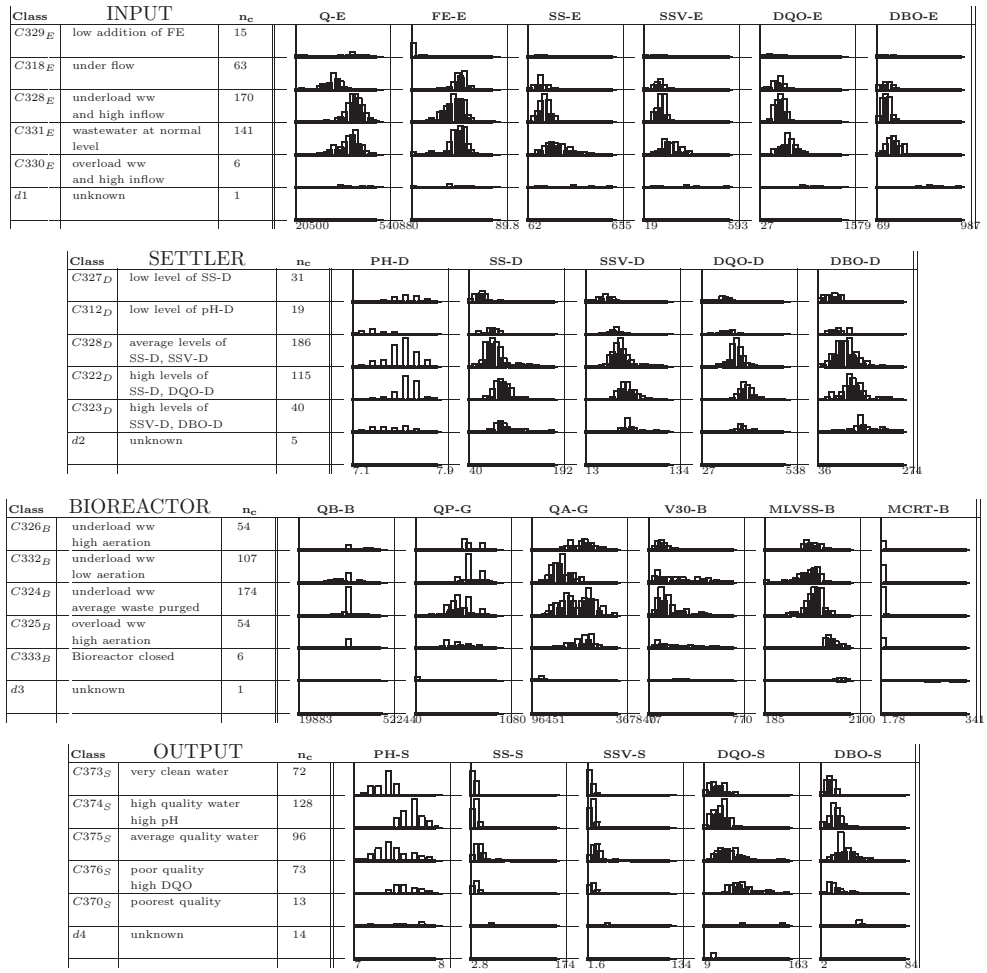


Figure 4: Class Panel Graph of ($\mathcal{P}^E$), ($\mathcal{P}^D$), ($\mathcal{P}^B$) and ($\mathcal{P}^S$)

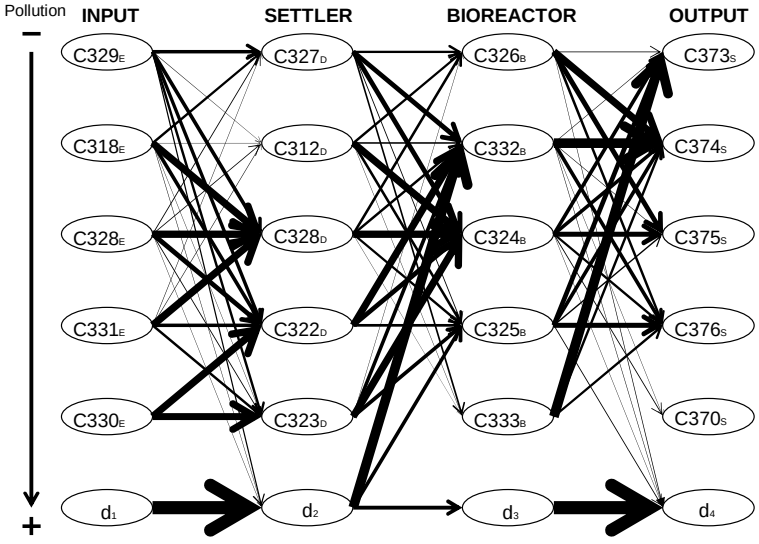


Figure 5: Transition Diagram between states

Table 2: Frequencies of state conditioned by $\mathcal{P}^D \mid \mathcal{P}^E$, $\mathcal{P}^B \mid \mathcal{P}^D$ and $\mathcal{P}^S \mid \mathcal{P}^B$.

P_{ED}	$C327$	$C312$	$C328$	$C322$	$C323$	$d2$	tot.	useful
$C329$	0.2667	0.0667	0.2	0.2667	0.1333	0.0667	1	15
$C318$	0.1429	0.0635	0.4762	0.2698	0.0317	0.0159	1	63
$C328$	0.0941	0.0706	0.4588	0.3059	0.0588	0.0118	1	170
$C331$	0.0142	0.0142	0.5319	0.2766	0.1631	0	1	141
$C330$	0	0	0	0.5	0.5	0	1	6
$d1$	0	0	0	0	0	1	1	1
margin	0.0783	0.048	0.4697	0.2904	0.101	0.0126	1	
useful	31	19	186	115	40	5		396

P_{DB}	$C326$	$C332$	$C324$	$C325$	$C333$	$d3$	tot.	useful
$C327$	0.1613	0.3226	0.2903	0.0968	0.129	0	1	31
$C312$	0.1579	0.1053	0.5263	0.1579	0.0526	0	1	19
$C328$	0.1882	0.1882	0.4946	0.1237	0.0054	0	1	186
$C322$	0.087	0.4174	0.3565	0.1391	0	0	1	115
$C323$	0.025	0.225	0.525	0.225	0	0	1	40
$d2$	0	0.6	0.2	0	0	0.2	1	5
margin	0.1364	0.2702	0.4394	0.1364	0.0152	0.0025	1	
useful	54	107	174	54	6	1		396

P_{BS}	$C373$	$C374$	$C375$	$C376$	$C370$	$d4$	tot.	useful
$C326$	0.1111	0.4074	0.3148	0.1296	0	0.037	1	54
$C332$	0.0654	0.5701	0.0654	0.243	0.0187	0.0374	1	107
$C324$	0.2126	0.1839	0.3851	0.1322	0.0632	0.023	1	174
$C325$	0.3333	0.2222	0.0926	0.2963	0	0.0556	1	54
$C333$	0.6667	0.1667	0	0.1667	0	0	1	6
$d3$	0	0	0	0	0	1	1	1
margin	0.1818	0.3232	0.2424	0.1843	0.0328	0.0354	1	
useful	72	128	96	73	13	14		396

- Under Markov assumption $P(T_1)$ is calculated as:
  $P(T_1) = P_{C331_E} \cdot P_{C331_E C328_D} \cdot P_{C328_D C324_B} \cdot P_{C324_B C375_S}$

  $P(T_1) = p(_{C331_E}) \cdot \frac{p(_{C328_D, C331_E})}{p(_{C331_E})} \cdot \frac{p(_{C324_B, C328_D})}{p(_{C328_D})} \cdot \frac{p(_{C375_S, C324_B})}{p(_{C324_B})}$

  Looking at table 2:

  $P_E = p(_{C331_E}) = \frac{card(_{C331_E})}{\mathbf{n}} = \frac{141}{396} = 0.3561$

  $P_{ED} = p(_{C328_D | C331_E}) = \frac{p(_{C328_D, C331_E})}{p(_{C331_E})} = \frac{\frac{card(C331_E \cap C328_D)}{\mathbf{n}}}{\frac{card(C331_E)}{\mathbf{n}}} = \frac{\frac{75}{396}}{\frac{141}{396}} = 0.5319$

  $P_{DB} = p(_{C324_B | C328_D}) = \frac{p(_{C324_B, C328_D})}{p(_{C328_D})} = \frac{\frac{card(C328_D \cap C324_B)}{\mathbf{n}}}{\frac{card(C328_D)}{\mathbf{n}}} = \frac{\frac{92}{396}}{\frac{186}{396}} = 0.4946$

  $P_{BS} = p(_{C375_S | C324_B}) = \frac{p(_{C375_S, C324_B})}{p(_{C324_B})} = \frac{\frac{card(C324_B \cap C375_S)}{\mathbf{n}}}{\frac{card(C324_B)}{\mathbf{n}}} = \frac{\frac{67}{396}}{\frac{174}{396}} = 0.3850$

  Substituting in eq. 1 $P(T_1) = 0.3561 \cdot 0.5319 \cdot 0.4946 \cdot 0.3850 = 0.0361$

- Under Frequentist approach $P(T_1)$ is calculated as:
  $P'(T_1) = \frac{\frac{card(C331_E \cap C328_D \cap C324_B \cap C375_S)}{\mathbf{n}}}{\frac{card(C331_E \cap C328_D \cap C324_B)}{\mathbf{n}}}$

  This is equivalent to use the following expression:
  $P'(T_1) = p(_{C331_E}) \cdot p(_{C328_D | C331_E}) \cdot p(_{C324_B | C328_D \wedge C331_E}) \cdot p(_{C375_S | C324_B \wedge C328_D \wedge C331_E})$

  $p(_{C331_E}) = \frac{card(_{C331_E})}{\mathbf{n}} == \frac{141}{396} = 0.3561$

  $p(_{C328_D | C331_E}) = \frac{p(_{C328_D \wedge C331_E})}{p(_{C331_E})} = \frac{\frac{card(C331_E \cap C328_D)}{\mathbf{n}}}{\frac{card(C331_E)}{\mathbf{n}}} = \frac{\frac{75}{396}}{\frac{141}{396}} = 0.5319$

  $p(_{C324_B | C328_D \wedge C331_E}) = \frac{p(_{C324_B \wedge C328_D \wedge C331_E})}{p(_{C328_D \wedge C331_E})} = \frac{\frac{card(C331_E \cap C328_D \cap C324_B)}{\mathbf{n}}}{\frac{card(C331_E \cap C328_D)}{\mathbf{n}}} = \frac{\frac{61}{396}}{\frac{75}{396}} = 0.8133$

  $p(_{C375_S | C324_B \wedge C328_D \wedge C331_E}) = \frac{p(_{C375_S \wedge C324_B \wedge C328_D \wedge C331_E})}{p(_{C324_B \wedge C328_D \wedge C331_E})} = \frac{\frac{card(C331_E \cap C328_D \cap C324_B \cap C375_S)}{\mathbf{n}}}{\frac{card(C331_E \cap C328_D \cap C324_B)}{\mathbf{n}}}$

  $= \frac{\frac{45}{396}}{\frac{61}{396}} = 0.7377$

  Substituting in eq. 2 $P'(T_1) = 0.3561 \cdot 0.5319 \cdot 0.8133 \cdot 0.7377 = 0.1136$

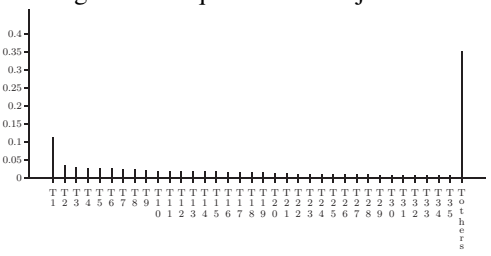
In general, $p(T) \neq p'(T)$ what means that Markov assumption is not hold in WWTP because the probability of a transition between two consecutive states depends not only on the previous state but on the whole sequence of previous states followed by wastewater. So, the most typical trajectories will be identified using the frequentist approach.

4.4 Typical Trajectories In the presented case study, 139 different trajectories are observed (from the 1296 possibilities of different trajectories). The frequencies of those 139 trajectories are shown in figure 6, and for this particular application trajectories with $p > 0.025$ are selected and interpreted by the experts :

$\tau_3 = \{C331_E, C323_D, C324_B, C375_S\}$	normal levels of pollutants at the input, high levels of SSV-D and DBO-D in the primary settler, underload WW with average waste purged flow rate and exits with average quality
$\tau_1 = \{C331_E, C328_D, C324_B, C375_S\}$	normal levels of pollutants at the input, still keeps average levels of solids in the primary settler, underload WW with average waste purged flow rate and exits with average quality
$\tau_7 = \{C328_E, C322_D, C332_B, C374_S\}$	underload WW and high inflow rate at the input, high levels of SS-D and DQO-D in the primary settler, underload WW with low aeration in the bioreactor and exits with high quality and high pH-S
$\tau_5 = \{C328_E, C322_D, C324_B, C370_S\}$	underload WW and high inflow rate, high levels of SS-D and DQO-D in the primary settler, underload WW with average waste purged flow rate and exits with poorest quality
$\tau_4 = \{C328_E, C328_D, C324_B, C373_S\}$	underload WW and high inflow rate, average levels of solids in the primary settler, underload WW with average waste purged flow rate and exits with high quality
$\tau_2 = \{C328_E, C328_D, C326_B, C374_S\}$	underload WW and high inflow rate, average levels of solids in the primary settler, underload WW with high aeration in the bioreactor and exits with high quality and high pH-S
$\tau_6 = \{C328_E, C328_D, C326_B, C375_S\}$	underload WW and high inflow rate, average levels of solids in the primary settler, underload WW with high aeration in the bioreactor and exits with average quality
$\tau_8 = \{C318_E, C328_D, C332_B, C374_S\}$	underflow WW at the input, average levels of solids in the primary settler, underload WW with low aeration in the bioreactor and exits with high quality and high pH-S

5 CONCLUSIONS

Figure 6: Frequencies of Trajectories



In this work ClBRxE was used to identify typical situations in the states of the WWTP. From the results, a transition diagram displays the more frequent transitions along the process allowing to face the analysis of trajectories as a whole. To identify the more typical trajectories, probabilities associated to each one is required. From this work it has been shown that Markov process is not a valid model for WWTP and direct frequentist estimation is more suitable for this purpose. So identification of more typical trajectories has been done using this criteria. The more probable trajectories with probability between 0.025 and 0.1136 have been selected identifying different processes in the WWTP. Currently, a methodology for automatically produce the interpretation of selected trajectories is being developed.

ACKNOWLEDGMENTS

This research has been partially financed by TIN 2004-01368.

REFERENCES

Fayyad, U. e. a. *Advances in Knowledge Discovery and Data Mining*, chapter From Data Mining to Knowledge Discovery: An overview. AAAI/MIT Press, 1996.

Gibert, K. The use of symbolic information in automation of statistical treatment for ill-structured domains. *AI Communications*, 9(1):36–37, 1996.

Gibert, K. and I. Roda. Identifying characteristic situations in wastewater treatment plants. In *Workshop in Binding Environmental Sciences and Artificial Intelligence*, volume 1, pages 1–9, August 2000. ECAI.

Gibert, K. and G. Rodríguez. Knowledge discovery in a wastewater treatment plant with clustering based on rules by states. *Frontiers in Artificial Intelligence and Applications*, 163, 2007.

Holger, H. *Interactive Markov chains: and the quest for quantified quality*. Springer Berlin, 2002. Lecture notes in computer science ; 2428.

Metcalf and Eddy. *Wastewater engineering treatament. Disposal and reuse.* McGraw-Hill, 2003. 4th Ed. revised by George Tchobanoglous, Franklin L. Burton NY.US.

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Identifying State Variables in Multivariate Hydrological Time Series Using Time Series Knowledge Mining

O. Gronz [a b], **M. Casper**[a], **P. Gemmar**[b]

[a] *Department of Physical Geography, University of Trier, Behringstrasse, 54286 Trier, Germany (gronz@uni-trier.de, casper@uni-trier.de)*

[b] *University of Applied Sciences (FH) Trier, Institute for Innovative Informatics Applications i3A, Schneidershof, 54293 Trier, Germany (o.gronz@fh-trier.de, p.gemmar@fh-trier.de)*

Abstract: The challenge of modeling rainfall-runoff processes is to define a suitable functional relationship between input variables representing precipitation measurements and the output runoff. Depending on the system's state, the amount of precipitation resulting in direct flow varies strongly. Thus, state variables indicating the system's actual state can enhance the accuracy of rainfall-runoff models significantly. Fuzzy models of Takagi-Sugeno-type are one of the effective approaches to simulate rainfall-runoff processes regarding the different system states. The design of a fuzzy rainfall-runoff model consists of two essential steps: the definition of the structure (input quantities, state variables, rules) and the identification of parameters in the conclusion. The latter one can be solved automatically using data-driven techniques in an optimal way concerning root mean square deviation. However, to solve the task of defining the structure, expert knowledge is mandatory to identify those time series that can effectively be used in a fuzzy model. This expert knowledge is not always available and not necessarily complete or correct. To identify effective state variables semi-automatically, the method Time Series Knowledge Mining (TSKM) has been used. TSKM discovers patterns representing the temporal concepts of duration, coincidence and order. Especially the patterns representing coincidence are valuable as the temporal concept of coincidence is used in fuzzy premises, too: the values of several state variables are evaluated simultaneously to determine the system's state. TSKM was applied to identify state variables with data from a 7 km^2 catchment in the northern Black Forest in Germany. From a set of more than 100 time series that were measured resulting in a huge set of possible state variable configurations, two soil moisture time series were identified. The fuzzy models generated using these two state variables were more efficient than all other models previously generated. Additionally, periods of snowfall and snowmelt could be reliably identified.

Keywords: Flood Prediction; Fuzzy Model; Time Series Knowledge Mining; Process Identification

1 INTRODUCTION

Especially in small catchments, the reaction on precipitation varies strongly depending on the catchment's actual state. In dry situations after long periods without precipitation, only a small part of rainfall will result in direct flow. In contrast, after periods with plenty of precipitation, a large part of an additional intensive rainfall will result almost completely in direct flow, indicating a high runoff coefficient. To identify the state of the catchment, different physical values can be measured. Additionally, these measurements allow for the identification of runoff processes and support the modeling of a catchment. Besides the usual values like discharge, precipitation, air temperature, radiation etc., soil moisture measurements can enhance the determination of the system's state significantly. All those measurements result in a large set of available time series.

Unfortunately, the examination of these time series is time consuming and challenging. On the other hand, the knowledge resulting from thorough examination is mandatory to build efficient models. A possible approach to model the rainfall-runoff processes is the usage of fuzzy logic. But for the creation of an efficient fuzzy model, expert knowledge is required in most cases: Only if a good combination of state variables is used in the premises to represent the system state, an efficient fuzzy model can be developed. Thus, there arise some disadvantages: gaining expert knowledge is time consuming; furthermore, this knowledge is not necessarily complete or correct; for each new catchment, the examination needs to be performed again. To solve this problem, a method needs to be applied that examines the various time series automatically and helps thereby identifying state variables.

2 FUZZY RAINFALL-RUNOFF MODELS

Catchments are complex systems. Modeling these systems using analytical approaches is challenging as most processes involved in discharge generation are hidden, not measurable, nonlinear or too complex. Thus, the effort in developing or computing a conceptual model is large. But elementary dependencies between the different variables precipitation and runoff are known and can be described: Depending on the system's state, a certain amount of precipitation will result in discharge. Additionally, historical time series with measurements of different values are available. These preconditions facilitate the usage of a fuzzy model (e. g. Cox [1994]). In fuzzy systems of Tagkagi-Sugeno type (Takagi and Sugeno [1985]), the description of the system state in the premise is separated from the quantitative influence of the variables used in the consequent. Depending on the system's state, these variables, e. g. precipitation, are differently weighted.

Figure 1 shows a possible design of a fuzzy model. The system uses m state variables in the premise like shown in part a). In this example, a soil moisture time series and a temperature time series is used. For each state variable, different fuzzy sets are defined, assigning each value from the range of values a degree of membership by membership function μ. A linguistic term can be applied to each fuzzy set, e.g. *dry*, *wet*, *medium* or *warm*. Overall, n rules are defined, following the general form IF ... AND ... THEN .... In the premise (IF part), different combinations of the individual fuzzy sets for each state variable are used to determine the state of the system (part b). For it, the degree of membership $\mu_{i,j}, 1 \leq i \leq n, 1 \leq j \leq m$ for each crisp value of the state variables is determined and all degrees are combined by calculating the product (part c) resulting in $\mu_i, 1 \leq i \leq n$. Furthermore, the result of each consequent is calculated. In our example, we assume a fixed base flow Q_B. The precipitation time series is convolved with a previously determined unit hydrograph, scaled by factor p and added to Q_B. The overall result Q_t is a linear combination of the degree of fulfillment of each rule and the individual consequent results $Q_{i,t}, 1 \leq i \leq t$ (part d).

Of course, this approach is extremely simplified. Processes like snowfall or snowmelt are not included. Nevertheless, it showed promising results like shown in Casper et al. [2007]. In designing such a model, one question arises: which configuration of state variables results in the most efficient models? Only soil moisture time series or accumulated precipitation or a combination of both of them? If n different time series are observed within a catchment, $2^n - 1$ different state variable configurations can be used as there are 2^n different subsets.

Optimizing fuzzy models can be subdivided into two different parts: the definition of the structure and identification of parameters. The latter one can automatically be solved data-driven in an optimal way (concerning root mean square deviation, Gemmar et al. [2006]). For the first step, the definition of the structure, expert knowledge is required (e.g. Cox [1994]). Once the state variables have been identified, different approaches can be applied to optimize structure like shown by Vernieuwe et al. [2003]. In our former models, state variables have been identified using expert knowledge resulting in a large set of rules (Casper et al. [2007]). In the following, we will present an alternative method.

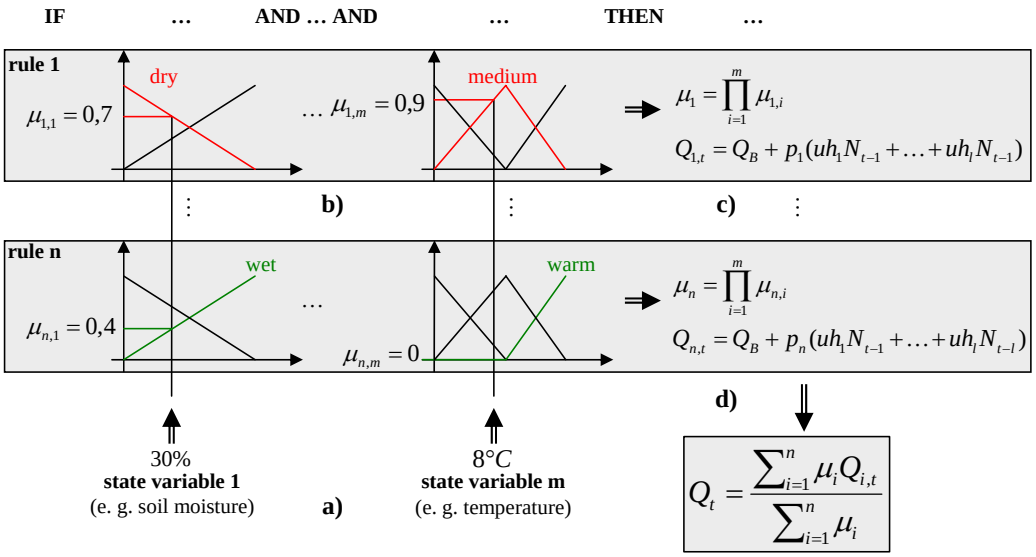


Figure 1: Possible structure of a rainfall-runoff fuzzy model of Takagi-Sugeno-type

3 Time Series Knowledge Mining

The purpose of the method Time Series Knowledge Mining (TSKM) as proposed by Mörchen [2006] is to extract temporal knowledge from multivariate time series: Using TSKM, patterns can be found that help to understand the underlying process and therefore allow modeling this process easier. The patterns are described using Time Series Knowledge Representation (TSKR), a hierarchically structured pattern language. The constructs of TSKR are named using terms from musicology, as these terms describe the underlying temporal concept vividly: Tones represent the concept of duration, overlapping parts of Tones, the Chords, stand for coincidence and partial order of Chords is represented by Phrases. The patterns found by TSKM are new, useful, understandable to humans, more compact and more abstract than the original time series. The discovery of TSKR patterns using TSKM is an interactive and iterative process consisting of five steps: pre-processing, defining Aspects, finding Tones, finding Chords and finding Phrases (Figure 2).

The goal of the first step, the pre-processing, is to remove systematic and random errors like noise, outliers, drift etc. As this is a problem regularly occurring in knowledge mining with plenty of different available approaches, there is no need for introducing new methods. Furthermore, the choice of the most suitable method to remove errors is highly application dependent and thus Mörchen [2006] does not suggest specific methods. A suitable well-known method can be chosen after examining the specific characteristics of a time series and the included errors.

In the second step, the dimensionality of the d-dimensional input space is reduced by grouping the multivariate time series to semantic blocks by selecting k subsets of the dimensions $[1, \ldots, d]$. These subsets are called *Aspects*. A descriptive, unique label is assigned to each Aspect allowing for an intuitive interpretation. Multivariate Aspects can profit from further processing like reducing the dimensionality by using e.g. PCA or calculating other derived time series.

Afterwards, in the third step, *Tones* are mined, representing the persistent occurrence of a state. Each Tone pattern contains a unique symbol, a descriptive label allowing for easy interpretation and a characteristic function indicating the occurrence of a state in an Aspect on a given time interval. The definition of the characteristic function is not restricted; various different types are possible, e.g. value-based functions, trend-based, shape-based etc. Using this characteristic function, the numerical Aspect is transformed into a symbolic interval series indicating the occurrence of Tones at each time point. Usually, this symbolic interval series will often be shortly interrupted e.g. due to noise. These gaps will split the potentially longer interval into two parts. To remove the gaps, a filter can be applied resulting in a symbolic time interval series of marginally interrupted

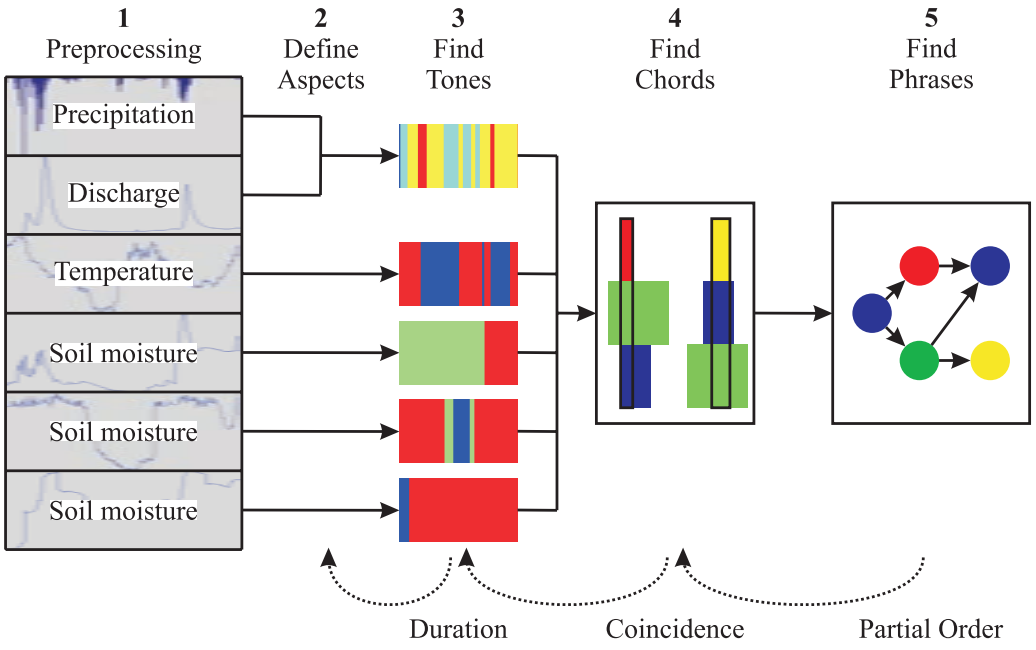


Figure 2: The five steps of Time Series Knowledge Mining (modified from Mörchen [2006])

occurrences. The filter can be parameterized by defining the maximum absolute length of a single gap and the amount of all gaps to be filled to the complete length of the time series. Finally, short Tones can be filtered, too.

Chord patterns also consist of a unique symbol, a label and a characteristic function indicating the simultaneous occurrence of Tones. A partial order is defined to allow for sub- and super-Chord relations between Chords. E.g. the Chord D consisting of Tones a and b is a sub-Chord of the Chord E consisting of Tones a, b and c - E is super-Chord of D. The hierarchy resulting from this relations can be visualized in a diagram and can be useful as we will see in section 4. To allow for tolerance concerning small differences of the support, margin-closed Chords are defined with respect to a threshold α indicating the allowed deviation of the ratio of the compared supports. Hence, the algorithm to mine Chords uses the following parameters: a threshold for the minimum duration of a Chord, the minimum and maximum size of a Chord and the threshold α for mining margin-closed Chords.

Finally, *Phrases* are mined. A Phrase pattern consists of a unique symbol, a descriptive label and a characteristic function indicating the subsequent occurrence of Chords according to a partial order on a given time interval. Again, a partial order is defined to allow for sub- and super-Phrase relations. The algorithm to mine Phrases uses the following parameters (extract): the minimum support of a Phrase and the minimum length of the paths in a Phrase.

4 Application of Time Series Knowledge Mining

In the following, the application of TSKM with data from a 7 km^2 catchment in the Northern Black Forest in Germany is described. The data contains various time series representing discharge, precipitation, air and soil temperature, soil moisture, radiation etc. in 1 h samples covering 2 years. Within this period, an extreme flood with a return period of 150 years was captured. The data has been examined thoroughly in different studies before (e.g. Casper [2002]).

To allow for unique and compact notation, the following nomenclature is used: Numeric, univariate time series are represented by vectors. E.g. $\vec{n}$ denotes a vector containing precipitation measurements. v_t describes the tth discrete sample of $\vec{v}$ corresponding to time point t. The convolution of two vectors $\vec{a}$ and $\vec{b}$ is denoted by $\vec{a} * \vec{b}$.

4.1 Pre-Processing

Although the database containing the different time series stores a quality flag for each value, some errors like outliers, drift or noise do remain. To identify and - if possible - to remove those errors, different techniques were applied. The effectiveness of the methods differed for the different physical values. For the soil moisture time series, outliers were identified reliably by applying the 2-sigma rule, indicating a value as an outlier if the difference to the mean value of the time series is bigger than the standard deviation times two (e.g. Runkler [2000]). For the runoff time series, this method failed as it indicates rare flood peaks as outliers. Another method is to plot the time series in a suitable way and use the human eye for outlier detection. For all time series, single outliers were replaced by linear interpolation of neighboring values. For some soil moisture time series, drift effects were visible, possibly due to changes in the soil matrix surrounding the TDR probes caused by water or animals. This was identified comparing the absolute values of different obviously saturated situations. Usually, these values will increase. Filters to remove noise were not applied, as tolerant algorithms are used in the following steps of TSKM.

4.2 Defining Aspects

For the following processing steps, time series representing discharge, precipitation, air temperature and soil moisture were selected from the available hydrometeorological data. The time series containing discharge measurements of the catchment's main outlet was used as Aspect *discharge*. The two available precipitation time series measured at two different sites within the catchment were combined by calculating the mean values element-by-element and used as Aspect *precipitation 1 h*. As the air temperature was measured at the same two sites, the two time series were combined similarly and used as Aspect *temperature*. Soil moisture was measured at four different sites in different depths resulting in 18 different time series. For each of the four sites, one probe was selected manually by an expert. As most of the probes of one site showed similar behavior, the probe showing least noise and drift was selected. Furthermore, the selected probe has to represent the typical behavior of the site from a hydrologic point of view. Thus, four Aspects were defined: *soil moisture 1* to *soil moisture 4*.

Besides those 7 Aspects, new Aspects were defined using derived times series. The precipitation time series contains samples representing the accumulated precipitation during one hour. Usually, those values will differ in consecutive samples resulting in a time series without persistent values. Even in periods with plenty of precipitation, the single samples can strongly vary. Hence, searching for persistent states within this time series will fail. Then again, the amount of precipitation within a certain time appears to be a good state variable as the system's state will change with the amount of precipitation. Unfortunately, the most appropriate duration of such a period is unknown. Regarding all these assumptions, new Aspects were derived to estimate the optimal size and to damp the high frequency properties. New Aspects *precipitation 2 h*, *precipitation 4 h*, *precipitation 8 h*, *precipitation 16 h* ... *precipitation 128 h* were derived by convolving the original time series: e.g. *precipitation 2 h* is $\vec{n} * (1\ 1)^T$, *precipitation 4 h* is $\vec{n} * (1\ 1\ 1\ 1)^T$ etc.

In the premises of the fuzzy system, the amount of precipitation resulting in discharge is scaled depending on the system state. This parameter, which might be compared to the runoff coefficient of the corresponding time point, is not yet incorporated in the previously defined Aspects. The exclusive usage of the amount of discharge is not adequate as a great amount of discharge resulting from a great amount of precipitation does not indicate a critical system state. On the other hand, flood waters resulting from a small amount of precipitation only occur in critical system states. Thus, a time series containing the runoff coefficient of each hour would be beneficial, but is not available. To approximate a comparable time series, the system approach of the fuzzy system is used in a modified version. Besides the amount of precipitation resulting in discharge, the constant base flow is scaled by the same factor:

$$Q_t = Q_B + Q_B \cdot p_t + p_t \cdot F_t, \tag{1}$$

where $\vec{F} = \vec{n} * \vec{uh}$. This approach can be transformed as follows:

$$p_t = \frac{Q_t - Q_B}{F_t}. \tag{2}$$

Using this approach, a time series containing the continuous values of p_t can be calculated. For the application of TSKM to determine state variables, this time series is useful as it indicates the proper value of p_t in each time step. In the following, this time series is used as Aspect *discharge disposition*.

4.3 Finding Tones

In the third step of TSKM, Tones are mined. Each Tone represents a persistent symbolic state of an Aspect. As described in section 3, various different types of characteristic functions can be defined indicating the occurrence of a Tone in a certain time interval. In the premises of fuzzy systems, the value of a sample is used to determine the degree of membership for each fuzzy set. Thus, value-based characteristic functions are defined for the 15 Aspects described in section 4.2.

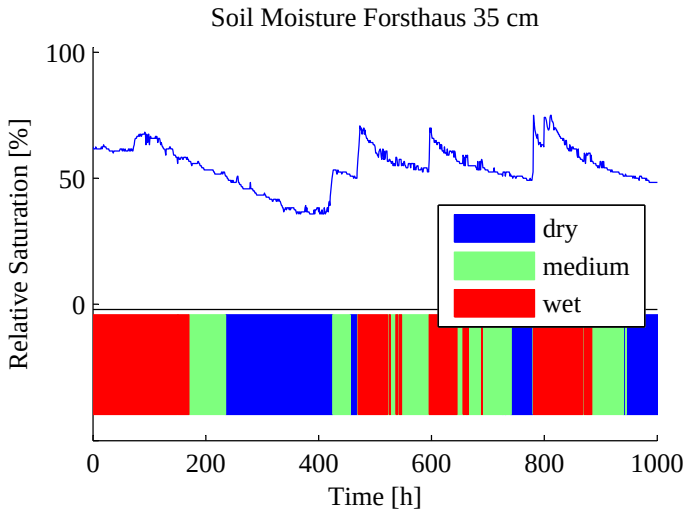


Figure 3: Transforming the numerical soil moisture time series to value-based Tones

For each of the different precipitation Aspects, three Tones were defined with labels *little*, *medium* and *plenty* to allow for an intuitive interpretation. The bin for each Tone has been determined manually in a way resulting in many Tones of type *little*, a couple of type *medium* and few of type *plenty*. The Aspect *discharge* has been subdivided into three bins similarly with labels *low*, *medium* and *high*. For *temperature*, only two Tones *cold* and *warm* were defined using the freezing point as threshold. The range of *discharge disposition* has been subdivided into 5 equidistant bins to allow for a more detailed resolution. To determine suitable bins for the soil moisture Aspects, the PERSIST algorithm has been used (Mörchen and Ultsch [2005]), which performs discretization observing the temporal order of samples. The algorithm produces bins that are similar to bins defined manually by an expert but result in fewer and more persistent occurrences of Tones. Unfortunately, this algorithm did not produce reasonable bins for the other Aspects. An example of descretization for Aspect *soil moisture 1* (site Forsthaus, depth 35cm) is shown in figure 3.

The symbolic interval series of Tones shown at the bottom of figure 3 in red, green and blue still contains many short Tones, e.g. at $t \approx 550$ and $t \approx 700$. These short Tones may be due to noise and do not represent a separated symbolic state and should be removed as a large set of short Tones will result in a huge set of Chords. Thus, this symbolic interval series is filtered resulting in a set of marginally interrupted occurrences. The first parameter of the filter, the maximum absolute length of any interruption, has been determined using the value of time to peak of the catchment. The second parameter, the maximum relative total length of all interruptions, has been limited to 10 %. The result of filter application is shown in figure 4 in the middle. As a last step, remaining short Tones are removed as shown at the bottom.

4.4 Finding Chords

In the next step, Chords are mined in the symbolic interval sequence containing Tones from the previous step. The algorithm that mines Chords uses different parameters. The first two parameters limit the size of a Chord. A Chord should at least consist of three Tones: besides Tones of Aspect *discharge* or *discharge disposition*, two further Tones should be contained as at least two time series should be used to represent system state accurately. The maximum size of a Chord is limited by the number of Aspects, that is to say 15. Again, the value of time to peak has been used to limit the minimum duration of a Chord. The threshold that limits the maximum deviation of support is 10 %. The flag whether to mine closed Chords has been set to *false* as the hierarchy of sub- and super-Chords will be used later to reduce the number of state variables. Using theses parameters, 2492 Chords were found.

4.5 Finding Phrases

To identify state variables, Chords are useful as they represent the simultaneous occurrence of Tones. For this procedure, Phrases are not necessary. However, they can be used to identify hydrologic processes like snowfall and melt within the multivariate time series.

For it, the set of Aspects has been reduced to *discharge*, *precipitation*, *temperature* and *discharge disposition*. The set of symbolic interval sequences for those Aspects has been reused without modifications. From the resulting set of Chords, specific ones were labeled manually: e.g. the Chord consisting of Tones *discharge is low*, *temperature is low* and *precipitation is medium* has been labeled *potential snowfall*. Using these Chords, Phrases were mined that indicate the different sequences of snowfall and following flood waters without precipitation; snowfall, a break were nothing happens (cold, no precipitation, low discharge) and subsequent flood water; snowfall and following precipitation during snowmelt etc. Those Phrases and especially the interval series indicating their occurrence are helpful in identifying flood waters influenced by snow as they are not included in our approach yet.

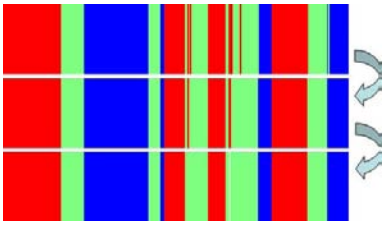


Figure 4: Filtering the symbolic interval series of Tones: value-based Tones at the top, marginally interrupted occurrences in the middle and after filtering short Tones at the bottom

4.6 From Chord to state variable

There is one critical point in the usage of TSKM to identify state variables: The occurrence of Chords does not imply that the included time series are adequate state variables; Chords state that the Tones of which they are made up occur simultaneously. Thus, the applicability of the included time series to represent system state has to be derived manually. For this, we propose the following approach: The set of Chords should not be limited by any parameter in the first four steps even if a huge set of Chords is produced. From this set of Chords, all Chords are automatically discarded that do not contain the Tone *discharge disposition is extreme* resulting in a small set of Chords that can be inspected manually. These Chords and their hierarchy are inspected manually by an expert and a Chord is selected that contains promising potential state variables besides the *discharge disposition*. Again, the initial huge set of Chords is inspected discarding automatically all Chords that are not made up of the *discharge disposition* and the identified potential state variables. The remaining set of Chords is again inspected manually by an expert. If the Chords cover the complete range of system states and low Tones of potential state variables coincide with low *discharge disposition* and high Tones of potential state variables coincide with high *discharge disposition*, the potential state variables can be used to generate a fuzzy model. If this fuzzy model shows poor results or if the set of Chords remaining after the last step contains conflicting Chords, the search has to be performed again, selecting different potential state variables.

5 RESULTS AND OUTLOOK

The set of possible state variable configurations resulting from the 15 Aspects defined in section 4.2 contains $32,767$ different configurations. From this large set, two soil moisture time series were identified as applicable state variables using TSKM and the approach described in section 4.6. The Nash-Sutcliffe efficiency of the generated fuzzy models using these two state variables is 0.38. During the identification, an expert assisted the knowledge mining process but no further explicit expert knowledge was used. In former approaches, the state variables were selected manually by an expert (Casper et al. [2007]). The Nash-Sutcliffe efficiency of the resulting model was -0.25. The expert knowledge was gathered during several years of intensive work within the catchment and two additional years of modeling. Comparing the two different approaches, the application of TSKM enhances the model's efficiency significantly. Furthermore, the amount of time that needs to be spent in gathering knowledge is decreased considerably.

The efficiencies mentioned above seam to be low compared to other rainfall-runoff models of other catchments but the small test catchment in the Black Forest shows extreme dynamics which complicates modeling. Additionally, the approach used in the fuzzy model does not yet include snowfall and snowmelt.

Up to now, TSKM has only been used with data of a single, small catchment. For evaluation purposes, the application of TSKM in other catchments is necessary. This evaluation includes the influence of parameters used in the different steps of TSKM, their general applicability and the approach to extract state variables from a set of Chords.

TSKM allows for mining temporal patterns within multivariate time series automatically. The interpretation of these patterns to identify state variables is a manual step. Thus, a huge set of possible state variable configurations can be mined but the identified state variables configuration is not necessarily the most efficient one.

REFERENCES

Casper, M. *Die Identifikation hydrologischer Prozesse im Einzugsgebiet des Dürreychbaches (Nordschwarzwald)*. PhD thesis, Universität Karlsruhe, Germany, 2002.

Casper, M., P. Gemmar, O. Gronz, M. Johst, and M. Stüber. Fuzzy logic-based rainfall-runoff modelling using soil moisture measurements to represent system state. *Hydrological Sciences Journal*, 52(3):478–490, 2007.

Cox, E. *The Fuzzy Systems Handbook*. Academic Press, Inc., 1994.

Gemmar, P., O. Gronz, and M. Stüber. Effiziente Erstellung und praktischer Einsatz von NA-Modellen mittels Fuzzy-Logik und automatisierter Entwicklungsverfahren. In Casper, M. and Herbst, M., editors, *Forum für Hydrologie und Wasserbewirtschaftung*, volume 16.06, pages 27–40, Trier, Germany, 2006.

Mörchen, F. *Time Series Knowledge Mining*. PhD thesis, Philipps-Universität Marburg, Germany, 2006.

Mörchen, F. and A. Ultsch. Optimizing time series discretization for knowledge discovery. In *KDD '05: Proceeding of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 660–665, New York, NY, USA, 2005. ACM.

Runkler, T. A. *Information Mining*. Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiesbaden, 2000.

Takagi, T. and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, Cybern.*, 15, 1985.

Vernieuwe, H., O. Georgieva, B. D. Baets, V. R. N. Pauwels, and N. E. C. Verhoest. Fuzzy models of rainfall-discharge dynamics. *International Fuzzy Systems Association World Congress*, 2003.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Prediction of Significant Wave Height Based on Regression Trees

A.Etemad-Shahidi [a], ***J.Mahjoobi*** [b]

[a] *College of Civil engineering,Iran University of Science and Technology*
(etemad@iust.ac.ir)
[b] *College of Civil engineering,IranUniversity of Science and Technology*
(jmahjoobi@gmail.com)

Abstract: One of the most important factors in design of coastal and offshore structures is significant wave height. Thus, an accurate prediction of wave height is of great importance. In this paper, an alternative approach based on regression trees was applied for prediction of significant wave height. The data set used in this study comprises of wave and wind data gathered from deep water location in Lake Michigan, from 15 September to 10 December, 2002. In this study the data set was divided into two groups. The first one that comprises of 58 days (1392 data point) wind and wave measurement was used as training data to develop the regression tree. The second one that comprises of 29 days (686 data point) wind and wave measurement was used as testing data to verify the model. Wind speeds belonging up to six previous hours were given as input variables, while the significant wave height *(H_s)* was the output parameter. *CART* algorithm was employed for building and evaluating regression trees and outputs of models with different lags were compared. Result showed that regression trees can be used successfully for prediction of H_s. In addition it was found that error statistics of the models for prediction of H_s decrease as wind speed lag increases. Finally, the results of *CART*-based model, was compared with artificial neural networks, Results indicated that error statistics of neural networks were marginally more accurate than regression trees.

Keywords: Wave Prediction; Decision trees; Regression Trees; CART Algorithm, neural networks

1. Introduction

The estimation of significant wave height is essential for almost any engineering activity in the ocean. Different methods such as empirical, numerical and soft computing approaches have been proposed for significant wave height prediction. Recently, artificial neural networks have been widely used to predict wave parameters [e.g. Makarynskyy et al. 2005, Agrawal and Deo 2002, Makarynskyy 2004]. A review of neural network applications in ocean engineering is given in Jain and Deo, [2006]. Recently, other soft computing techniques such as Fuzzy Inference System (FIS) and Adaptive-Network-based Fuzzy Inference System (ANFIS) have been used to develop wave prediction models (e. g. Kazeminezhad et al. 2005, Ozger et al. 2007). These studies have shown that the wind speed is the most important parameter in wave parameters prediction. Prediction of significant wave height is basically an uncertain and random process and hence is difficult to accomplish by using deterministic equations. Therefore, it is ideally suited to regression trees since it is primarily aimed at recognition of a random pattern in a given set of input values. Regression trees are useful to model a random input with the corresponding random output and their application does not require knowledge of the underlying physical process as a precondition. In this paper, regression trees were invoked for prediction of significant

wave height using different lags. For this purpose, *CART* algorithm was employed for building and evaluating regression trees.

2. Regression Trees

Decision trees are powerful and popular tools for classification and prediction. The advantage of decision trees is due to the fact that, decision trees represent rules. Rules can readily be expressed so that humans can understand them. A decision tree is a tree in which each non-leaf node is labelled with an attribute or a question of some sort, and in which the branches at that node correspond to the possible values of the attribute, or answers to the question. Leaf nodes are labelled with a class. Decision trees are used for classifying instances. One starts at the root of the tree, and taking appropriate branches according to the attribute or question asked about at each branch node, one eventually comes to a leaf node. The label on that leaf node is the class for that instance. The Classification and Regression Trees (*CART*) method of Breiman et al. [1984] generates binary decision trees. Regression tree building centers on three major components: (1) a set of questions of the form, Is $X \le d$? Where X is a variable and d is a constant. The response to such questions is yes or no; (2) goodness of split criteria for choosing the best split on a variable; and (3) the generation of summary statistics for terminal nodes. Here, the least squared deviation (*LSD*) impurity measure is used for splitting rules and goodness of fit criteria**.** The *LSD* measure *R(t)* is simply the weighted within node variance for node *t*, and it is equal to the resubstitution estimate of risk for the node. It is defined as:

$$R(t) = \frac{1}{N_W(t)} \sum_{i \in t} \omega_i f_i \left(y_i - \bar{y}(t)\right)^2 \tag{1}$$

where $N_W(t)$ is the weighted number of records in node *t*, ω_i is the value of the weighting field for record *i* (if any), f_i is the value of the frequency field (if any), y_i is the value of the target field, and $\bar{y}(t)$ is the mean of the dependent variable (target field) at node *t*. The *LSD* criterion function for split s at node t is defined as:

$$Q(s,t) = R(t) - R(t_L) - R(t_R) \tag{2}$$

where $R(t_R)$ is the sum of squares of the right child node, and $R(t_L)$ is the sum of squares of the left child node. The split s is chosen to maximize the value of $Q(s, t)$.

Stopping rules control how the algorithm decides when to stop splitting nodes in the tree. Tree growth proceeds until every leaf node in the tree triggers at least one stopping rule. Any of the following conditions will prevent a node from being split:

a. All records in the node have the same value for all predictor fields used by the model.
b. The number of records in the node is less than the minimum parent node size.
c. If the number of records in any of the child nodes resulting from the node's best split is less than the minimum child node size.
d. The best split for the node yields a decrease in impurity that is less than the minimum change in impurity.

In regression trees, each terminal node's predicted category is the weighted mean of the target values for records in the node. This weighted mean is calculated as:

$$\bar{y}(t) = \frac{1}{N_W(t)} \sum_{i \in t} \omega_i f_i y_i \tag{3}$$

where $N_W(t)$ is defined as:

$$N_W(t) = \sum_{i \in t} \omega_i f_i \tag{4}$$

3. Employed Data

The data set used in this study comprises of wave and wind data gathered from deep water location in Lake Michigan, from 15 September to 10 December, 2002. The data set was collected by National Data Buoy Center (NDBC) in station 45007 at 42° 40´ 30´´ N and 87° 01´ 30´´ W (Figure 1), where water depth is 176.4 m. Wind and wave data were collected using 3-meter discus buoy at 1-hour intervals. The wind speed at buoy was measured at a height of 5 meter above the mean sea level. Tables 1 and 2 show ranges and average values of different parameters of training and testing data sets.

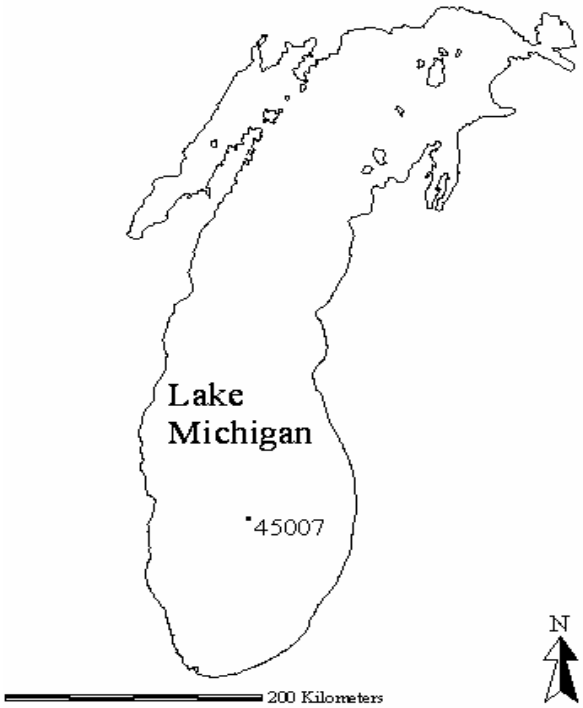


Figure 1. Lake Michigan bathymetry and location of *NDBC* buoy 45007 located at 43°37´09´´ N and 77°24´18´´ W.

Table 1: Ranges and average values of different parameters in training data

Parameter	Range	Average
Wind speed (m/s)	0.1-16.6	7.19
Significant wave height (m)	0.15-3.43	1

Table 2: Ranges and average values of different parameters in testing data

Parameter	Range	Average
Wind speed (m/s)	0.7-16.5	7.84
Significant wave height (m)	0.17-3.36	1.19

4. Building and Evaluating Regression Trees

Here the data set was divided into two groups. The first one that comprises of 58 days (1392 data point) wind and wave measurement was used as training data to develop the regression tree. The second one that comprises of 29 days (686 data point) wind and wave measurement was used as testing data to verify the model. *CART* algorithm was employed to build the regression tree. Wind speeds belonging up to six previous hours were given as input variables, while the significant wave height (H_s) was the output parameter. SPSS Clementine software was used to apply *CART (http://www.spss.com/clementine/).*

The performances of models were evaluated using three statistical measures: (1) bias, which shows the mean error;

$$bias = \bar{y} - \bar{x} \tag{5}$$

(2) Scatter index (*SI*), is the root mean square error normalized by the mean of observed values of the reference quantity;

$$SI = \frac{\sqrt{\frac{1}{n}\sum\left(\left(y_i - \bar{y}\right) - \left(x_i - \bar{x}\right)\right)^2}}{\bar{x}} \tag{6}$$

and (3) Coefficient of correlation (*R*), which is a measure of strength of the linear relationship developed by a model.

$$R = \frac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum\left(x_i - \bar{x}\right)^2 \sum\left(y_i - \bar{y}\right)^2}} \tag{7}$$

where x_i is an observed value, y_i is a predicted value and n is the number of observations, $\bar{x}$ is the mean of x and $\bar{y}$ is the mean of y.

Time series plots of the significant wave height obtained from observations and predicted by the regression trees are displayed in Figures 2 and 3. As seen, the model prediction of H_s matched very well with the observed data.

Table 3 shows the error statistics of models with different wind speed lags (for testing data). As can be seen, the models slightly underestimate significant wave heights (bias=-0. 1 m) in the studied case. Results also indicate that error statistics of the models for prediction of H_s decrease as wind speed lag increases.

Table 3: Error statistics of prediction significant wave height by regression trees

Wave parameter/ model	bias	*SI* (%)	*R*
H_s (m) / 6hr- time lag	-0.08	22.56	0.926
H_s (m) / 5hr- time lag	-0.095	24.71	0.922
H_s (m) / 4hr- time lag	-0.092	24.72	0.922
H_s (m) / 3hr- time lag	-0.1	25.23	0.92
H_s (m) / 2hr- time lag	-0.1	25.75	0.916
H_s (m) / 1hr- time lag	-0.106	27.8	0.9
H_s (m) / no- time lag	-0.109	30.94	0.86

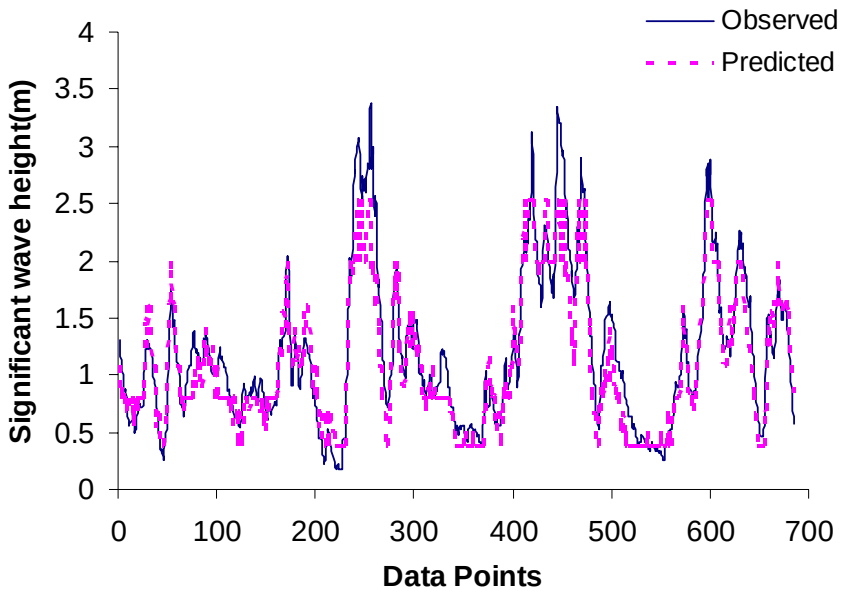


Figure 2. Time series of measured and predicted (6 hour wind speed lag) significant wave heights.

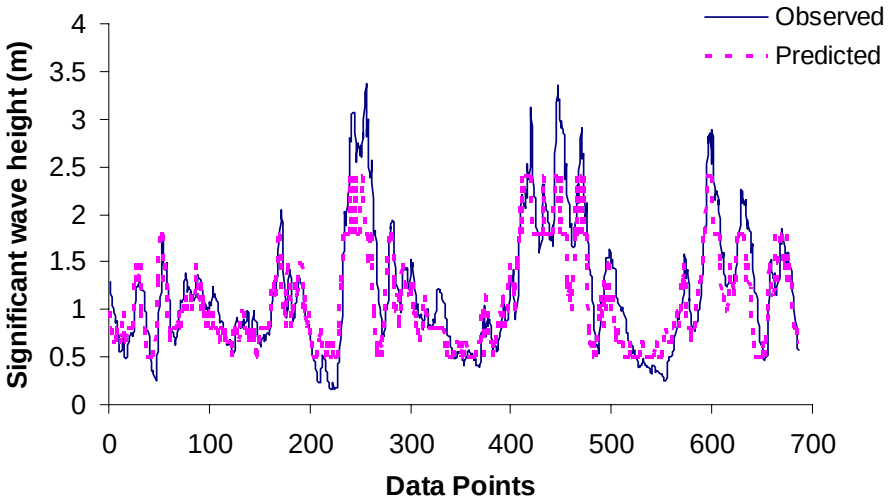


Figure 3. Time series of measured and predicted (0 hour wind speed lag) significant wave heights.

In addition, for comparison between regression trees and artificial neural networks, a three-layer feed-forward neural network (the back-propagation network) [Haykin, 1999] with the sigmoid transfer functions was used. Wind speeds belonging up to six previous hours were given as input variables. From 2 to 20 nodes for the hidden layer were examined. The best topology was found to be 7×15×1 (neurons in the input × hidden × output layers). After testing network, Coefficient of correlation, scatter index and bias were obtained 0.94, 20.96 and -.092 m, respectively. Results indicated that artificial neural networks were marginally more accurate than regression trees.

5. Summary and Conclusions

Significant wave height is an important parameter in the design of coastal and offshore structures. In this study, regression trees were used successfully for prediction of significant wave height variation responding to wind forcing. *CART* algorithm was employed for building and evaluating regression trees. The data set used in this paper comprises of wave and wind data gathered from deep water location in Lake Michigan. Wind speeds belonging up to six previous hours were given as input variables. Result show that *CART* algorithm is skilful in prediction of significant wave heights in the studied case. Furthermore, it was found that the error statistics of the models for prediction of H_s decrease as wind speed lag increases. Also, results of regression trees were compared with those of artificial neural networks (ANNs). Results show that ANNs models were marginally more accurate than regression trees. Therefore, the regression trees can be used as a cost effective and easy to use tool for engineers and scientists with much less effort required for the implementation of process-based models.

References

Agrawal J.D., Deo M.C. On-line wave prediction, Marine Structure, 15, 57–74, 2002.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. Classification and Regression Trees. Belmont, CA: Wadsworth Statistical Press, 1984.

Haykin S. Neural Networks: a Comprehensive Foundation. 2th Ed. Prentice-Hall, NJ, 842pp, 1999.

Jain, P and M. C. Deo, M.C. Neural networks in ocean engineering, International Journal of Ships and Offshore Structures, 1, 25-35, 2006.

Kazeminezhad M.H., Etemad-Shahidi A., Mousavi S.J. Application of fuzzy inference system in the prediction wave parameters. Ocean Engineering. 32, 1709-1725,2005.

Makarynskyy O. Improving wave predictions with artificial neural networks. Ocean Engineering, 31, 5–6, 709–724, 2004.

Makarynskyy O., Pires-Silva A.A., Makarynska D., Ventura-Soares, C. Artificial neural networks in wave predictions at the west coast of Portugal, Computers & Geosciences,; 31, 4, 415-424,2005.

Ozger M., Sen Z. Prediction of wave parameters by using fuzzy logic approach. Ocean Engineering., 34, 460-469,2007.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A spatial decision support system for the optimal environmental reclamation of open-pit coal mines in Greece

Dr. F. Pavloudakis[a], Dr. M. Galetakis[b] and Dr. Ch. Roumpos[a]
[a] Public Power Corporation of Greece – Mines Division.
Kifisou & Dirachiou 89, 104 43 Sepolia, Athens - Greece.
[b] Technical University of Crete - Department of Mineral Resources Engineering. University Campus, Akrotiri, 73 100 Chania - Greece. (galetaki@mred.tuc.gr)

Abstract: Mine reclamation is an integral part of the mineral development process. The selection of land uses after the mine closure is a difficult decision, which is complicated further due to the variety of parameters that must be taken into account trying to provide the local community with a viable development plan. Conventional methods used for reclamation planning are characterised by the lack of data integration and by time-consuming analysis. In this study we propose a spatial decision-support system (SDSS) that minimizes these problems, as data integration and analysis are offered within one computerized environment. A geographical information system and multi-criteria decision-making methods, based on binary integer linear programming models, have been integrated to select the appropriate land use in different parts of a post-mining area taking into account social, technical, economic, environmental and safety criteria. The proposed SDSS was used for the selection of the optimal landscape reclamation strategy of the Amynteon lignite surface mine located at West Macedonia Lignite Centre, Northern Greece. Based on the developed mine maps, the model variables are assessed and incorporated into the objective optimization function. Emphasis is placed on the spatial diversification of the model variables. The application demonstrates that the decision-support system allows the mining company to determine in an efficient way the specific land use (agricultural land, forest, recreational area and industrial zone) that is considered the most suitable for every part of the study area.

***Keywords**:* mining; reclamation; decision support systems; linear programming;GIS.

1. INTRODUCTION

Coal deposits exploitation in open-pits is an activity that impacts adversely the environment. No matter the effectiveness of the environmental management practices that are applied by the mine operator, often, these mining activities led to severe environmental degradation just because of their enormous size. Moreover, these activities last for many decades. Otherwise there is not enough pay-back time for the capital invested for mine development. As a consequence, mining activities affect irreversibly the local communities, which adjust their life-style and depend their economic prosperity on the development plans of the mining company [Pavloudakis et al., 2004, 2006].

In this context, the selection of land uses after the mine closure is a difficult decision, which is complicated further due to the variety of parameters that must be taken into account trying to provide the local community with a viable development plan. Many of those decision-making parameters are characterized by spatial diversification. For instance, parameters such as geological structure, hydrological pattern, land morphology, slopes stability and orientation, distance from public utilities, accessibility from existing roads, etc. are considered critical for deciding for the development of agricultural, recreational or residential activities in a rehabilitated mining area.

Conventional methods used for reclamation planning and design are characterised by the lack of data integration and by time-consuming analysis. Therefore, the main objective of the present paper is to demonstrate how Spatial Decision-Support System (SDSS) minimize these problems providing a single computer platform of data integration and analysis: a geographical information system (GIS) and a multi-criteria decision-making (MCDM) tool are combined to select the appropriate land use in different parts of a post-mining area taking into account technical, economic, environmental and safety criteria. Since SDSS aided reclamation planning and design will be time efficient, several reclamation strategies can be generated and evaluated with the use of different priority schemes offering significant help in the selection of the optimal land reclamation method [Fortner et al., 2004, Giove et al., 2007].

2. GIS AND MULTICRITERIA DECISION ANALYSIS

Environmental decisions are often complex, multi-faceted, and involve many different stakeholders with different priorities or objectives. Moreover, environmental decisions typically draw upon multidisciplinary knowledge bases, incorporating natural, physical, and social sciences, medicine, politics, and ethics. This fact, and the tendency of environmental issues to involve shared resources and broad constituencies, means that group decision processes are called for. For environmental management projects, decision makers may currently receive four types of technical input: modelling/monitoring, risk analysis, cost or cost-benefit analysis, and stakeholders' preferences. Every technique incorporates information coming in different forms. While modelling and monitoring results are usually presented as quantitative estimates, risk assessment and cost-benefit analyses allow a higher degree of qualitative judgment by the project team. [McDaniels et al., 1999; Kiker et al., 2005; Linkov et al., 2004]. The common purpose of these diverse methods is to be able to evaluate and choose among alternatives based on multiple criteria using systematic analysis that overcomes the observed limitations of unstructured individual and group decision-making. Some techniques rank options, some identify a single optimal alternative, some provide an incomplete ranking, and others differentiate between acceptable and unacceptable alternatives [Guitouni and Martel, 1998; Diwekar and Small, 2002; Linkov et al., 2004].

Focusing on Multi-Criteria Decision Analysis (MCDA) models, these allow for more accurate representation of decision problems, by accounting for several objectives. These often exhibit a hierarchical structure. The highest level represents the broad overall objectives (e.g., improving the soil fertility in reclaimed mine areas). They may be broken down into more operational and practically assessed lower level objectives (e.g., increasing the content of nutrients, enhancing the mechanical properties). Sometimes only proxies are available (e.g., if the objective is to enhance recreation opportunities, the number of recreation days can be used). In MCDA models, value judgments may be required in choosing the proper attribute, but measurement does not have to be in monetary terms, like in the case of using cost-benefit analysis as a single criterion. Moreover, a critical advantage of MCDA modelling in group decisions is the capacity for calling attention to similarities or potential areas of conflict between stakeholders with different views, which results in a more complete understanding of the values held by others [Kiker et al., 2005].

The conceptual idea, on which most of GIS-based multicriteria analyses rely, is to use the GIS capabilities to prepare an adequate platform for using multicriteria methods. The GIS-based multicriteria analysis starts with the problem identification, where the capabilities of the GIS are used to define the set of feasible alternatives and the set of criteria. Then, the overlay procedures are used in order to reduce an initially rich set of alternatives into a small number of alternatives which are easily evaluated by using a multicriteria method [Chakhar and Martel, 2003; Chakhar and Mousseau, 2007]. The most common integrations of GIS with MCDM systems include the integration with expert systems and mathematical models such as linear programming (LP). Guerra and Lewis [2002] have developed a spatial optimisation model using GIS and an external linear programming solver for selecting the site of a reserve for reintroducing wildlife species. Moreover, McCarthy et al. [2008] have integrated GIS and an expert system platform for processing large amounts of

spatial data deriving from sensors for making decisions relevant to slopes stability and geotechnical hazards. Other examples of SDSS in land use management and risk analysis are provided by Chen et al. [2001] and Roetter et al. [2005].

The optimization model (a GIS in combination with a LP model), presented in the following paragraphs is capable of supporting decisions about land uses that must be developed in different parts of an area affected by surface mining operations. For this purpose, it employs numerical scores to communicate the benefits resulting from one option in comparison to the others, based on a single scale. Aggregated scores are developed from the performance of alternative land uses with respect to all individual criteria. Individual scores are averaged using a weighting mechanism which favours the most important criteria.

3. DESCRIPTION OF THE STUDY AREA

The study area is located in Ptolemais lignite-bearing basin, in the Region of West Macedonia, Northern Greece. In particular, the lignite deposit of Amynteon extends over an area of 14 km². The surface mine that has been developed and operates in this area during the last 25 years occupies currently 6,000 ha and supplies with lignite a thermal power plant with an installed capacity of 600MW. The Amynteon mine pit is developed from a depth of 40-50m, in the eastern boxcut area, and reaches the depth of 250m in the central section of the western rim slope and the depth of 180m in the area of the southern rim slope. After the mine closure, the final pit will be filled up with water. For stability reasons, the inclinations of the final pit slopes are planned to be 1:5 on the excavation side and 1:7 on the waste dumping side. These values will be probably reviewed after a thorough investigation of the type of earth materials that will be dumped in the wave-attack zone of the artificial lake.

The mine and the power plant are surrounded by a relatively flat agricultural area, with an extensive network of irrigation channels and three lakes: Himaditis, Petron and Vegoritis. Some areas nearby the water bodies are controlled by laws and regulations relevant to the preservation of wildlife and of sensitive environmental components, which set specific restrictions regarding the development of human activities. Moreover, in the vicinity of the mine there are several villages. According to the National Regulation of Mining and Quarries Works, mining activities must keep a clearance of 250m from residential areas.

3.1 Environmental aspects

Open-pit lignite mines are complex operations that can affect numerous environmental constituents in various ways. Pavloudakis and Roumpos [2004] have described the “processes” and “waste streams” that are considered to have greater potential of causing environmental damage in lignite surface mining areas. The resulting impacts can be classified to those related to the mine operation and last until the mine closure, such as the deterioration of ambient air quality due to dust emissions, and those of permanent character unless proper land reclamations measures are applied. In the latter category the following impacts are included: alteration of morphology, changes in hydrological pattern, loss of wildlife habitat, degradation of landscape value, reduction of property value, loss of topsoil and loss of agricultural land. The above impacts must take increased interest if the planning of land uses for the post-mining era is in question.

3.2 Social and economic aspects

The development of every large-scale mining project aims at creation of wealth and employment, boosting, in this way, the National, Regional and local economies. Nevertheless, there are many adverse socioeconomic impacts that can violate the quality standards for the communities located in the vicinity of mining sites. These impacts can be summarised in the following three broad categories: changes in economic activities,

occupation of land for a long-period (usually many decades) and limited access to public utilities. For a mine operator, the development of public relations on the basis of mutual respect and understanding with local authorities is a critical parameter for the acceptance of a surface mining activity and reclamation plan. For this reason, mining companies must provide every stakeholder with all information necessary to explain not only the expected environmental impacts but also the benefits resulting for the local communities.

3.3 Planning for the post-mining period

According to the environmental permits that regulate the mining operations in the study area, the mine operator is obliged to carry out and finance land reclamations works to a certain extent that is described in the permits and in supplementary studies. However, these works are possible to be proved inadequate for assuring the economic prosperity of the local communities. Therefore, from the early stages of the land reclamation works, the local authorities must agree with the mine operator on a development strategy, which will be facilitated by all mine-land reclamation activities. To this extent, the proposed SDSS model provides to the public, local authorities and experts the possibility of expressing their opinions in two ways: (i) setting restrictions or presenting their preferences regarding the percentage of the coverage of the study area from every land use, and (ii) proposing specific land uses for specific parts of the study area.

4. METHODOLOGY

4.1 Information sources

Spatial information was obtained mainly from an orthophotomap (Figure 1). For the purposes of land reclamation planning and co-operation with local and supervisory authorities the orthophotomap is printed in scales of 1:15,000 and 1:32,000 allowing positional accuracy down to 7.5m. Spatial information was also obtained from maps and pictures that represent parts of the study area prior to the mine development, as well as from plots and sections produced from the mining company, determining the boundaries and the inclination of the final slopes of the pit and the waste heaps at the time of mine closure. Moreover, other sources of information that were used are:

- meetings and discussions with experts, citizens and representatives of local authorities
- the terms and conditions of the Amynteon mine environmental permits
- legal documents (e.g. National Regulation of Mining & Quarrying Works)
- laboratory analyses carried out for determining the fertility and other properties of the soil, which are closely related to the possibility of developing certain land uses.

Additional information deriving from monitoring of several geohydrological, water quality and wildlife parameters can be included in the system as long as reliable data bases will be available.

4.2 Selection of criteria

Once the problem has been delineated and the information sources have been identified, the spatial multi-criteria analysis focuses on the selection of criteria. This is a process that depends on the characteristics of the decision problem, even though there are some techniques which facilitate the choice of the parameters which influence decision making [Keeney & Raiffa, 1993]. In the case of Amynteon lignite mining area the determination of land uses is based on a multi stage procedure and a series of criteria presented in Figure 2. The criteria are classified into those showing spatial variation and those with no spatial character. In step 1, which includes all these criteria with no spatial character, priorities and limitations are declared in a way that affects the formulation of the ultimate goal of the modelling procedure. Furthermore, the spatial analysis criteria are different from the criteria of step 2, which are used for selecting one of the following land uses: ‘agriculture land’, ‘forest’, ‘recreational area’ and ‘industrial zone’, and the criteria of step 3, which are relevant to the suitability of soil for supporting the rapid development of a vegetative cover and are used for determining the type of trees or the type of cultivation, in case that the

selected land use is ‘forest’ or ‘agricultural land’ respectively.



Figure 1. Orthophotomap of the greater area of Amynteon mine

4.3 Ranking of criteria

In the examined case the criteria do not have similar importance. To stress their differences, each criterion must be assigned to a specific weight that corresponds to its relative significance. For this purpose several methods are offered. The most popular are the ‘ranking’, ‘rating’, ‘pairwise comparison’ and ‘trade-off’ methods [Hair et al., 1995]. For the particular model and for the criteria used in the first step, the ranking method was considered as the most suitable. Starting from the most important factor, the decided rank was: 1. legal restrictions, 2. public opinion, 3. experts opinion, 4. potential of wealth creation, 5. cost, and 6. former land uses. Regarding the spatial criteria, these were assigned specific weights using again the ranking method. For the criteria of Step 2 the rank is: 1. slope, 2. soil fertility, 3. proximity to lakes, 4. proximity to archaeological sites, 5. proximity to villages, while for the criteria of Step 3 the rank is: 1. irrigation infrastructures, 2. pH, acidity, alkalinity, 3. nutrients concentration, 4. mechanical properties, 5. metals concentration. It worth to be noticed that the criteria of slope, proximity to lakes and metals concentrations are closely related to legal or safety constrains, which can be prohibitive for certain land uses. For this reason, these criteria must be distinguished from others, which may influence the model either positively or negatively, depending on their values. For instance, if the slope of a reclaimed surface is higher than 1:5 this surface is unsuitable for any other land use but reforestation, while if an area is within the preservation zone that surrounds a lake (e.g. zone of “Natura 2000”) is unsuitable for any other land use but reforestation and low-intensity agricultural use.

4.4 Spatial analysis

The spatial analysis of the examined decision-making problem starts with the definition of restrictions and preferences regarding the land uses of the reclaimed mine area. For instance, a preference can be expressed as: ‘at least 1,000 ha of agricultural land must be developed’, while a restriction as: ‘the industrial zone will be less than 500 ha’. In case of an opposite opinion about land uses, the one that will be taken into account will be chosen based on the rank of importance presented in paragraph 5.3 for the criteria of step 1. Then, the study area is divided into small squares. The dimensions of those squares can be decided taking into account the size of the study area and the variability of the examined criteria. For assessing the fertility of the reclaimed mine surfaces four soil samples per hectare are suggested by the literature [Hansen, 1996]. So, the square dimensions can be 50m x 50m unless this choice requires enormous sampling, analytical and data processing work.

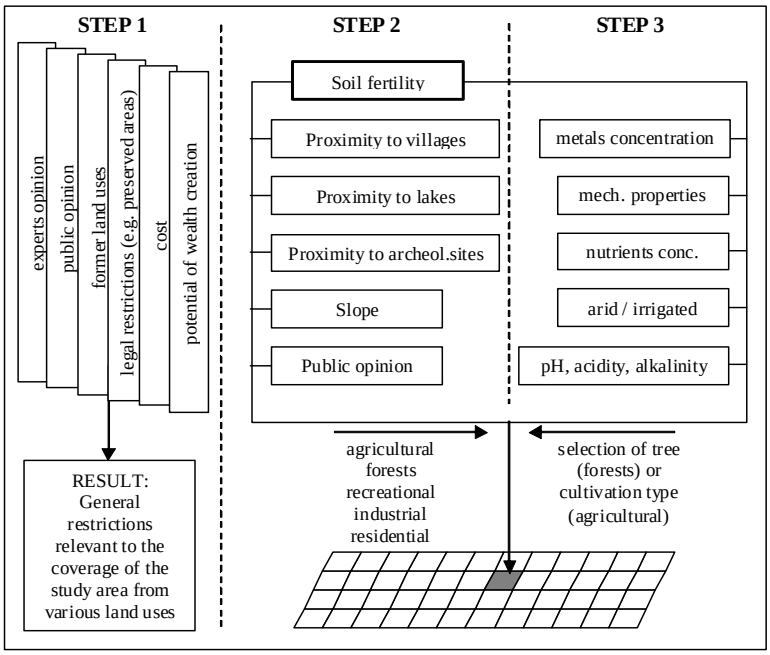


Figure 2. Criteria incorporated at the successive steps of the decision-making procedure.

For every square of the reclaimed mine land, the most suitable land use is proposed by applying the following rating scale: 0=low, 1=medium, 2=high. Table 1 presents a matrix with the values of the step 2 criteria which are considered the optimum ones for the selection of each one of the four alternative land uses. Using the above presented optimum values, for each square-element of the study area and for each criterion the absolute value of the deviation of the actual rate from the optimum one is calculated (Table 2). Then, the sum of deviations of the five criteria are calculated for each land use and the land uses are ranked starting from the one that scores the least sum of deviations from the optimum criteria values. In case of equal scores for two or more land uses, the ranking procedure takes into account the criteria ranking of step 2 presented in the paragraph 4.3. The final step of the proposed SDSS regards the review of the above-selected land uses for all square elements of the study area in a way that meets the limitations set in step 1. For this purpose, the decision-variables are defined as shown in the Table 3, which derives from Table 2 by ranking the four alternatives land uses according to their sum of deviations from the optimum criteria values.

Table 1. Optimum criteria values for the selection of the four alternative land uses.

Criteria	Agricultural land	Forest	Recreational area	Industrial zone
Slope	0	1 or 2	0	0
Soil fertility	2	0 or 1	1 or 2	0
Proximity to lakes	0 or 1	0 or 1 or 2	2	0
Proximity to archaeological sites	0 or 1	0 or 1 or 2	2	0
Proximity to villages	0 or 1	0 or 1 or 2	2	0

Table 2. Deviation from the preferable land uses' criteria values (best rates in bold)

Squares (i, j)	Agricultural land	Forest	Recreational area	Industrial zone
	Deviation from the preferable criteria values			
0,0	**0**	1	6	3
0,1	3	**1**	5	4
0,2	2	1	**0**	3
…	…	…	…	…
n, m	2	2	7	**0**

Table 3. Ranking of the four alternative land uses according to their suitability

Squares (i, j)	Agricultural land	Forest	Recreational area	Industrial zone
	C_{ij1}	C_{ij2}	C_{ij3}	C_{ij4}
0,0	**1**	2	4	3
0,1	2	**1**	4	3
0,2	3	2	**1**	4
…	…	…	…	…
n, m	2	3	4	**1**

Based on Table 3, the ranking of land uses of each square (i,j) is possible. For instance means that for the square-element (0,0) the most suitable land use is 'agricultural land'. It is now possible to use a minimization algorithm to find out the optimal land uses combination for all square-elements of the study area (discrete optimisation problem solved by applying binary linear integer programming). The objective function Z which must to be minimised is as follows:

$$MinZ = \sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=1}^{4} C_{ijk}X_{ijk} \qquad X_{ijk} = 0 \; or \; 1, \forall i = 0,...,n, \quad j = 0,...,m, \quad k = 1,..,4 \tag{1}$$

under the constraints:

$$\sum_{k=1}^{4} X_{ijk} = 1, \forall i = 0,...,n, \quad j = 0,...,m \tag{2}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} X_{ij1} \le K \qquad \sum_{i=0}^{n}\sum_{j=0}^{m} X_{ij2} \le L \qquad \sum_{i=0}^{n}\sum_{j=0}^{m} X_{ij3} \ge M \qquad \sum_{i=0}^{n}\sum_{j=0}^{m} X_{ij4} \le N \tag{3}$$

where, K=maximum agricultural land area (i.e. max number of square-elements), L=maximum forest area, M=minimum recreational area and N=maximum industrial zone area

The final result of this procedure is illustrated in Figure 3, which refers to a part of the study area.

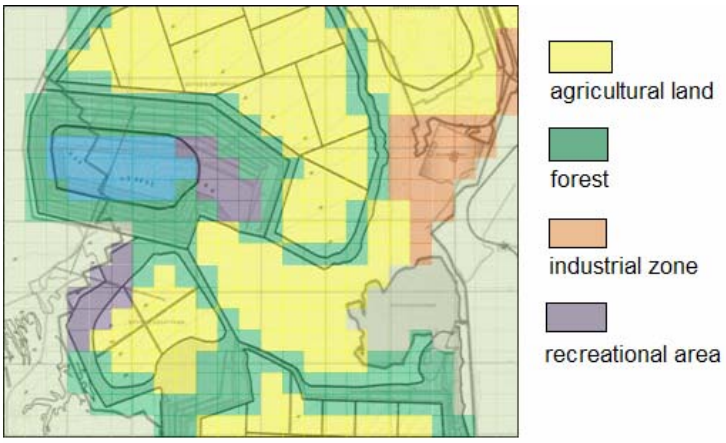


Figure 3: Final proposal regarding land uses in the study area.

5. CONCLUSIONS

The primary objective of the study was to facilitate the land use processes, of areas affected by mining, in the presence of multiple and diverse decision criteria using GIS in combination with a LP model. To achieve this, effort focused mainly on the following sub-objectives: (i) to use GIS as tool to provide all the necessary data and to visualize the results (ii) to develop a multiple criteria decision support system based on binary integer LP, and (iii) to integrate the LP model and the GIS in a SDSS to facilitate the land use suitability analysis. The application of the developed SDSS in the selection of the optimal landscape reclamation strategy of the Amynteon lignite surface mine located at West

Macedonia Lignite Center, Northern Greece indicated that the decision-support system allows the mining company to evaluate different reclamation schemes maximizing the long-term sustainability of the broader mining area.

REFERENCES

Chakhar, S., Martel, J.M., Enhancing geographical information systems capabilities with multi-criteria evaluation functions, *"Journal of Geographic Information and Decision Analysis"*, Vol. 7(2), 47-71, 2003.

Chakhar, S., Mousseau, V., Multicriteria Spatial Decision Support Systems. In Shashi Shekhar and Hui Xiong (eds.), *Encyclopedia of Geographical Information Science*. Springer, 2007.

Chen, K., Blong, R., Jacobson, C., MCE-RISK: integrating multicriteria evaluation and GIS for risk decision-making in natural hazards, *Environmental Modelling & Software*, 16(4), 387-397, 2001.

Diwekar, U., Small M.J., Process analysis approach to industrial ecology, Chapter 11 in *"A Handbook of Industrial Ecology"*, R.U. Ayres and L.W. Ayres, eds., Edward Elgar Ltd, Cheltenham, UK, 114-137, 2002.

Fortner A., Smith, D., and Rohrer C., GIS as a prioritization and planning tool in abandoned mine Reclamation, paper presented at Advanced Integration of Geospatial Technologies in mining and Reclamation, Atlanta, December 6-10, 2004.

Giove, S., Agostini P., Critto, A., Semenzin, E., and Marcomini A., Decision support systems for the management of contaminated sites. A Multi-Criteria Approach, *Environmental Security in Harbors and Coastal Areas*, I. Linkov et al. (eds.), 267–273, 2007.

Guerra, G., Lewis, J., Spatial optimisation and GIS – Locating an optimal habitat for wildlife reintroduction, *ArcUser*, April-June, 2002.

Guitouni, A., Martel. J.M., Tentative Guidelines to Help Choosing an Appropriate MCDA Method, *European Journal Operations Research*, Vol. 109, 501-521, 1998.

Hair, J.F., Anderson, R.E., Tatham, R.L. & Black, W.C., *Multivariate Data Analysis with Readings*, Prentice Hall, New Jersey, 1995.

Hansen M.M., (edt), *Handbook of western reclamation techniques*, University of Wyoming & US Office of Surface Mining, 504pp, 1996.

Keeney, R.L., Raiffa, H., *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Cambridge University Press, Cambridge, 1993.

Kiker, G.A., Bridges, T.S., Varghese, A., Seager, T.P., Linkov, I., Application of multicriteria decision analysis in environmental decision making, *Integrated environmental assessment & management*, Vol. 1(2), 95-108, 2005.

Linkov, I., Varghese, A., Jamil, S., Seager, T.P., Kiker, G., Bridges, T., Multi-criteria decision analysis: A framework for structuring remedial decisions at contaminated sites, in Linkov, I. and Ramadan, A. (eds.), "*Comparative Risk Assessment and Environmental Decision Making*", Kluwer, 2004, p. 15-54.

McDaniels, T.L., Gregory, R.S., Fieldsm D., Democratizing risk management: Successful public involvement in local water management decisions, *Risk Analysis*, 19, 497–510, 1999.

McCarthy, J.D., Graniero, P.A., Rozic, S.M., An integrated GIS – expert system framework for live hazard monitoring and detection, *Sensors*, 8, 830-846, 2008.

Pavloudakis, F. Roumpos, C. Evaluation of land reclamation and environmental protection strategies in open-pit lignite mines, Paper presented at Advances in Mineral Resources Management and Environmental Geotechnology, Chania, Greece, June 9-11, 2004.

Pavloudakis, F., Roumpos C., Sachanidis C. & Galetakis M., Environmental impacts assessment of surface mining operations: A review of the Greek experience, Paper presented at *Mine Planning and Equipment Selection*, Torino, Italy, September 22-26, 2006.

Roetter, R.P., Hoanh, C.T., Laborte, A.G., Van Keulen, H., Van Ittersum, M.K., Dreiser, C., Van Diepen, C.A., De Ridder, N., Van Laar, H.H., Integration of Systems Network (SysNet) tools for regional land use scenario analysis in Asia, *Environmental Modelling & Software*, 20(3), 291-307, 2005.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Design and Implementation of the Land Use Management Support System (LUMASS)

A. Herzig

Department of Geography, Division of Landscape Ecology and Geoinformation Sciences, Christian-Albrechts-Universität zu Kiel, Germany (herzig@geographie.uni-kiel.de)

Abstract: Because of the increasing complexity of spatial planning processes there is a need for more powerful and more intelligent planning tools. This involves the development of environmental decision support systems (EDSS), or, in a more general term, spatial decision support systems (SDSS). Their design, implementation, and application have been discussed in many articles within the last 15 years or so. A summary of the most important aspects is given here, which distinguishes between functional and structural requirements of SDSS. With regard to the latter, it is argued that, from an end user's point of view, only a tight coupling or full integration strategy with respect to the involved software components fulfils the functional demands on SDSS. Further, the Land Use Management Support System (LUMASS) is introduced, which was developed in consideration of the functional and structural specifications discussed above. It integrates (i) the commercial GIS package ArcGIS™, (ii) ecological models focusing on soil and watercourse protection, and (iii) the Open Source Mixed Integer Linear Programming System lp_solve [Berkelaar et al., 2004] and follows a tight coupling approach. Finally, a sample application is presented that reveals the optimization of land use patterns with a view to minimizing soil erosion and sediment discharge into adjacent creeks and rivers.

Keywords: Spatial Decision Support System; LUMASS; Ecological Modelling; Land Use Pattern Optimization; Soil and Watercourse protection

1 INTRODUCTION

From the landscape ecologist's point of view, land use management aims at the sparing use of natural resources as well as the sustainable protection of the efficiency of the landscape budget. In the medium or long term, this is achieved by optimizing land use patterns according to environmental standards, e.g. as prescribed by the proposal of the EU Soil Framework Directive and the EU Water Framework Directive. In this context, the Land Use Management Support System (LUMASS) was developed in order to assist scientists and environmental managers in assessing land use impacts on the landscape budget with respect to specified criteria (e.g. minimizing soil erosion). Based on these assessments and additional constraints (i.e. area shares of individual land uses) optimum land use patterns can be generated using the multiobjective optimization module of LUMASS.

The underlying framework for the design and implementation of LUMASS is derived from the SDSS literature covering the last 15 years or so. In the following, the most important aspects are summarized and differentiated according to the functional and structural requirements of SDSS. Subsequently it will be described how LUMASS was designed and implemented to meet the aforementioned functional and structural requirements. Finally the application of the system is demonstrated using an example of the optimization of land use patterns so as to minimize soil erosion and sediment discharge, taking into account the specified area shares of specified land use alternatives.

2 REQUIREMENTS OF SPATIAL DECISION SUPPORT SYSTEMS

2.1 Functional Requirements

In order to identify the functional requirements of spatial decision support systems, it is useful to first take a closer look at the general decision-making process itself. According to Wessels and Wierzbicki [2000, p. 9], there are three vitally important issues that have to be addressed within the decision-making process: "(i) information about the current situation and history; (ii) the relation between basic processes and actions or decisions; and (iii) the decision process." In the course of a decision-making process, these issues are treated within different phases, which go back to Simon [1960]. In the broadest sense, they may be characterized as follows [cf. Wessels and Wierzbicki, 2000; Makowski and Wierzbicki, 2000]:

Intelligence: The intelligence phase of the decision-making process is devoted to the identification and investigation of the decision problem. Besides data gathering and as accurate a delineation of the decision problem as possible, this phase includes the determination of problem specific evaluation criteria as well as the choice of methods and models for assessing criterion scores with respect to the decision problem.

Design: The design phase involves the development and evaluation of decision alternatives. Therefore, the processes affecting decisions are analysed and different decision alternatives are generated with respect to different magnitudes of preference concerning the evaluation criteria. Finally, each alternative is evaluated in terms of the given criteria.

Choice: In the choice phase, decision support methods and models are applied to the problem in order to make the best choice among the given alternatives.

Although the aforementioned approach looks like a sequence of steps to be performed one after another, in reality, it is accomplished in a recursive manner. That is, each phase may itself lead to a complex decision problem with its own intelligence, design and choice phases [Simon, 1960]. Furthermore, at each stage of the process it is possible to return to a preceding one in order to revise it and to repeat the successive stages with modified settings [Janssen, 1996], thus generating decision alternatives or evaluating the weighting scheme of objective functions following an iterative procedure.

Since spatial decision-making may be regarded as a special case of decision-making in general, the aforementioned approach seems to be suitable for spatial decision-making as well. In fact, a multi-stage or multi-level approach is also used in the spatial decision-making literature [cf. Jankowski, 1995; Janssen, 1996; Malczewski, 1999; Poch et al., 2004]. Even if there are some deviations from the above characterization concerning the number of stages and the assignment of tasks [e.g. Jankowski, 1995; Poch et al. 2004], there seems to be a consensus about the essential phases and tasks of spatial decision-making. However, what distinguishes spatial decision-making from non-spatial decision-making? It is the dependence of choices on one or more of the following spatial characteristics of given alternatives: (i) *location*, i.e. the absolute position given by two or three-dimensional coordinates [e.g. Carver, 1991]; (ii) *shape*, i.e. the geometry or contiguity [e.g. Aerts and Heuvelink, 2002]; and (iii) *topology*, i.e. the relative position in relation to other objects [e.g. Tourino et al., 2003].

So, what does this imply with regard to the functional requirements of spatial decision support systems? Generally speaking, the task of (spatial) decision support systems is to provide the user with appropriate methods and tools to support each phase of decision-making [Janssen, 1996; Wessels and Wierzbicki, 2000]. Therefore, the implementation of spatial decision support systems entails the integration of the following capabilities [cf. Fedra and Reitsma, 1990; Jankowski, 1995; Fedra 1996; Djokic, 1996; Malczewski, 1999, Denzer, 2002, Poch et al. 2004]: (i) spatial data management, analysis, and presentation (i.e. GIS), (ii) modelling of spatial processes, and (iii) assisting the user in solving complex spatial decision problems. A more detailed set of distinguishing features of spatial decision support systems is provided by Densham [1993, based on Geoffrion, 1983] as well as by Rizzoli and Young [1997]. Figure 1 shows how the most important of these features are integrated with the 3-phase model of decision-making described above.

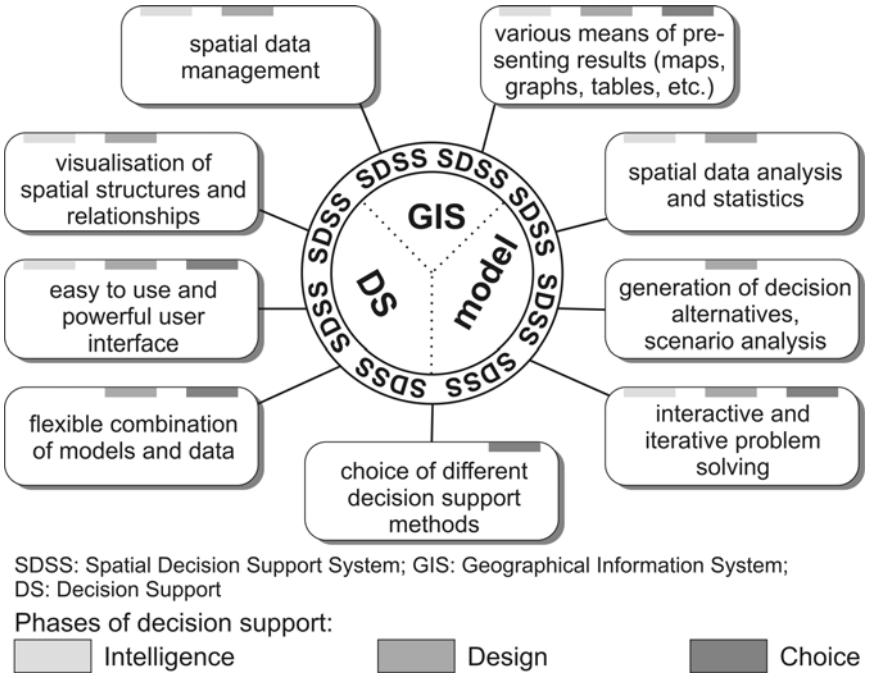


Figure 1. Functional requirements of spatial decision support systems [Geoffrion, 1983; Densham, 1993; Janssen 1996; Rizzoli and Young, 1997; Makowski and Wierzbicki, 2000].

One of the most important features in figure 1 is the ability to carry out the spatial decision-making process in an iterative and interactive manner, thus allowing the user to explore the set of feasible solutions by adjusting inherent model parameters and criterion weights or by preselecting spatial alternatives. Another important feature is the flexible combination of data and models. That is, it should be possible to run different process models on a once configured database in order to analyse the interdependencies of processes and decisions and to develop decision alternatives. Since both of the just mentioned features require some user input and configuration, they also need to have a powerful and easy-to-use user interface providing access to the relevant model parameters as well as to the underlying database. Another subset of important SDSS features comprises the already mentioned capabilities of spatial data management and processing. Here special emphasis should be placed on the methods and tools for analysing spatial structures and relationships as well as their visual representation in the form of maps, tables, graphs, etc.

A review of a random sample of articles dealing with SDSS showed, that the minority of systems integrate all of the aforementioned functional components (i.e. GIS, spatial process models, decision support methods) into the decision making process [e.g. Lam and Pupp, 1996; Joerin and Musy, 2000; Ostfeld et al., 2001; Mendoza et al., 2002a,b; Eldrandaly et al., 2003; Tourino et al., 2003; Li et al. 2005]. The majority of articles focus on either GIS and model integration or GIS and decision support (DS) integration.

2.2 Structural Requirements

In principal, there are three different approaches to the implementation of spatial decision support systems [cf. Djokic, 1996, modified; see also Rizzoli and Young, 1997]: (i) completely new development, (ii) coupling of existing components, (iii) partial new development, reusing existing components. Of these, completely new development is certainly the most flexible, but also the most expensive of the three alternatives [Rizzoli and Young, 1997], and it is rarely described in the literature [e.g. Lam and Pupp, 1996]. Coupling existing software components is much more economical [Denzer, 2005] and minimizes the need for new development. The functionality of the components in question is predetermined by the functional demands on spatial decision support systems as

described in the previous section**.** Denzer [2005, p. 1218] considers four different building blocks: (i) models, (ii) geographical information systems (GIS), (iii) decision support systems (DSS), and (iv) data management systems. The challenge of the coupling approach lies in realizing data exchange between the individual components [cf. Johnston, 1990; Djokic, 1996; Reitsma and Carron, 1997; Ungerer and Goodchild, 2002; Eldrandaly et al., 2003; Denzer, 2005]. Depending on the systems involved, different interfaces may be used, resulting in different levels of integration. In general, the literature distinguishes three main types of coupling strategies [cf. Nygeres, 1992, cited in Jankowski, 1995; Fedra, 1996; Ungerer and Goodchild, 2002], which are roughly characterized below. For a more detailed discussion, please refer to the cited articles.

Loose coupling. In loose coupling, the user operates separate stand-alone applications, and data exchange relies on data files (usually in text format) being exported and imported.

Tight coupling. In tight coupling, the user operates separate applications or software components sharing a common user interface. Data exchange is automated and may be implemented by a variety of interfaces such as DDE (Dynamic Data Exchange), COM (Component Object Model), CORBA (Common Object Request Broker Architecture), etc.

Full integration. In fully integrated systems, there is no need for data exchange since the user operates a single application with a single database.

With respect to the structural demands on SDSS, only tightly coupled and fully integrated systems are capable of fulfilling the functional requirements of SDSS as described in the previous section. This becomes particularly apparent with regard to the need for flexible combination of data and models and, more obviously, for a common and easy-to-use user interface providing access to all functional components of the SDSS. Among the previously cited articles dealing with SDSS (cf. section 2.1), which fulfil the functional requirements, only some of them also fulfil the structural requirements discussed above [e.g. Lam and Pupp, 1996; Tourino et al., 2003; Eldrandaly et al., 2003; Li et al., 2005]. This implies, that only a minority of SDSS described in the literature seems to be ready-to-use systems tailored to the end-user for use in daily planning processes.

3 THE LAND USE MANAGEMENT SUPPORT SYSTEM (LUMASS)

3.1 Functional Characteristics and Features

In order to implement a SDSS for the purpose of land management or land use management respectively, appropriate models or methods have to be chosen for implementation in accordance with the general functional requirements of SDSS. With regard to the spatial modelling capabilities, land management systems, which are described in the literature, exhibit a broad range of used models. They range from watershed-scale hydrologic models [e.g. Kaur et al., 2004] to noise propagation models [e.g. Joerin and Musy, 2000]. In fact, due to the environment's complexity, there exists no generic model for assessing the infinite variety of spatial processes. Hence, only specific models may be implemented, which cover the decision problem at hand.

LUMASS was developed primarily as a planning tool in the area of soil and watercourse protection. It focuses on model-based prediction of soil erosion risks on agricultural land and on estimating the erosion related transport of sediments into the surface water. Additionally, it offers further modelling capabilities for assessing land use impact on the landscape budget (cf. table 1). To be applicable as a tool for supporting decisions in the planning of concrete measures, it was designed with the following additional requirements in mind: (i) it must be possible to carry out query, calculation and assessment operations at the level of individual parcels, (ii) the system must have flexible parameter settings for land use and cultivation scenarios, and (iii) the geometry of parcels and other areal and linear landscape elements must be freely changeable (e.g. roads, ditches, erosion prevention strips) via the user interface.

For the purpose of spatial data processing, analysis and presentation, LUMASS is implemented as an extension to the commercial GIS package ArcGIS™. Thus, LUMASS provides the user with the functionality of a full featured geographical information system.

The decision support methods of SDSS, which are tailored to land management, predominantly focus on the following issues: (i) resource allocation among existing parcels (i.e. spatial alternatives) [e.g. Janssen, 1996], (ii) resource allocation including the generation of new spatial entities [e.g. Aerts and Heuvelink, 2002; Tourino et al., 2003], and (iii) site selection (i.e. generation of suitability maps) [e.g. Joerin and Musy, 2000]. Typical decision support methods used to address the aforementioned decision problems may be summarized as follows (in the same order): (i) linear programming, (ii) search heuristics (e.g. simulated annealing), and (iii) multicriteria analysis.

With regard to LUMASS, the challenge is to allocate the specified land use options among the existing pattern of parcels with a view to minimizing landscape impacts (e.g. minimizing soil erosion and sediment discharge) whilst taking into account specified area shares of land use options. For this, LUMASS provides the module "Multiobjective Optimization", which provides the user with an easy-to-use user interface for specifying the optimization problem. It maps the spatial allocation problem into the linear program format required by the underlying solver library (lp_solve) [Berkelaar et al., 2004]. If a solution is found by the solver, LUMASS displays the optimization result as a polygon layer within the ArcMap™ GIS environment. Due to the generic implementation of the optimization module, it may also be utilized for any areal related resource allocation problem in general. Thus, other criteria besides those focused on here (e.g. socioeconomic criteria) may be taken into account within the optimization procedure.

Table 1. Functional components and features of LUMASS.

Component	*Features*
Modelling Analysis and assessment of landscape processes	Simple and complex relief parameters (e.g. specific catchment area)
	SCS curve number (e.g. direct runoff per parcel, cascaded direct runoff)
	Soil erosion by water (USLE/RUSLE approach)
	Surface matter transport including localization of discharge points
	Soil water budget (e.g. groundwater recharge, usable field capacity)
	Soil compaction (including pressure propagation within the profile)
	Additional soil parameters (e.g. effective rooting depth)
GIS	Input, output, analysis, visualization, etc.
DS	Multiobjective optimization of areal resource allocation

3.2 Structural Characteristics and Integration

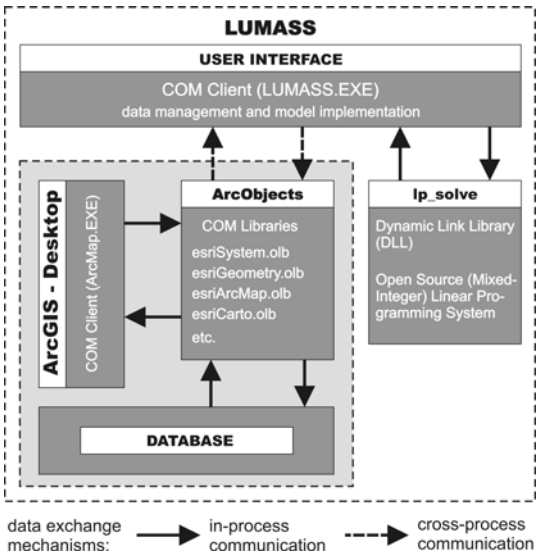


Figure 2. Structural components of LUMASS.

In order to fulfil the general functional requirements of SDSS as described in section 2.1, LUMASS follows a tight coupling strategy (cf. figure 2). It integrates five different software components and applications: (i) the LUMASS user interface, (ii) the ArcMap™ GIS application, (iii) the ArcObjects™ libraries, (iv) the geographical database, and (v) the Open Source Mixed-Integer Linear Programming System lp_solve [Berkelaar et al., 2004]. In this the LUMASS user interface, which is implemented as Visual C++® executeable, plays the key role in integrating the aforementioned software components. On the one hand, it serves as central user interface for configuring the spatial database and model specific parameters and for modelling the spatial allocation problem. Further, it implements the logic of the integrated spatial process models. On the other hand, it

manages the bidirectional data exchange between the LUMASS user interface and the ArcMap™ GIS application as well as the communication between the LUMASS user interface and the lp_solve library. For the communication with the GIS, the LUMASS user interface utilizes the ArcObjects™ COM (Microsoft Component Object Model) interface to access the spatial data, which is actually loaded into the ArcMap™ GIS application. Since the decision support component lp_solve is available as dynamic link library (DLL), the data exchange with the LUMASS user interface is implemented on the Visual C++® level.

4 LAND USE OPTIMIZATION USING LUMASS

The application of the LUMASS module "Multiobjective Optimization" is demonstrated using a rather simple but comprehensible optimization problem. The task is to optimize the land use pattern of the investigation site (cf. figure 3) with a view to minimizing the overall soil erosion. Therefore, two scenarios, which differ in the area shares allotted to the land use alternatives (cf. table 2), are investigated.

Table 2. Scenario definition of the sample application.

Scenario [1)]	WW-WB-R (c-factor: 0.072)	Winter wheat (c-factor: 0.115)	Maize (c-factor: 0.505)	Pasture (c-factor: 0.004)
1	≥ 40	≥ 20	≥ 6	≥ 9
2	≥ 35	≥ 15	≥ 6	≥ 20

[1)] the values represent the specified area shares as percentage of the area of the investigation site
ww: winter wheat, wb: winter barley, r: rape

Here, the disposition to water driven soil erosion of each land use alternative is expressed in terms of the c-factor. A high value denotes a high disposition to soil erosion and vice versa. Then for each optimization criterion (here: soil erosion), each spatial alternative has to be evaluated with respect to the given land use alternatives. In case of LUMASS, the integrated soil erosion model is used to assess the criterion scores as input for the optimization module. Figure 3 (left) maps the results of the model run in terms of a winter wheat–winter barley–rape rotation, indicating the parcel specific potential erosion risk.

When solving the optimization problem according to scenario 1 (cf. table 1), it becomes evident that the optimization procedure assigns the land use alternatives that exhibit a relatively high disposition to soil erosion to those parcels showing a relatively low potential erosion risk and vice versa. When the area shares are adjusted in accordance with scenario 2, the optimization module produces a similar land use pattern (cf. figure 3, right), except that the overall area of parcels providing a relatively high potential erosion risk has decreased in size due to the higher percentage of pasture.

5 CONCLUSIONS

Many articles have been published covering different aspects of the design and implementation of SDSS. Together, they provide valuable clues of functional and structural demands on the development of SDSS. However, only a limited number of SDSS fulfil these strict requirements. The introduced land use management support system (LUMASS) was developed in accordance with the functional and structural requirements of SDSS, thus representing a ready-to-use planning tool even for smaller private planning offices as well as for agricultural consulting agencies. With respect to its modelling capabilities, it includes rarely found methods such as the localization of non-point source sediment discharge. Furthermore, it assists the user in finding an optimum land use pattern with a view to minimization of landscape impacts whilst taking into account specified area shares of specified land use options. The sample application with the objective of minimizing soil erosion and sediment discharge in the investigation area showed plausible results for both, the modelling component as well as the optimization module.

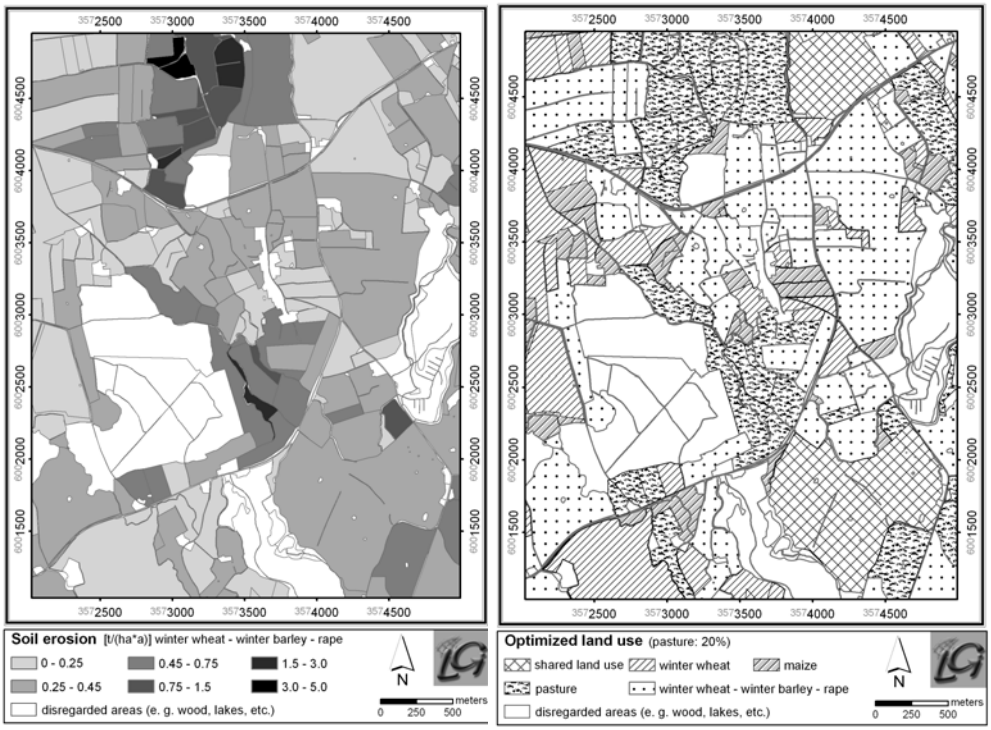


Figure 3. Left: soil erosion modelled for a winter wheat-winter barley-rape rotation; right: land use pattern optimized to minimize soil erosion and taking into account the given area shares (scenario 2).

REFERENCES

Aerts, J.C.J.H., Heuvelink, G.B.M., Using simulated annealing for resource allocation, *International Journal of Geographical Information Science*, 16(6), 571-587, 2002.

Berkelaar, M., Eikland, K., Notebaert, P., lp_solve version 5.1 - open source (mixed-integer) linear programming system, http://groups.yahoo.com/group/lp_solve (date of access: 04.01.2005), 2004.

Carver, S., Integrating multi-criteria evaluation with geographical information systems, *International Journal of Geographical Information Science*, 5(3), 321-339, 1991.

Densham, P.J., Spatial decision support systems, Maguire, D.F., Goodchild, M.F., Rhind, D.W. (Eds.), Geographical Information Systems, Volume 1, 403-412, 1993.

Denzer, R., Generic integration of environmental decision support systems – state-of-the-art, *Environmental Modelling & Software*, 20, 1217-1223, 2005.

Djokic, D., Toward a general-purpose decision support system using existing technologies, Goodchild, M.F., Steyaert, L.T., Parks, B.O., Johnston, C., Maidment, D., Crane, M., Glendinning, S. (Eds.), *GIS and Environmental Modeling: Progress and Research Issues*, 353-356, 1996.

Eldrandaly, K., Eldin, N., Sui, D., A COM-based spatial decision support system for industrial site selection, *Journal of Geographic Information and Decision Analysis*, 7(2), 72-92, 2003.

Fedra, K., Reitsma, R.F., Decision support and geographical information systems, Scholten, H.J., Stillwell, J.C.H. (Eds.), *Geographical Information Systems for Urban and Regional Planning*, 177-188, 1990.

Fedra, K., Distributed models and embedded GIS: integration strategies and case studies, Goodchild, M.F., Steyaert, L.T., Parks, B.O., Johnston, C., Maidment, D., Crane, M., Glendinning, S. (Eds.), *GIS and Environmental Modeling: Progress and Research Issues*, 413-417, 1996.

Geoffrion, A.M., Can OR/MS evolve fast enough? *Interfaces*, 13, 10-25, 1983.

Jankowski, P., Integrating geographical information systems and multiple criteria decision-making methods, *International Journal of Geographical Information Science*, 9(3*)*, 251-273, 1995.

Janssen, R., Multiobjective decision support for environmental management, Environment & Management, 2, Dordrecht, 1996.

Joerin, F., Musy, A., Land management with GIS and multicriteria analysis, *International Transactions in Operational Research*, 7, 67-78, 2000.

Johnston, K.M., Geoprocessing and geographic information system hardware and software: looking toward the 1990s, Scholten, H.J., Stillwell, J.C.H. (Eds.), *Geographical Information Systems for Urban and Regional Planning*, 215-227, 1990.

Kaur, R., Srivastava, R., Betne, R., Mishra, K., Dutta, D., Integration of linear programming and a watershed-scale hydrologic model for proposing an optimized land-use plan and assessing its impact on soil conservation – a cae study of the Nagwan watershed in the Hazaribagh district of Jharkhand, India, *International Journal of Geographical Information Science*, 18(1), 73-98, 2004.

Lam, D., Pupp, C., Integration of GIS, expert systems, and modeling for state-of-environment reporting, Goodchild, M.F., Steyaert, L.T., Parks, B.O., Johnston, C., Maidment, D., Crane, M., Glendinning, S. (Eds.), *GIS and Environmental Modeling: Progress and Research Issues*, 419-422, 1996.

Li, L., Wang, J., Wang, C., Typhoon insurance pricing with spatial decision support tools, *International Journal of Geographical Information Science*, 19(3), 363-384, 2005.

Makowski, M., Wierzbicki, A.P., Architecture of decision support systems, Wierzbicki, A.P., Makowski, M., Wessels, J. (Eds.), *Model-Based Decision Support Methodology with Environmental Applications*, 47-70, 2000.

Malczewski, J., GIS and multicriteria decision analysis, New York, 1999.

Mendoza, G., Anderson, A.B., Gertner, G.Z., Integrating multi-criteria analysis and GIS for land condition assessment: Part I - evaluation and restoration of military training areas, *Journal of Geographic Information and Decision Analysis*, 6(1), 1-16, 2002a.

Mendoza, G., Anderson, A.B., Gertner, G.Z., Integrating multi-criteria analysis and GIS for land condition assessment: Part II - allocation of military training areas, *Journal of Geographic Information and Decision Analysis*, 6(1), 17-30, 2002b.

Ostfeld, A., Muzaffar, E., Lansey, K., HANDSS: The hula aggregated numerical decision support system, *Journal of Geographic Information and Decision Analysis*, 5(1), 16-31, 2001.

Poch, M., Comas, J., Rodríguez-Roda, I., Sànchez-Marré, M., Cortés, U., Designing and building real environmental decision support systems, *Environmental Modelling & Software*, 19, 857-873, 2004.

Reitsma, R.F., Carron, J.C., Object-oriented simulation and evaluation of river basin operations, *Journal of Geographic Information and Decision Analysis*, 1(1), 10-24, 1997.

Rizzoli, A.E., Young, W.J., Delivering environmental decision support systems: software tools and techniques, *Environmental Modelling and Software*, 12(2-3), 237–249, 1997.

Simon, H.A., The new science of management decision, New York, 1960.

Tourino, J., Parapar, J., Doallo, R., Boullon, M., Rivera, F.F., Bruguera, J.D., Gonzales, X.P., Crecente, R., Alvarez, C., A GIS-embedded system to support land consolidation plans in Galicia, *International Journal of Geographical Information Science*, 17(4), 377-396, 2003.

Ungerer, M.J., Goodchild, M.F., Integrating spatial data analysis and GIS: a new implementation using the component object model (COM), *International Journal of Geographical Information Science*, 16(1), 41-53, 2002.

Wessels, J., Wierzbicki, A.P., Model-based decision support, Wierzbicki, A.P., Makowski, M., Wessels, J. (Eds.), *Model-Based Decision Support Methodology with Environmental Applications*, 9-28, 2000.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Sustainable Use of Pesticides: an Innovative Experience in Lombardy Region, Italy

C. Riparbelli [a], A. Di Guardo [b], A. Serrano [b], E. Capri [c] and S. Brenna [a]

[a] *Ente Regionale per i Servizi all'Agricoltura e alle Foreste, Milano – Italy (carlo.riparbelli@ersaf.lombardia.it)*
[b] *Informatica Ambientale, Milano – Italy (diguardo@iambientale.it)*
[c] *Istituto di Chimica Agraria ed Ambientale, Università Cattolica del Sacro Cuore, Piacenza – Italy*

Abstract: Pesticide contamination may affect the quality of different environmental compartments as soil, water and air and, consequently, may pose a risk to human beings and ecosystems.
The current environmental and agricultural policies of the European Union are aimed at the development of sustainable agriculture production. Therefore, decision makers need innovative instruments to plan pesticide use, while farmers need directions to fight pests in a sustainable way in order to maintain reasonable cost-benefit balances for the farm.
A decision support system called SuSAP Network – Supplying Sustainable Agriculture Production – has been developed for an intensive agricultural area of North Italy (Lombardy Region) to promote the sustainable use of pesticides at farm scale.
SuSAP Network is a web based software which integrates existing environmental data from different sources to calculate the environmental risk of pesticides, an innovative web service for managing model simulations and a WebGis which displays geographic data.

Keywords: Pesticide; Risk indicator; Web services; Decision support system.

1. INTRODUCTION

Pesticide may have the potential to kill or control harmful organisms such as pests, but can also cause unwanted adverse effects on non-target organisms, human health and environment. The possible risks associated with the use of pesticides are accepted to a certain extent, because of the direct benefits that it generates. These substances are employed on a large scale and generally considered as essential in modern cropping systems since they contribute to ensuring reliable supplies of affordable and healthy agricultural products of high quality. Unwanted amounts of certain pesticides can however still be found in environmental media (in particular soil and water) and residues exceeding regulatory limits still occur in agricultural produce.
For these reasons, sustainable agricultural systems, characterised by technologies and crop protection strategies, able to safeguard the groundwater resources and aquatic ecosystems, have become one of the themes of modern agriculture.
A concrete and frequently applied example of a sustainable use of pesticides could be considered the so called Integrated Pest Management (IPM), an integrated farming system which offers real solutions to many of the problems encountered on farms throughout Europe today. A quantitative reliable demonstration of the beneficial effects of IPM eco-compatible practices could lead agronomists and consultants to suggest a reduction up to 30% of the pesticides applied to the most common crops like maize, rice and vineyard. This kind of practices should results in a lower pesticide impact on the environment, but also in a money safeguard for the farmers and enhancing of eco-compatible productions.

The most recent environmental and agricultural policies (see for example [Dir. 91/414], [COM2006 373]), at national and European level, pursue these aspects and try to develop strategies integrating environmental problems in the management of the natural resources.
Following this need for the decision makers, the first objective is to define and to carry out finalised action plans at regional and local level based on real data and scenarios.
The objective of this paper is to describe a decision support system, dedicated to the extension services and farmers, called SuSAP Network – Supplying Sustainable Agriculture Production – developed within an intensive agricultural area of North Italy (Lombardy Region) to promote the sustainable use of pesticides at farm scale.
At present the system takes into account the comprehensive information on the life cycles of pests and their interaction with the environment; it is expressly be tailored to address the IPM practices, characterized not only by a single pest control method but, rather, by a series of pest management evaluations, decisions and controls. The extensive use of SuSAP will suggest to the extensions services and to the farmers a pool of technically-based choices for pesticide treatments on the main crops of the region.

2. SuSAP NETWORK, THE SYSTEM

SuSAP Network has been conceived as a regional service and therefore it presents an easy-to-use web interface which exposes several functionalities to the users, mostly farmers and agronomical technicians. The main topic we discuss here relates to the possibility to find the best crop protection strategies (Figure 1) using different pesticide application scenarios and simulating their behaviour in the environment by the means of internationally acknowledged models.
SuSAP Network uses these different layers of data to feed its coupled models:

- at predefined time steps it acquires the regional cadastral map and all the related information, supplied by the Agriculture Information System of the Lombardy (SIARL);
- a layer of soil data at scale 1:50.000;
- 12 years of weather data for 13 meteorological station spread over the regional territory;
- phenological data and default irrigations for the main crops cultivated in the Region;
- a complex sub-database of pesticides and active ingredients constantly updated by an external source.

Each pesticide application scenario built with the information above can be simulated to provide an estimate of the feasible impacts to the environmental compartments considered - soil, groundwater, surface water, air - and to provide a set of recommendations. [Riparbelli *et al*. 2006].

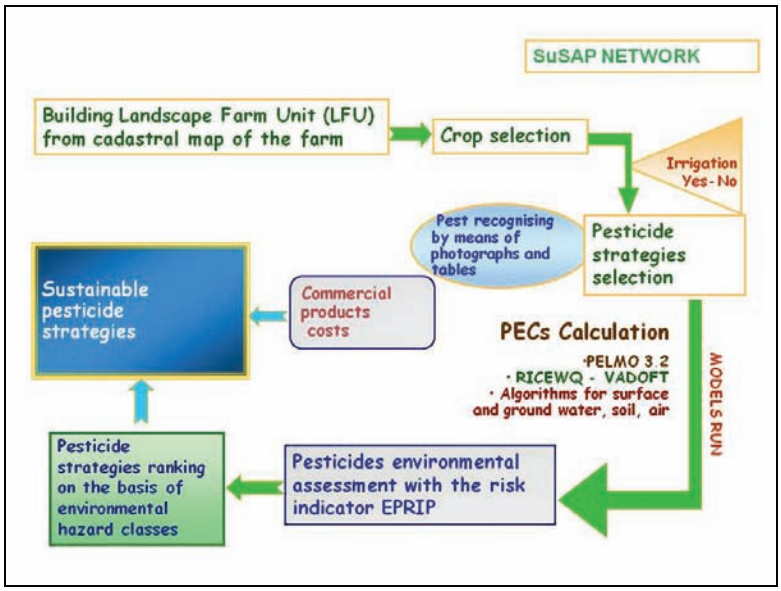


Figure 1. SuSAP Network flow charts.

Calculation is performed according to the spatial variability of soil properties, meteorological conditions, crops and patterns of pesticide applications for a user-defined homogeneous area - Landscape Farm Unit – (LFU) derived by the integration of the cadastral map and the cartographic databases of the region (Figure 2).

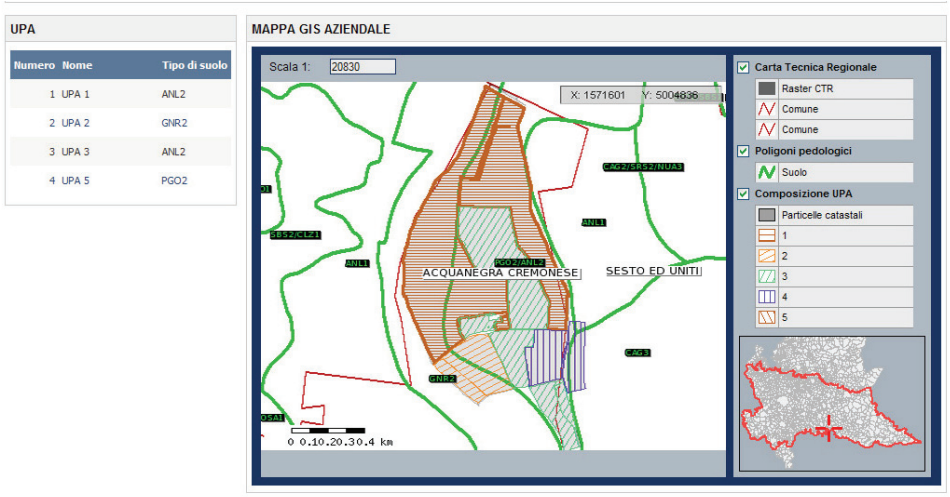


Figure 2. Visualisation of Landscape Farm Units (UPA in Italian) in a farm with SuSAP Network.

The current release of SuSAP Network calculates the Predicted Environmental Concentrations (PECs) of the used pesticides for each environmental compartment by means of a two steps procedure; it calculates at first the EPRIP index using EPRIP 2.1 [Padovani et al. 2004] [Trevisan et al. 2007], which rapidly estimates the overall impact of the pesticide to the environmental compartments and provides an estimation of the environmental risk using six classes (none, negligible, small, present, large and very large). The user can later refine the EPRIP evaluation performing a model simulation of pesticide leaching to groundwater using the one dimension models RICEWQ-VADOFT 2.0 [Williams et al. 1999] [Capri et al. 2002] for the rice flooded crop or PELMO 3.2.2 (PEsticide Leaching MOdel, [Klein *et al.*,1995]) for the other crops. The two models have been integrated in SuSAP to estimate the pesticide leaching potential through the soil horizons based on an extended cascade model; moreover, RICEWQ calculates the runoff of pesticide to surface water for flooded rice; these processes include estimation of soil

temperatures, pesticide degradation, sorption, volatilisation, and potential evapotranspiration.
The above models have been selected among those indicated and validated by the FOCUS (FOrum for the Co-ordination of pesticide fate models and their USe), jointly established since 1993 by the European Commission and the European Crop Protection Association to provide guidance to the Member States in EU pesticide registration process.
At the end of the simulation process, the user can compare the predicted environmental risk, the efficacy of the treatments simulated and the costs of each plant protection product and therefore evaluate the sustainability of crop protection strategies. Hence using their own databases as well as their pesticide plans, farmers can identify the best calibrated strategies, since all the available data in SuSAP Network have been validated and is up-to-date.

3. THE SOFTWARE ARCHITECTURE

SuSAP Network is a web portal that follows the standard n-tier architecture that guarantee the higher level of software modularity, allowing any of the tiers to be upgraded or replaced independently as requirements or technology changes. The architecture is composed of three main tiers: the user interface tier ("presentation"), the functional process logic tier ("business rules") and the data access tier ("computer data store"). The business rules tier is composed by the application logic and two additional levels conceived as web services: an environmental data provider service and an easy-to-use model engine service. The next sections will describe each of them.

3.1 The Web Portal and the Database

The web portal has been developed using Microsoft ASP.NET 2.0, enhanced with the AJAX extension library v1.0 whenever possible to provide fast and smooth server response to user interactions. Using the portal, the user can describe and configure different pesticide application scenarios that can be simulated in order to evaluate the risk of the selected pesticides to the environmental compartments and to provide a set of recommendations. Pesticides can be selected from the set of those registered in Italy, information that is kept up to date by the means of an online interaction with the pesticide database of the ICPS Institute (International Centre for Pesticides and Health Risk Prevention).
The portal allows the users to keep a pesticide treatment notebook that can be used to generate a yearly report due by the current legislation.
Coupled to the web portal, the user can find a WebGIS client that serves as an interactive tool to visualize and select several cartographic parameters, such as the soil type or the fields to analyze, the cadastral spatial information of the "in use" farm from where he can create the LFU. The cadastral and farm GIS layers used to by the WebGIS client are constantly updated by SIARL (Sistema Informativo Agricolo della Regione Lombardia) and are linked to detailed soil and climatological region-wide informational layers.
The core of the system is the database, which stores the data and defines the complex relationships relating the different datasets.
Basic soil data have been surveyed at scale 1:50,000 (Figure 3) where several representative soil profiles have been characterised by field and laboratory analysis in order to collect specific hydrological parameters necessary to calibrate model applications. The collected samples were compared with data calculated by a set of pedo-transfer functions (PTF); the best fitting PTF was used to estimate field capacity, wilting point [Rawls *et al.*, 1982] and bulk density [Manrique & Jones, 1991], necessary for the models execution.
Rainfall, temperature, wind, solar radiation, relative humidity and evapotranspiration data from 125 meteorological stations covering the Lombardy region were collected and interpolated with geostatistical methods to obtain spatial patterns of rainfall and temperature. The integrated analysis of soil and meteorological data allowed us to identify 15 macro-areas, each one defined by a reference meteorological station.

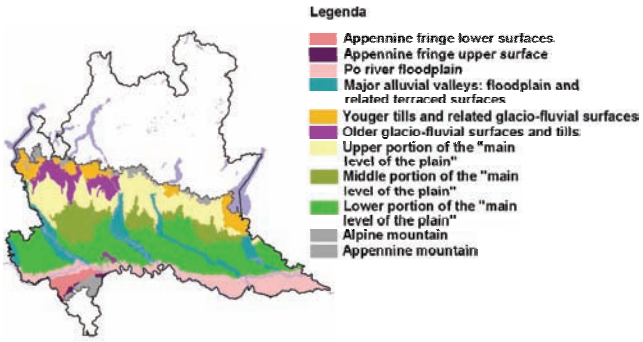


Figure 3. A map layer example: the soil mapping units.

More than 500 pesticides, including their chemo-physical properties, ecotoxicological data and the related plant protection products along with their information on the date, number and rate of application and the efficacy against pests have been stored.
The physiological characteristics of each crop, its phenological stages, weeds and pests affecting it, complete the SuSAP database. Several set of irrigation amounts, dates and types have been calculated by means of CropSyst 3.04 model [Stöckle *et al.*, 1999], applied to 10 soil scenarios representative of the different soil and weather conditions of the Lombardy Region.

3.2 The Model engine service

In order to build a scalable and powerful system and to speed up the evaluation process, an innovative simulation engine service (Web Environmental Model Simulation Service – WEMSS) has been introduced. It is a separated and independent environmental model manager, capable to expose its functionalities through a remotely located web service, accessible via the Internet and the TCP/IP protocol. WEMSS permits to aggregate several environmental models (related to a single compartment or to different ones) providing a shared infrastructure to manage the input variables, the output ones and the internal parameterization, by the means of XML configuration files.
Since each model uses a custom set of environmental variables, the first need was to harmonise them among models. To accomplish this, an ontology structure [Athanasiadis et al., 2006] has been developed that unambiguously define the environmental concepts used throughout the software. Since an ontology can be defined as a representation of a set of concepts within a domain and the relationships between them, we defined for each concept a descriptor that qualify it with important information like its validity range, its unit of measure and its default value (among others).
For each concept, WEMSS provide a generic descriptor, which exposes validity ranges foreseeable in real world. The modeller can then modify these descriptors adjusting their values to the needs of the model used or to the location in which they will be applied (for example the temperature in the real world ranges generally between -70 °C and 60 °C, but a particular model can treat only values between -10 °C and 40 °C).
Since all of the ontology concepts are stored in xml files (see an example in Fig. 4), it is simple to publish their values and descriptor in different formats using XSLT transformation files.
The structure of the WEMSS web service provides a set of functionalities that can be used independently of the models it exposes, hence providing a common and useful context that simplify the interactions with the models. This aspect is really important because it allows the user to focus more on the environmental aspects and the quality of the input data rather than arranging and organizing these input data to satisfy the model specifications, operations transparently accomplished by the service. Each model in facts interacts with the WEMSS infrastructure by the means of a specific wrapper class, which implements some functionalities (web methods) inherited by an interface of the model engine. These methods provide a common way to read and manage input values from external sources and to provide output values requested by the user. Moreover, each wrapper class implements a

"validate" method that uses the variable descriptors to check the data structure values for consistency, and an "execute" method that prepares the input data for the wrapped model and run it each time it is needed.

```
<?xml version="1.0" encoding="utf-8"?>
<VariableDescriptorsSet Name="Generic">
  <Variables>
    <Variable type="Double"  Name="Temperature" Description="Soil temperature"  Nullable="false"
      MinValue="-30" MaxValue="100" DefaultValue="20" UnitOfMeasure="°C" />

    <Variable type="Int" Name="GiulianDay" Description="Day of year in giulian format"  Nullable="false"
      MinValue="0"  MaxValue="366" DefaultValue="0" UnitOfMeasure="day">

    <Variable type="String" Name="Speed" Description="Qualitative speed" Nullable="false"
      DefaultValue="Medium" UnitOfMeasure="">
      <Values>
        <Value>VeryFast</Value>
        <Value>Fast</Value>
        <Value>Medium</Value>
        <Value>Slow</Value>
        <Value>VerySlow</Value>
      </Values>
    </Variable>
  </Variables>
</VariableDescriptorsSet>
```

Figure 4. An example of a three generic variable descriptor

When a simulation requests a model calculation, the SuSAP Network application fills the input data structure provided by the model wrapper class, sets the correct parameters and instantiate a new executor class; when the model terminates its calculations, the wrapper class fills the output data structure of the model and make available the results to the caller. Using this approach the models can be piped and combined to create unique composed models to get better results. The WEMSS can automatically map the output variables of one model to the input variables of the next one using the variable descriptors for each piped model. For now just linear pipes are allowed, but as a future enhancement nonlinear, conditional and cyclical model executions will be permitted, as well as the possibility to persist these created pipes to be reused at future executions.
Each model run is executed in a separated pool providing the facility to better monitor the engine (for example making error checking and exception handling) and to parallelize the computational work by the means of an executor's grid architecture (Grid Computing) , for example in regional-wide applications.
WEMSS has been used as the model engine for the SuSAP Network portal: the two models PELMO and RICEWQ/VADOFT are inserted into the framework and their functionalities can be reached through the common interface provided.

3.3 Environmental Data Web Service

Linked to WEMSS, an environmental data web service has been developed to store and provide soil and weather data.
The web service is simply a remote repository and it has two main functionalities: to increase its capacity in terms of environmental information and to provide data to external users. To accomplish this, the service uses an internal structure optimised for massive saving and loading operations of environmental data, in particular time series. The data can be manipulated by external users using a set of web methods; Figure 5 shows two of these web methods related to the meteorological data supply (ImportMeteoFile) and data retrieve (GetMeteoFile). Each operation provides several informations to the web service by the means of an XML file:

- the meteo station;
- the data types of the variables to be supplied or retrieved;

- the format of the file provided of returned;
- the initial and ending dates;
- the time step;

The access to the repository can be filtered using a strong authorisation mechanism to guarantee the protection of the stored data and its authorized use.
The web service can be coupled with external software components in order to offer data generation capabilities: for example the new PTF component [Acutis et al., 2008] recently made available can be used to generate soil structure data using the internal set of Pedo-Transfer functions and provide them in real time to the user.

```
/// <summary>
/// Imports a meteo file containing meteo series into
/// the internal storage.
/// </summary>
/// <param name="meteoFile">Lines of the meteo file.</param>
/// <param name="descriptor">XML descriptor of the
/// time series file.</param>
/// <param name="overwrite">true to overwrite the
/// existing data, false to leave the existing data.</param>
/// <returns>An object describing the result of
/// the operation.</returns>
[WebMethod]
public ClsFileMeteoWSResponse ImportMeteoFile(
    string[] meteoFile,
    XmlDocument descriptor,
    bool overwrite)
{
}

/// <summary>
/// Builds and retrieves a meteo file containing
/// meteo series from the internal storage.
/// </summary>
/// <param name="descriptor">XML descriptor of
/// the time series file.</param>
/// <returns>An object describing the result of
/// the operation.</returns>
[WebMethod]
public ClsFileMeteoWSResponse GetMeteoFile(XmlDocument descriptor)
{
}
```

Figure 5. Web methods that manage meteorological files

4. CONCLUSIONS

The starting up of SuSAP Network service and its diffusion in the agriculture of Lombardy could assure a better environmental protection as well as a higher food safety, resulting in higher quality of agricultural productions. Actually the use of this system could give to farmers an important contribution to the certification of eco-compatible productions, in which pesticides are used under Integrated Pest Management (IPM).
These objectives are consistent with the regional strategic projects, aimed at safeguarding cultivated areas and at upgrading the quality of the agricultural environment, as is stated in the Regional Rural Development Program.
SuSAP Network users, as farmers, farm operators and local authorities will benefit when proper advices may be given to them about the use of the more environmentally safe products.
Three types of benefits are expected:

- better pesticide environmental impact assessment due to the easier access to representative data sets;

- better use of pesticides because adviser and farmers would have access to this information;
- increase of the awareness about the use of soil and climate data for pesticide use.

It is impossible to exactly compute the amount of money saved by the SuSAP Network application in Lombardy, but it will be very significant because a better use of pesticides is becoming more and more a critical point in modern agricultural systems management.

REFERENCES

Acutis M., M. Donatelli and G. Lanza Filippi, PTF: an Extensible Component for Sharing and Using Knowledge on Pedo-Transfer Functions, *These Proceedings*, 2008

Athanasiadis, I. N., A. E. Rizzoli, M. Donatelli, and L. Carlini, Enriching software model interfaces using ontology-based tools iEMSs congress, Vermont, July 2006 (*in press on Int. J. of Adv. Systemic Studies*)

Capri E., D.G. Karpouzas and E. Papadopoulou Markidou, Application of the RICEWQ-VADOFT model to simulate leaching of propanil in rice paddies in Greece, *Agronomy for Sustainable Development*, 25, 35-44, 2005.

Council Directive 91/414/EEC. *Official Journal of the EC*. L230, vol. 34, 1-32, 1991.

Jene B., PELMO 3.00 Manual extension, Staatliche Lehr- und Forschungsanstalt für Landwirtschaft, Weinbau und Gartenbau, Neustadt, 1998.

Manrique L.A. and C.A. Jones. Bulk density of soils in relation to soil physical and chemical properties. *Soil Sci. Soc. Am. J*. 55:476-481, 1991.

Padovani L., M. Trevisan and E. Capri, A calculation procedure to assess potential environmental risk of pesticides at the farm level, *Ecological Indicators*, 4, 111-123, 2004.

Rawls W.J., D.L. Brakensiek and K.E. Saxton, Estimation of soil water properties. *Trans. ASAE*, 25, 1316 – 1320, 1982.

Riparbelli C., M.N. Cambareri, M. Pastori and S. Brenna, Water quality management in agriculture: the case of lombardy - North Italy. *Management of Intentional and Accidental Water Pollution - 2005, 8 - 11 May, Sofia, Bulgaria. NATO Security to Science Series: Environement series. 211-230 pp.* G.Dura et al. (eds.). Springer NL, 2006.

Stöckle C. & R. Nelson, CropSyst 3.04. Department of Biological Systems Washington State University USA, 1999.

Commission of the European Communities Proposal for a regulation of the European Parliament and of the Council concerning the placing of plant protection products on the market Thematic strategy on sustainable use of pesticide, COM(2006) 388 final, Brussels, 2006

Trevisan M., A. Di Guardo and M. Balderacchi, EPRIP 2 - Environmental Potential Risk Indicator for Pesticide to drive sustainable pest management practice, *Environmental Modelling & Software*, submitted, 2008.

Williams W., A. M. Ritter, J. M. Cheplick and E. C. Zdinak, *RICEWQ: Pesticide Runoff Model for Rice Crops***,** Waterborne Environmental. Inc**.,** Leesburg, VA, 1999.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Influencing Factors for Scheduling Optimal Pipeline Replacement with Budgetary Constraints

C.D. Alonso[a], R. Pérez[a], P. A. López[a], X. Delgado[a]

[a]Multidisciplinary Center Modeling Fluid (MCMF) – Technical University of Valencia-Camino de Vera s/n. Valencia (España) Teléfono +34.96.387.98.90
E-mail: cdalonso@posgrado.upv.es.

Abstract. We propose an instrument to aid in the decision-making process that allows for the prioritization of needs for the water distribution network renewal planning and to identify these priorities applying the methodology Multicriteria Decision Support Systems (MDSS) and Genetic Algorithms. We consider the various influencing factors involved in decision-making. The instrument developed proposes an acceptable sequence of interventions in the water network, identifying the hierarchy of the groups of pipes before renovation work beings and takes into account all factors associated with the endogenous and exogenous factors and the technical constraints of functioning of the water network for a given planning horizon.

Keywords: *Water distribution network, renewal, hierarchy, decision-making.*

1 INTRODUCTION

Identifing the factors that influence the determination for prioritization in the renewal of the pipes is not an easy task, given that the selection should support the fundamental criteria at the time the pipe is considered for renewal. Age is undoubtedly the first criterion for inclusion in the list, but not the only one. Coupled with age is the availability of financial resources required and the financial resources available to water companies that operate water distribution networks. The aim is to improve the reliability and operation of the network. The emphasis on the renewal of the pipes is done according to a decision-making process that takes into account a number of variables related to the pipes and their environment, the conditions of operation of the network, technical restrictions and of course the financial implications.

2 OBJECTIVES AND SCOPE

The main objective is to study the factors of influence on the deteriorating pipes, in order to propose a model for the scheduling of work on the prioritization of the renewal of pipelines. We take into consideration a number of factors in order to prioritize of a group of pipes with similar characteristics including age and type of material. We also consider the annual budget amounts available for the rehabilitation of pipelines. This requires the development of the proposed methodology for the allocation of priorities for renewal of pipes, using the system to support the decision making of the weighted sums, detailing the weighting matrix of alternatives and evaluation of results. Once addressed the methodology referred, it was applied the method of weighted sums for the prioritization of renewal in eight groups of piping within the water distribution network in Celaya City, Mexico, where the grouping of partial scores of each pipe is made with respect to each of

the criteria obtained as a result of prioritizing the group of pipes. The results of sensitivity analysis are applied before changes in the consideration of the values on weights of the criteria. Finally we apply the method for ranking each of the groups of pipes for each factor considered applying Genetic Algorithm.

3 ELECTION OF THE CRITERIA

The selection of criteria that can really influence in the decision to replace or not to replace a pipe or to position it on a priority list is not immediate. This is undoubtedly a particular decision of the managers of each network. They must choose the criteria more representative in their environment and make an assessment in accord with the reality they live, depending of course, on their technical expertise of the business. Based on the above, here is the classification of the eight criteria that in the view of the network manager in Celaya City deserved to be considered. The criteria are classified into groups with different characteristics, but they allow for a quantitative assessment of the item analysis. They are the "Fundamental" group , "Hydraulic" group and the "Specials" group.

3.1 Fundamental criteria

The age of the pipe.- A network updates its status by changing its parameters, not just physical (the most important: internal diameter and roughness), but also hydraulics, both to adapt to new conditions imposed by the quality of water, and as a result of variations in changes in the demands occurring by the growth of the population served, bring with them a degree of inevitable ageing. Thus, the aging is natural and predictable to a certain extent, and thus should be valued.

History of each pipe breakage. -Here we evaluate the number and frequency of breaks that occurred over a period of time, that arise as a result not only of the age of the pipe, but also other factors that make two pipes with the same age have different ratios in occurrence of leak.

3.2 Hydraulic criteria

Transport capacity of each pipe.- Based on the theoretical flow of the pipeline, both in its state of "new pipeline", as in the current state (with some reduction in diameter) one calculates the loss of hydraulic capacity as the difference between the transport capacity "new" and transport capacity "today".

Index deficiency nodes.- The index reflects the impact of the unavailability of a pipeline in water supply to subscribers. Allows censoring all the consumer nodes where the supply is not assured (Wagner et al, 1988). We assume that under a certain pressure P_{low} the water supply is not assured. For each pipe "cut" there is a calculation of the pressure with the help of EPANET[1] ®, and we can then perform a comparison with the pressure P_{low}. The lowest threshold allowed for Celaya City is 10 mwc.,

Impact flow escaped by the behavior of the network, compared with an optimal level of efficiency.- It is a fact that in any water distribution system an unknown amount of water that is lost. This "uncontrolled" water is not lost through only leakage (although mostly), but is also due to other causes such as measurement errors and consumption not measured and / or unauthorized.

[1] Developed by EPA's (U.S.A) Water Supply and Water Resources Division, EPANET is software that models water distribution piping systems, EPANET is public domain software that may be freely used.

3.3 Special criteria

Optimal time to replace the pipe according to Shamir-Howard.- This approach considers the balance between the cost to repairs breaks and the cost of replacing the pipe, which, with the passage of time increases the occurrence of breaks.

Repairing all the breaks increases the cost of maintaining each pipe over time.
That rate increases to the point when it is economically more profitable to replace the pipes rather than continue with the repetitive process of repetitive "unoperability-repair-operability". Based on the theory presented by Shamir and Howard (Shamir and al, 1979) you find the optimal year of renewal, Being penalized pipe that has a renewal date earliest, from the reference year.

Influence of a water "cut" in different types of users of the water distribution network.- Subscribers directly connected to the pipe cut will have no water, while the others suffer the effects of lower water pressure, at various periods of time (from a few hours to a few days). It is therefore the intent of this section to quantify these effects before cutting service in the sector by closing the flow in each of the pipes studied especially the number of users considered "major" by the water company "JUMAPA".

Impact of a break in the environment.- The assessments about the importance of the traffic disruptions caused, by the disruption of the working hours of the affected community, are carried out using qualifications type as "strong", "moderate", "weak" and "extreme". There are, however, other effects that are more precisely quantifiable: economic damages above and below commercial premises, replacing the pavement, among others. It is thus a little more natural to think of the possibility of assigning numerical weights to each of these qualification types than was previously possible. Figure 1 shows the outline of the hydraulic model of the water distribution network on one of the sectors which corresponds to downtown.

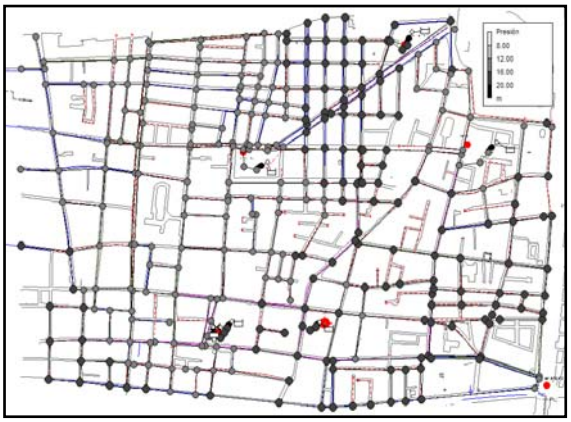


Figure 1. Outline hydraulic model

4 EVALUATION OF PIPES SELECTED ACCORDING TO THE CRITERIA DEFINED.

It comes with mathematical calculation that allows you to obtain a rating for each criterion of the eight groups of pipes on the same scale. Once the introduction of the selected criteria is realized, then you can evaluate the current status of the water network in the city of Celaya.

4.1 Fundamental criteria

The age of the pipe.- Is proposed a maximum value of 10 points for a pipe 50 years old and a minimum of 0 in the case of a new pipe (0 years). The rating associated with each age, as expressed in the equation 1 has been made assuming a parabolic relationship between the two ends. It seems logical to penalize in an increasingly proportionate way to the age of the pipe.

$$Value\,(age) = 10 \cdot \left(\frac{t}{60}\right)^2 \qquad (1)$$

Table 1. Factor values

Pipe ID	Lenght m	Age	Number of breaks (breaks/ Km. year)	Interpo-lation
AC 1	12605	40	4.09	10.00
AC 2	7243	30	3.77	9.18
Fo Fo 1	11613	50	3.44	8.36
Fo Ga 1	8600	50	3.79	9.24
Fo Ga 2	7540	40	4.06	9.92
PVC30	25201	30	2.81	6.77
PVC20	11548	20	3.00	7.25
Polythene	13865	10	3.73	9.09

History of each pipe breakage.- The number of breakages is proportional to the length of the pipe, for this reason the number of breaks is estimated per km., and the year of each pipe as a fundamental value to calculate the penalty (table 1). The values of breaking/ Km. and year are high, but for various reasons are considered high. We propose a linear interpolation where we assign a value of 10 to the maximum number of breakages and assigned a value 0 for the minimum number of breakage per km per year (Hirner, 1997).

4.2 Hydraulic criteria

Transport capacity of each pipe.- In this first hydraulic criterion it is desired to value the hydraulic capacity of the pipe regardless of their behavior and real possibilities in the network. Knowning both the pipes diameter and roughness characteristics, and that the roughness of the pipe increases with the passage of time as well as other features such as type of material, conducted water flow quality, etc. it decides the absolute roughness for a given moment of time through the following equation.

$$k_{si}(t) = k_{soi} + a_i(t + g_i) \tag{2}$$

Moreover, the actual diameter of the pipe for a given moment in time is calculated by the equation:

$$D_i(t) = D_{oi}(1 - red_i(t + g_i)) \tag{3}$$

With the formulation taken from the previous model posed by Sharp and Walski (1998) one can determine the operation of a water distribution network for a given period of time considering the increased roughness of the pipe and obstruction by the same inlay of materials, both linear with respect to time. Based on the above we can calculate the pressure loss per km. of pipe anytime, through a simulation model using EPANET ®.

Index deficiency nodes.- This procedure makes it possible to identify the group of affected nodes. Once these nodes are identified, one calculates each pipe cut and prepares the report between the number of non-communicating nodes and the total number of nodes that make up the network. There will be a penalty for the number of pipes that are below the lowest pressure allowed. The penalty was higher when the number of larger pipes fell below the minimum allowable pressure. The sector that had the highest number of "important" pipelines had the greatest penalty. Is assigned a value of 10 for the largest IDN and 0 for the lowest index, which is directly proportional to the value of the IDN. The evaluation was carried out through the expression 4:

$$IDN_j = \frac{Number\ nodes_{low}}{Total\ \ number\ nodes} \tag{4}$$

Impact flow escaped by the behavior of the network, compared with an optimal level of efficiency.- This approach influences the consumption of the affected nodes, in addition to the changes incurred in the pipe nodes analyzed where we had the leak. When comparing the simulation of EPANET ® with the 85% efficiency of rating of the physical network, it was considered a satisfactory ratio (AWWA Water Leak Detection and Accountability Committee, 2004).Comparing the simulation results with current actual data. The allocation of leakage is dependent on pressure, the same happens with the 85% efficiency level where

the escaped flow in each node is based on the pressure. The evaluation was conducted using the following expression.

$$IIQ_{leak} = \frac{Num.nodes_{lowreal} - Num.nodes_{low85\%}}{Total\ number\ nodes} \quad (5)$$

The score will be the greater value derived from the equation 5 for each group of pipe assigned the value of 10 and for the lower the value of zero, making an interpolation for the intermediate values.

4.3 Special criteria

Optimal time to replace the pipe according to Shamir-Howard..- The statistical methodology proposed here gathers pipes that have the same properties that are considered discriminatory: age (years) and size (diameter). Other factors such as corrosion, material, the method of union between pipes, the conditions of the bed of the pipe, and traffic are considered fairly homogeneous for all pipes in every sector of the Celaya network. Using a linear regression "age-size" we obtained for each group, the relationship between the number of breaks and time. For this type of analysis, Shamir and Howard (1979) proposed for the temporary development of breakages per unit length an exponential law as reflected in the following equation, where N (t) is the number of breaks / km in the year (t) and (t_0) represents the reference year.

$$N(t)_i = N(t_o)_i * e^{A_i(t+g_i)} \quad (6)$$

$N(t)_i$ is the number of breakages per unit length per year in pipe in time t (km^{-1}, $year^{-1}$), $N(t_0)_i$ is the $N(t)_i$ in the year of installation of the pipe (new pipe); t is the time in years; g_i is the age of the pipe in the present (years), and A_i is the ratio of the rate of growth of breakages in the pipe (years). Shamir and Howard (1979) derive an expression (7) to calculate the optimal timing of replacement of a pipe (t_s), adding the cost of replacing the pipe (C_s) and the costs of repair (C_r) occurred in the period $t_s - t_0$ and derive regarding t_s year to obtain the replacement year at the minimum cost. In the following equation the variable (*i*) represents the annual rate of inflation and (t_o) the reference year.

$$t_s = -g_i + \frac{1}{A_i}\ln\left[\frac{C_s + \ln(1+i)}{C_r N(t_o)}\right] \quad (7)$$

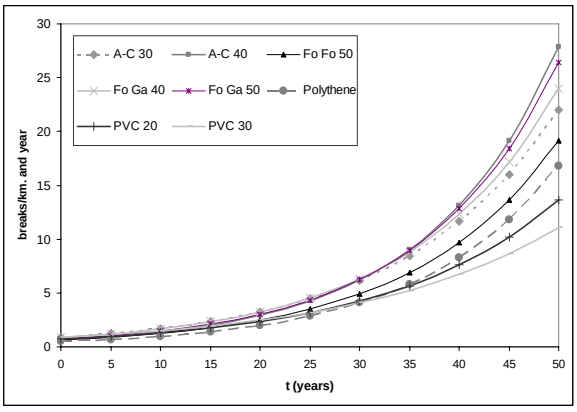


Figure 2. Prediction of breaks in the time

Now is the time to evaluate this approach. No penalty is given for pipes that have a year of optimal renewal greater than or equal to 25 years, but a penalty of 10 points is given to a value of ts - tc = 0 years, considering that the most recent pipes 20-30 years old are made with better materials except polyethylene with not very specific characteristics since its installation. the values about the prediction of breaks in the time are shown in figure. 2.

Influence of a water "cut" in different types of user in the water distribution network:.- We had considered using the JUMAPA classification for consumers called "majors" which are

users who consume at least 100m^3 monthly. In the simulated program EPANET ® (EPA, 2000) we cut each pipe by initiating main pipes at the network and analyze the number of affected nodes (at 09:00 h) when the greatest demand occurs. and the degree of influence of this court in each node, By we also measured the original pressure drop, as well as the number of affected nodes are located where the major consumers are located. Results are shown in figure 3.

	Demand	Closed pipelines					Closed pipelines				
		13				21	13				21
ID Node	>100						Presión m	Presión m	Presión m	Presión m	Presión m
node 1	0	no	no	no	no	no	6.75	0.00	12.32	14.72	14.79
node 2	0	no	no	no	no	no	6.74	0.00	12.49	14.76	14.82
node 3	0	no	no	no	no	no	6.64	0.00	12.97	14.84	14.89
node 4	0	no	no	no	no	no	6.84	0.00	13.42	15.13	15.17
node 5	0	no	no	no	no	no	6.85	0.00	13.63	15.21	15.25
node 6	0	no	no	no	no	no	7.47	0.00	11.82	15.09	15.17
node 7	0	no	no	no	no	no	7.08	0.00	11.11	14.57	14.66
node 8	100	si	si	no	no	no	8.32	0.00	11.04	15.07	15.16
node 9	100	si	si	no	no	no	8.47	0.00	11.13	15.17	15.26
node 10	0	no	no	no	no	no	16.38	0.00	15.78	16.28	16.3
node 11	100	no	si	si	si	si	12.87	7.52	0.00	0.00	0.00
node 12	0	no	no	no	no	no	12.46	7.14	0.00	0.00	0.00
node 13	100	no	si	si	si	no	13.83	8.69	0.00	0.00	15.29
node 14	0	no	no	no	no	no	14.04	9.06	0.00	0.00	14.9
Max	4	2	4	2	2	1	9	22	12	12	9
		0.5	1	0.5	0.5	0.25	0.41	1.00	0.55	0.55	0.41
		0.91	2.00	1.05	1.05	0.66					
Average	5.66	4.55	10.00	5.23	5.23	3.30					

Figure 3. Concentrate and scoring result in closure of pipeline

Impact of a break in the environment.- This factor will be the score with a valuation of 10 for a density of extreme traffic, 6 for a strong density, 3 in the case of a moderate density and 0 for a low density of traffic. A relative weight of 60% is considered for each of the 4 factors analysed. The same set of values is applied to the density of buildings and / or historic monuments, obtaining the score of each pipe.

5 METHODOLOGY

Method for prioritization pipelines renewal using- Decision Support System of the weighted sums.

5.1 Matrix weighing alternatives

Of the different approaches used to generate the set of efficient solutions, this was the first to develop. Zadeh in 1963 was the one who suggested this method. Its basic idea is to combine all the criteria in a unique role. To that end, you associate a weight or ratio of consideration to each criterion, at which time you add the value of an alternative to each criterion (adjusting these by the weights of these criteria), to generate and achieve a global solution of this alternative. Each alternative depending on the criteria evaluated, reaches a (V_{AJ}) valuation, which is affected by the weight assigned to this criterion, (W_{CrI}). Having said that, and given a set of criteria m and n alternatives, the alternative J will add m products as (V_{AJ} * W_{CrI}) represented in a summation. The matrix includes n amounts for n alternative as shown in figure 4.

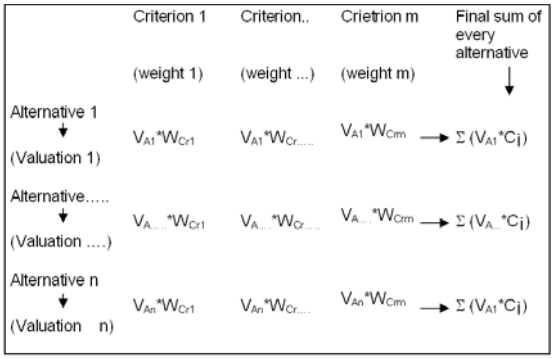


Figure 4. Typical structure of the matrix weighting alternatives

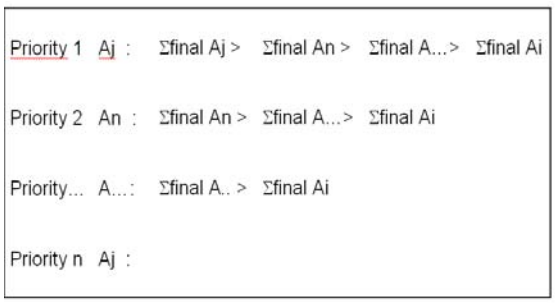


Figure 5. Prioritization of alternatives depending on the value of the sum end of each

5.2 Assessment final results

In the specific case of keeping a list of prioritized pipes ready to renew with greater urgency, given that the evaluation of each criterion implies greater need to renew when the score is higher in each case, we immediately conclude that greater sum of these products involves a greater need, and at the global level, to renew the pipe which is associated with the product. As a result, the final step in the process is to organize, from highest to lowest, each of the final results achieved by the alternatives (see figure 5), in order to know the final order for renewal.

Table 2. Order of prioritize on 100

Order of prioritize	Assessment on 100	Hierarchy
1	7.33	FoGa (C1)
2	6.71	FoGa (C2)
3	6.62	FoFo (C1)
4	6.11	AC(A1)
5	5.39	AC(A2)
6	4.80	PVC20
7	4.52	Polythene
8	4.48	PVC30

Implementation of the decision support system about weighted amounts for prioritization renewal in eight groups pipelines network Celaya City, Mexico .- We need to make some clarifications regarding the composition of the water distribution network on the diameters referred to on the majority of the pipes that make up the network, which range between 75 mm and 200 mm. The age of the pipes in the downtown sector are older (50 years) and there are different materials in the five areas such as: Fo-Fo, Fo-Ga, PVC and polyethylene. The eight groups of pipes selected were homogeneous in terms of the type of material and age.

Association of partial ratings of each pipe with respect to each of the criteria considered .- Once you evaluate the pipes for each criterion and all receive a value qualifier between 0 and 10 for each criterion, it is time to add these values, and average the weight of each criterion as described in the matrix above to achieve a final outcome for each alternative, which allows you to establish as a last step, the order of priority for renewal. Finally partial and weighted scoring of each pipe, is shown in table 2. Once you get the prioritization of the group of pipes, we first have to galvanized iron pipes that are 50 years old, followed by galvanized iron pipes 40 years old, which are older than the cast iron pipes FoFo (C1). The same is true for PVC pipes which are younger but are higher priority on this list. (i.e. PVC piping 20 years old get ratings higher than those pipes PVC30 in hydraulic and fundamental criteria.) Another significant data obtained are the very close values between the following pipes in the downtown sector: FoGa (C2) FoFo (C1), AC (A1) except the Asbestos Cement pipes AC (A2). This tells us what the homogeneous characteristics are in terms of the age and the material, and furthermore, their behaviour so far. Proof of this is the hydraulic behavior.

Table 3. Result of the sensibility analysis

Order of prioritize	Assessment on 100	Hierarchy
1	7.36	Fo Ga (C1)
2	7.07	FoFo (C1)
3	6.67	Fo Ga (C2)
4	6.11	AC (A1)
5	4.71	AC (A2)
6	4.00	PVC30
7	3.74	PVC20
8	3.54	Polythene

Sensitivity analysis of the results to changes in the consideration of the values on weights the criteria.- If the relative weight in the key considerations were interspersed values, (i.e. were rated 30% age and 20% at breakage history (original values were 20% age, 30% breakage history), another change is QII_{break} from 10% to 5%, and the special criteria that was rated at 10% optimal renewal, is now 5%. The system after these changes is reflected in table 3. Clearly, the first group of pipes to be considered for renewal is the group that corresponds to the galvanized iron pipes (FoGa 1). With the two assessments that have been made, we note that the prioritization changes considerably by changing the weight the ratings. Therefore it is important to conduct a deeper analysis of the next group of pipes to be considered for

Table 4 Assessment and scoring result of factors influence

FACTORS INFLUENCE										
ASSESSMENT	50		35			15				
	20	30	15	10	10	5	5	5		
Pipe ID	Age	Brekage history	Hydraulic capacity	NDR	QII_{break}	Times of replace	Influence break Q	Sorrounding impact		
AC 1	6,40	10,00	3,98	4,1	0	7,19	3,26	6	2	10,00
AC 2	3,60	9,19	4,64	5,8	2,45	0	3,04	4,8	7	3,04
Fo Fo 1	10,00	8,36	5,02	5,2	0,55	6,33	4,46	4,8	1	10,00
Fo Ga 1	10,00	9,25	4,73	4,6	4,62	3,85	4,57	10	8	10,00
Fo Ga 2	6,40	9,92	4,1	5,5	3,72	10	4,1	4,2	6	10,00
PVC30	3,60	6,78	4,92	3,6	3,24	0	0,12	6	3	4,92
PVC20	1,60	7,26	4,27	5,5	10	0	2,27	0	5	10,00
Polythene	0,40	9,10	2,86	10	2,23	0	1,12	0	4	10,00
										67,96

renewal, based on the factors set forth, define the pipe which has more impact on each factor individually. Based on the above, propose a method for ranking each of the groups of pipes for each factor considered, using genetic algorithms through the Evolver program. The results are shown in table 4.

6 CONCLUSIONS

Managers who make decisions on prioritization for the renewal of the pipes have a complicated task, because there are several factors of influence involved. For this reason, we believe that the method developed here can be a good tool for decision making in its simplicity. The results obtained in this study show us that the first group of pipelines on the prioritization is the group of pipe FoGa1 even with the sensitivity analysis. We note that this group of pipes score highest with established methods of scoring, Given that the factors for groups of the "fundamental" and "hydraulic" criteria obtained high values for both because of their characteristics of age and the number of breaks by the length, there is no doubt that it is the first group of pipes that should be considered for renewal. For the second and third place we will support the results of indexing and base the decision on the criteria JUMAPA decides.

ACKNOWLEDGMENTS

This article has been made possible through actions of the CMMF researchers, involved in the following project: DANAIDES: Desarrollo de herramientas de simulación para la caracterización hidráulica de redes de abastecimiento a través de indicadores de calidad del agua. REF. DPI2007-63424. Ministerio de Educación y Ciencia de España.

REFERENCES

Dridi, L, Mialhot, A., Parizeau, M., and Villenueve J. P. A strategy for optimal replacement of water pipes integrating structural and hydraulic indicators based on a statistical water pipe break model. Proceedings of the 8th International conference on Computing and control for the Water Industry, U. of Exeter, UK, September, 65-70, 2005.

Grablutz, F. and S. Hanneken,. Economic Modeling for prioritizing pipe replacement Programs. Presented at the AWWA Infrastructure Conference and Exhibition. Baltimore, MD, 14 March 2000.

Kleiner Y.; Adams; Rogers. Water network renewal planning. *Journal of Computing in Civil Engineering.* Vol. 15,No. 1. 2001.

Laky, Dora.; Water network reconstruction in Hungary to reduce number of pipe bursts. Rehabilitation management of urban infrastructure networks. Vol. December, ISBN 3 86005 430 1, 2003.

Mailhot, *et al.*, Optimal replacement of water pipes, *Water Resources Research*, vol. 39, núm. 5, may., pp. 1-13, 37, 2003.

Shamir U., and C.D.D Howard, An Analytic approach to scheduling pipe replacement, *Journal of AWWA*, Vol. 71, No 5, May, pp. 248-258, 1979.

Information about the Evolver program can be found online: http://www.palisade/educational/evolver/.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

The Emission Inventory Water, a Planning Support System aimed at reducing the pollution emissions in the surface waters of Flanders, Belgium

L. Van Esch[a], **G. Vos**[b] **and G. Engelen**[a]

a VITO, Vlaamse Instelling voor Technologisch Onderzoek (Flemish Institute for Technological Research), Mol, Belgium (leen.vanesch/guy.engelen@vito.be)

b VMM, Vlaamse MilieuMaatschappij (Flemish Environmental Agency), Erembodegem, Belgium (g.vos@vmm.be)

Abstract: The Emission Inventory Water-Planning Support System (EIW-PSS) enables the Flemish Environmental Agency to fulfil its day-to-day monitoring and reporting obligations regarding the pressures and impacts of point and diffuse emissions on surface waters vis-à-vis the Flemish, Belgian and European authorities. Moreover, it helps in the design, prioritization and evaluation of alternative measures to reduce water pollution. It consists of an advanced MS Excel/Visual Basic core intimately linked to an ArcGIS module. The latter is used for locating the sources of pollutants and their pathways to the surface waters at high spatial resolutions. Scenario and "what if" analysis featuring changes of a technical nature, as well as uses of space are supported. This paper discusses the aims of the newest version of the system, its generic approach to surface water pollution and its spatial component. Part of the discussion will dwell on a first case fully implemented namely the emission of heavy metals caused by the building stock. In particular, the heavy metals lead, zinc and copper released by corrosion in the outer skin of the buildings and in the plumbing systems are considered. Literature research and measuring campaigns were carried out to gather the required technical data alongside the development of the methodology and associated models. The outputs were calibrated and validated against independent data. The analytical capabilities of the new version, its explicitly geographical nature, as well as its direct usefulness for policy and planning support, were the main criteria for continuing its development aimed at the incorporation of more sectors and more pollutants.

Keywords: emissions; planning support systems; sewage infrastructure; surface waters.

1. INTRODUCTION

The Emission Inventory Water-Planning Support System (EIW-PSS) is foremost a water emissions accounting tool linking a wide variety of sources of pollutants to their impacts on the water quality. To that effect it couples spatially-explicitly mathematical models and their underlying databases. Pollutants considered are mostly nitrogen, phosphates, heavy metals and polycyclic aromatic hydrocarbons (PAHs). They are of concern because some compounds have been identified as carcinogenic, mutagenic and teratogenic. The EIW-PSS is developed by the Flemish Environmental Agency (VMM) to support it in its monitoring and reporting obligations vis-à-vis the Water Information System for Europe (WISE), especially under art. 3, 5 and 8 of the Water Framework Directive (WFD, 2000/60/EC). They need this tool for two reasons: firstly to determine the significant emission sources and their contribution to the pollution of the surface waters in order to formulate mitigation measures, and secondly, to control and monitor compliance with the

objectives to gradually terminate or stop and decrease pollution as defined in article 4, part 1, under a, point iv of the WFD. Instead of the emission reduction concept from the European Urban WasteWater Treatment Directive (UWWTD, 91/271/EEC), the WFD focuses on immision reduction. The proposed planning support system deals with this new concept, because it considers the pressure on the surface water and follows the pollution from source to sink. The innovative character consists of the generic approach for the quantification of emissions, discharges and losses of hazardous substances from a broad range of point and diffuse sources and the explicit geographical link between source and sink. The latter permits identifying the problematic zones on a high resolution and takes the existing inventories like in the Netherlands [Emissieregistratie Nederland, 2007] one step further [Crouzet and Bogestrand, 1999].

The current EIW-PSS is the result of experiences gained by the Flemish region of Belgium with the development of a tool meant to integrate sources and pollutants required for different reporting obligations. In 2005, the first complete emission inventory (EIW) for heavy metals was realised by Syncera water B.V. [2005] as an accounting system keeping track of the pollutants from a variety of sources towards their sinks in the surface waters. This instrument addressed the problem at a high level of abstraction and simplification. It consists of a sophisticated MS Excel/Visual Basic application and its underlying databases calculating results at the level of the so-called Smallest Geographical Units (SGU), unique polygons created as the intersection of municipalities, Water Treatment Areas (WTA) and sub-catchments. Deficiencies and unexplained loads in the results of the EIW pointed at the need for a more in-depth analysis based on a more geographically explicit approach. Thus, a new version of the EIW was developed as an operational prototype in December 2006 by the Flemish Institute For Technological Research [Engelen et al., 2006]. This prototype focused on quantifying the emissions of one particular type of pollutants only, namely heavy metals, caused by one single sector only, namely the building stock. It calculated and located the emissions and their transport at a 60 metre resolution. With a view to test the applicability, robustness and sensitivity of the tool and the models incorporated, a follow-up study [Engelen and Van Esch, 2007] focused on a historical analysis for the years 1998, 2002 and 2005 and carried out a number of policy relevant scenario-analysis. Based on the experiences gained in the context of both projects, a decision was made to extend the methodology to all pollutants and all emission generating sectors and stakeholders. Thus, the concept of a generic EIW-PSS was born.

2. EIW – PSS: CONCEPTUAL FRAMEWORK

In Figure 1 the elements of a complete Emission Inventory Water (EIW) are depicted schematically. The components of the system are the inventory of sources, the calculation of the gross emission, the pathway, the material flow analyses and the geographical allocation of the discharge into the surface water (pressure). The sources are classified in categories, with distinction between point (in italic) and diffuse sources (See also Figure 2). They are represented in the various boxes on the left side of the diagram. They can be opened to retrieve information at a more detailed level and new sources of pollution can be added in case relevant new information and knowledge becomes available. Emissions are calculated on a sector-by-sector basis including: agriculture, industry, households and traffic. Emissions caused by materials used in buildings and road infrastructure, as well as natural processes such as atmospheric deposition are also considered. Pollutants follow different pathways from the source to the receiving aquatic environment. These pathways may involve: direct losses to the surface water, discharge to the sewer system or running of at the surface. For wastewaters delivered to the urban wastewater treatment plants, technical coefficients are applied to calculate the amounts of various compounds removed as part of the treatment. The pollutants that finally arrive in the surface water generate pressure. Transport of emissions in water courses and loads thus entering the study region are currently beyond the scope of the EIW-PSS. Rather, they are dealt with by hydrological modelling tools running in parallel. Thus, the EIW-PSS focuses strongly on emission sources, material flows and the pathways to the surface waters.

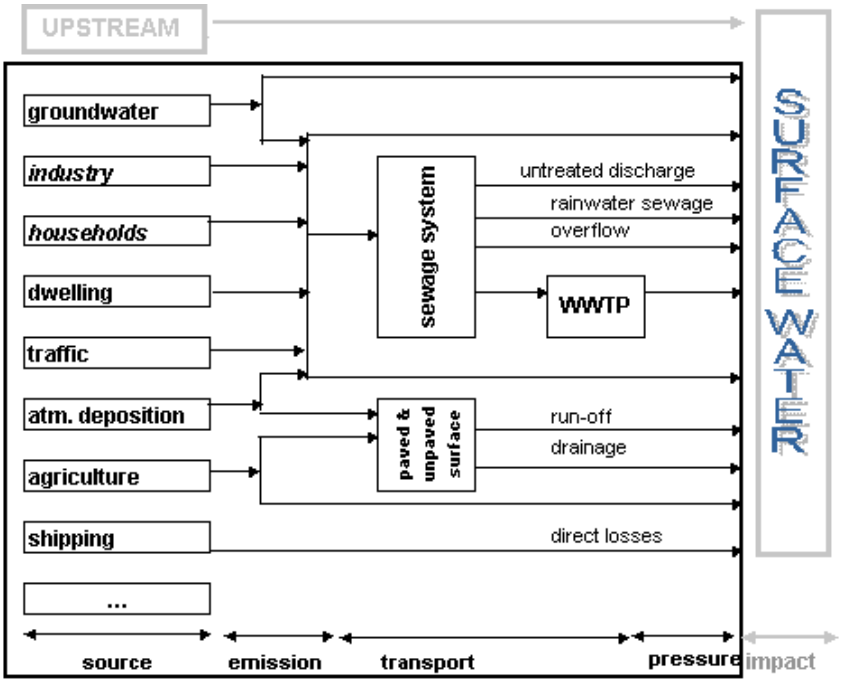


Figure 1. Functional scheme of the EIW-PSS: emission sources and pathways.

In order to implement the depicted EIW, various modules are invoked in function of the precise characteristics of the pollutant considered and its transport. Three sets of modules are available in the EIW-PSS to process the pollutant from its source to the surface waters (see Figure 2).

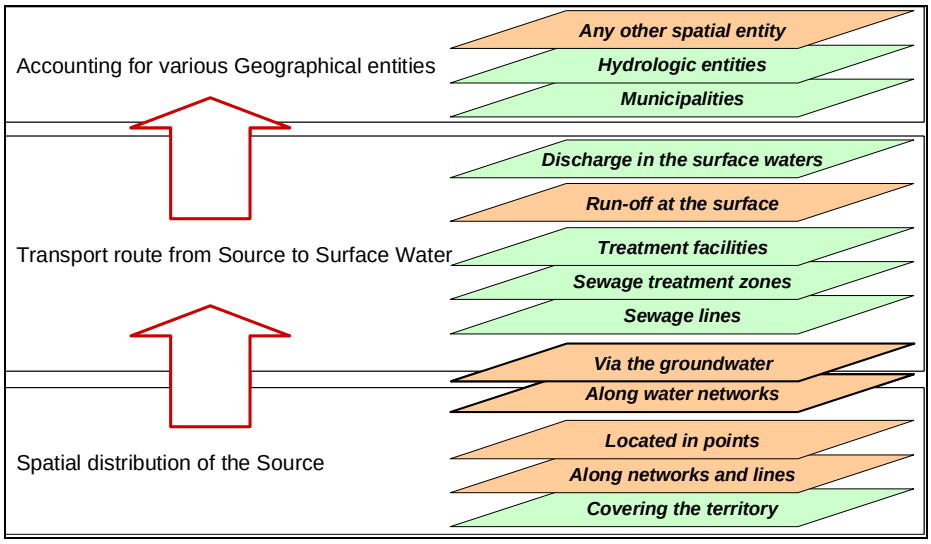


Figure 2. Representation of various modules integrated in the EIW-PSS.

The first set comprises modules enabling an adequate *geographical distribution* of the source itself and the amount of pollutant that it generates. Sources can be point sources, meaning that they are precisely located by means of their X-Y coordinates. Industrial facilities are a good example of this. But sources can just as well be diffuse and cover the entire territory in pockets of higher or lower density. This is the case among others for atmospheric deposition. Finally sources can be distributed along networks or line objects. Various forms of traffic and road infrastructure are examples of the latter. A special case are rivers and canals. Sources associated with the latter, such as PAHs emitted by the skin

of boats, end up directly in the surface waters and need no further processing relative to their transportation.

The second set of modules enable the *transport* of the pollutants from the source to the surface waters. Overland transport of diffuse sources will be handled by a run-off algorithm. Pollutants thus transported will end-up directly in the surface waters, in ditches and other drainage networks and/or in the sewage network. Point sources will often deliver their waste waters directly to the sewage system. Parts of the sewage networks will transport the waste waters and dissolved pollutants to the sewage treatment plants where the pollutants are partly removed. The remainder will be discharged in the surface waters as part of the effluent. Part of the pollutants will infiltrate into the soil and further into the groundwater. They are transported dissolved in the groundwater to the surface waters.

The characteristics of the treatment infrastructure and the territorial coverage subject to waste water treatment is know to the *accounting* modules of the third set. Accounts are produced at various nodes in the Material Flow Scheme (See Section 3.3) for pollutants removed and those reaching the surface waters and for relevant administrative or hydrological entities by aggregating the information available at the cellular level, polygons and/or line segments.

Fundamental principles underlying the accounting and the calculation of the transport of pollutants in the EIW-PSS are Conservation of Mass and Material Flow Analyses. They are not discussed here, rather will be dealt with in the example of Section 3.

For every source and every pollutant the relevant modules are selected from the three sets discussed. They are processed in a linked and sequential manner. In Figure 2 the modules are represented by means of map-layer symbols as they are predominantly geographically explicit sub-models implementing various spatial allocation, networking, dis-aggregation and aggregation algorithms. The modules in green operate currently as part of the EIW-PSS in a more or less final form, while layers in orange still need to be developed. Section 3 of this paper will discuss in greater detail the application of the above-mentioned as it has been applied to heavy metals caused by the building stock sector. This is a typical example of a diffuse source with transportation mainly through the sewage system. Models to calculate run-off at the surface and infiltration are not (yet) incorporated in this first prototype (orange in Figure 2).

In order to calculate the gross emission of a particular chemical compound generated by a sector, the *Emission Explanatory Variables (EEV)* and associated *Emission Factors (EF)* are estimated. The beauty of both concepts resides in their generic and the recursive applicability. Simply stated, an Emission Explanatory Variable is a further specification of the size of the source, a statistical number which can be updated on a for instance yearly basis. When applied to PAHs generated by transportation (exhaust of vehicles, wear of tyres, wear of the asphalt, oil spilled on parkings, etc.), EEVs include the number of driven kilometres, the surface of the parkings, etc.. EEVs have an explicit geographical distribution, meaning that they can be located on a map to result in areas of higher and lower densities. Emissions Factors are estimates of the amount of pollution generated per unit length, surface, weight or any relevant unit used to represent the EEVs. The more information is available about the pollution in the sector the more EEVs can be subdivided in smaller components and sub-components. For example for the sector building stock EEVs were several hundreds of hierarchically nested components exposed to atmospheric influences and water in the plumbing systems of various types of buildings. Multiplication of the EEV, its size (expressed as a length, a surface, a volume, …) and its associated EF results in the Gross Emission Value (GEV). In the EIW-PSS the GEV are calculated per unit length or area. Hence, the aforementioned multiplication requires a further summation, if not a multiplication with the units of EVV per unit length or area.

3. SPATIAL ANALYSIS OF THE BUILDING STOCK

The spatial component of the EIW-PSS allocates the main sources of pollution in a raster representation at a 60 metre resolution, incorporates the detailed representation of the sewage collection systems canalizing the pollutants to Urban WasteWater Treatment

Plants (UWWTPs), generates gross (prior to treatment) and net (after treatment) emission maps and produces accounts for catchment and administrative entities at different hierarchical levels.

In order to clarify the concepts introduced in the previous section as well as their practical implementation in the EIW-PSS, a closer look is taken at the example of the building stock currently implemented. Building stock refers here to all buildings for residential, commercial and social activities. It also includes the housing and office buildings of small industrial production units and farms, but excludes the industrial installations. The latter are an integral part of the analysis at the level of the industrial and/or agricultural sector itself. Building stock also includes materials used for fencing off and for equipping the terrains surrounding the buildings. The discussion is limited to the spatial component of the EIW-PSS. Readers interested in the estimation of the EEVs and EFs and in technical details about the mathematical models are referred to the original study [Engelen et al., 2006].

The spatial component of the EIW-PSS can be described in three successive steps (3.1, 3.2 and 3.3, Figure 3) described below. They are related to the three modules of Figure 2: 'spatial distribution of the source' in 3.1 and 3.2 and the modules 'transport route from source to surface water' and 'accounting' in 3.3. Chapter 3.4 shows an example of validation, based on the measured influent at UWWTPs.

3.1 Step 1: Spatial distribution of the building stock

Based on the data available, and taking into consideration typical function, morphology, location in space and application of building components, the following eight building types were withheld:

- row house;
- semi-detached house;
- detached house, farm or castle;
- apartment building;
- commercial building;
- industrial building;
- shack, garage and other small buildings including greenhouses;
- other building.

From the Belgian Federal Statistics data are available at the municipality level (308 in Flanders) for all building types. For the first four types, data are also available at the level of the so-called statistical sector which is a sub-division of the municipality (some 9500 in Flanders). The most detailed information (municipality or statistical sector) was used to spatially allocate the eight building types at a higher spatial resolution. A dasymetric mapping algorithm nearly identical to that of Mennis [2003] was chosen to that effect. It is generic, hence it also applies to other sectors covering the full territory.

Dasymetric mapping [Eicher and Brewer, 2001], [Mennis, 2003] is a technique enabling a cartographic representation of a spatial feature, in the given case buildings of a particular type, for which data are available at the level of relatively arbitrary spatial units, in the given case municipalities and statistical sectors, reflecting more realistically the precise spatial location of the feature within those spatial units. To that effect, use is made of ancillary (cartographic) information representative of the location of the feature. In the given case, the ancillary data is derived from the land-use map of Flanders at the 15 metre resolution available from the AGIV, the Agency for Geographical Information Flanders. The original map is resampled at the 60 metre resolution. The latter is in part a pragmatic choice enabling a straightforward aggregation of the original 15 metre cells, but it also recognizes the Flemish environmental legislation stating that buildings within a 50 metre buffer from a sewage collection pipe are to connect to that pipe. When overlaying the detailed location of the sewage system (in Step 3 of the procedure), it thus is possible to distinguish between grid cells that should and should not connect. Prior to the resampling, the original 20 land uses of the map were grouped into seven classes most relevant for the location of the building types.

The dasymetric mapping algorithm is performed in MS Excel. A distribution code has been established to link the eight building types with the seven land use types. Row houses are for instance for almost 100% allocated in the urban fabric land use category. The distribution equation contains also a component to correct for the fact that every land use type is not equally present in every statistical sector.

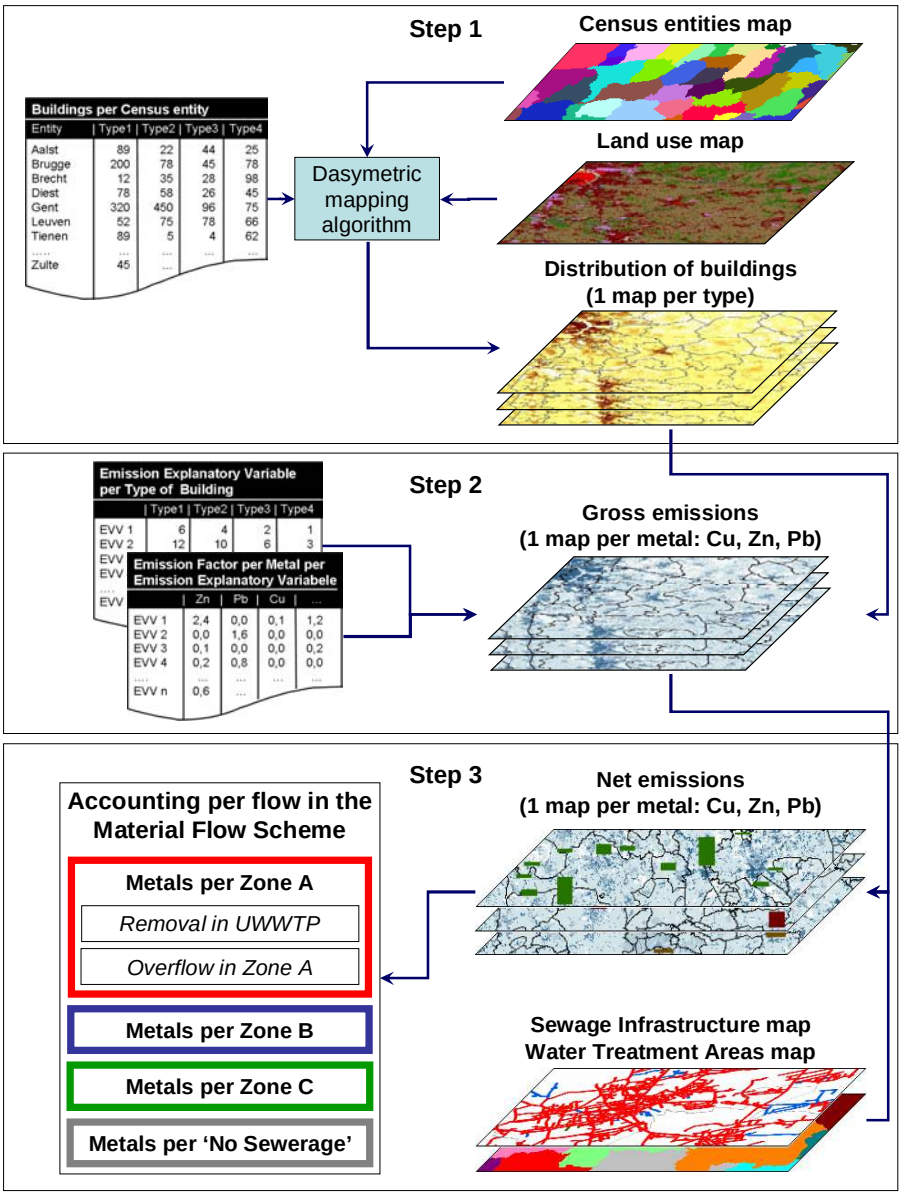


Figure 3. The spatial component of the EIW-PSS applied to the building stock.

In the calibration and validation of the procedure, the amount of buildings allocated per cell was compared with reality and the distribution code together with other technical coefficients were adjusted through iteration until a best fit with reality was obtained. The typical density of buildings allocated per land use cell could be estimated based on statistical sectors where only one land use type was present. It was further adjusted based on expert knowledge. Around 2,5 million buildings (situation 2005) were thus allocated. The result of Step 1 is a set of eight maps representing the distribution of each building type.

3.2 Step 2: Generating gross emission maps

Once the detailed distribution of each building type is known, the next step consists in calculating the gross emissions generated in situ. To that effect, the Belgian Building Research Institute developed an elaborate database of building components used now and in the recent past (+/- 30 years ago). Regional differences in the building practice and tradition were taken into consideration. Only components in the outer shell exposed to the atmosphere and those in the plumbing circuits were withheld. Data was collected relative to quantities used in each of the eight types and the amount of metals emitted on a yearly basis. Metals withheld were copper (Cu), zinc (Zn), lead (Pb), aluminum (Al), chromium (Cr) and nickel (Ni) but sufficient information to perform all steps of the analysis was only available for the first three. The information is obtained from literature research, a survey among building practitioners and major producers and distributors of zinc and copper. In addition, high resolution remote sensing images from Google Earth combined with field visits were used to define prototypical buildings with their associated building components.

Gross Emission Values (GEV) were established for the eight building types and for the three metals. A distinction was made between metals corroded in the outer shell of the buildings, hence feeding the surface run-off, and the sanitary waters originating in the plumbing circuits. As a result, the EIW-PSS provides maps for both the surface and sanitary waters for both gross and net emissions (see Step 3). The reason for this distinction is based on the growing importance given to decoupling in Flanders involving separate systems for the transport of surface run-off and sanitary waste waters.

The GEV are multiplied with the number of buildings (obtained from Step 1) on a cell-by-cell basis to result in gross emission maps per building type, per metal and per type of transport. The sum over all building types results in the gross emission maps per metal and per type of transport. The latter are input in Step 3.

3.3 Step 3: Transporting the emissions from source to sink and accounting

The Material Flow Scheme (Figure 4) is applied to the gross emission maps. This is another generic cornerstone of the methodology that can be applied for each sector individually. It shows how pollutants flow from the source to a sink via a number of nodes, mostly intermediate stations relevant for monitoring and reporting purposes. Conservation of mass applies, hence, the amount of pollutant entering a node is equal to that leaving the node. In this first prototype, no chemical reactions or leakages in the sewer system during transportation are taken into account.

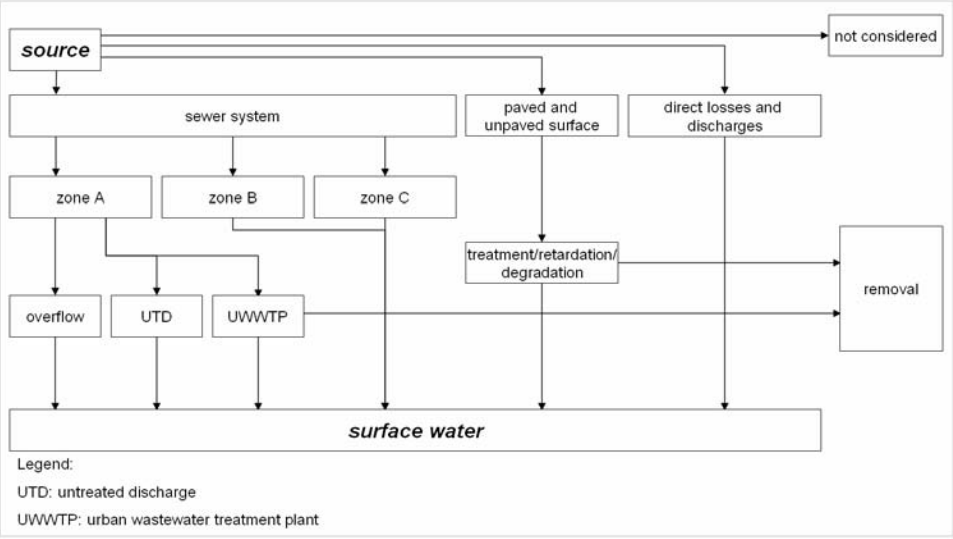


Figure 4. Material flow scheme of the EIW-PSS.

When applied to the building stock, pollutants running off at the surface to end up directly in the surface waters or infiltrating in the groundwater (node 'Direct losses and discharges') are estimated on the basis of coefficients obtained from hydrological models external to the system. This will change in the future when run-off and infiltration algorithms will be incorporated in the EIW-PSS (See Section 2). Most importantly for VMM are pollutants physically transported and treated by means of the sewage infrastructure (Figure 4: node 'Sewer system'). In the spatial analysis, this is implemented by overlaying the gross emissions with the vector layer representing the highly detailed sewage collection system of Flanders. A reasonably limited number of cells deliver their waste waters and emissions to a sewage collection system to be transported to an UWWTP (Figure 4: zone A). Zone B has a sewer system that is not yet connected to an UWWTP, but will be in the near future while for zone C there is no intention at all to connect the sewer system to an UWWTP. The application of the EIW-PSS to the building stock revealed that 57 per cent of the waste waters generated by the population of Flanders is treated in an UWWTP ('Zone A') and that 78 per cent of the population is connected to a sewage system ('Zone A' + 'Zone B' + 'Zone C)'. These numbers match very closely the ones used by the Flemish Environmental Agency based on alternative sources (respectively 60 and 80 per cent). Technical coefficients concerning treatment efficiency of the UWWTP (purification rates for each UWWTP and each metal, measured by the VMM), physical pre-treatment (Figure 4: node 'UTD', also available for each UWWTP and each metal) and storm weather overflow (Figure 4: node 'Overflow') are applied to calculate net emissions. This results in a net emission map per metal and per type of transport. By way of example, Figure 5 shows the net emission map for zinc summed for the two types of transport (surface and sanitary waters). The discharges of pollutants at the outlet of the UWWTPs are represented by bars.

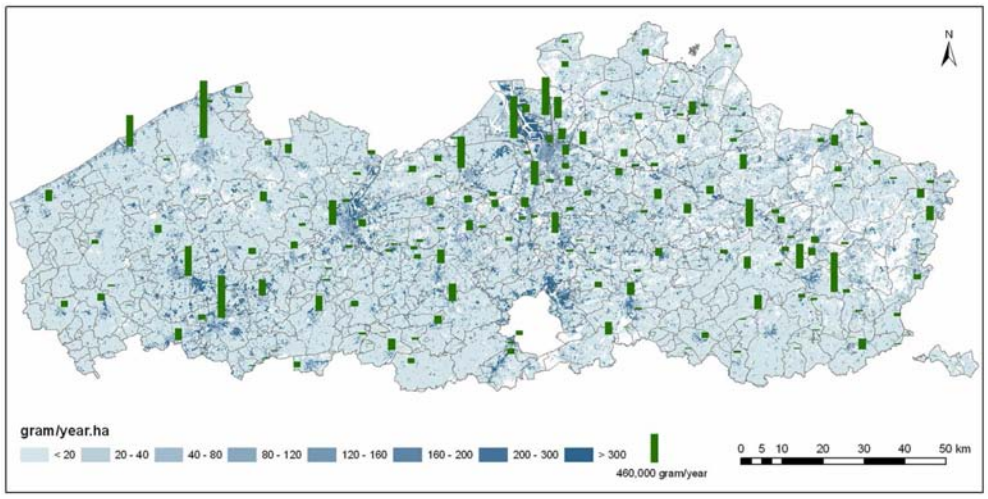


Figure 5. Net emissions of zinc due to the corrosion of the building stock.

3.4 Example of calibration and validation

It is possible to examine the quality of the model by comparing the calculated amounts of a heavy metal with the measured amounts at the inlet of the UWWTPs. The measured influent involves all sources while the prototype for the moment only calculates for the building stock. However, the first version of the EIW (Syncera water B.V., 2005) calculated emissions for all sectors and these estimates can be used to complete the prototype temporarily for the missing sectors. Figure 6 shows this validation applied onto zinc. In the first version of the EIW, the building stock represented 56% of the total calculated load. In the above-described prototype, completed with data of the first EIW, 71% of the calculated influent originates from the building stock. When confronting the total calculated influent with the total measured influent of zinc (181 ton/year), 64% of the

measured load could not be linked to a source in the first version of the EIW. In the above-described model, this proportion already decreased to 45%.

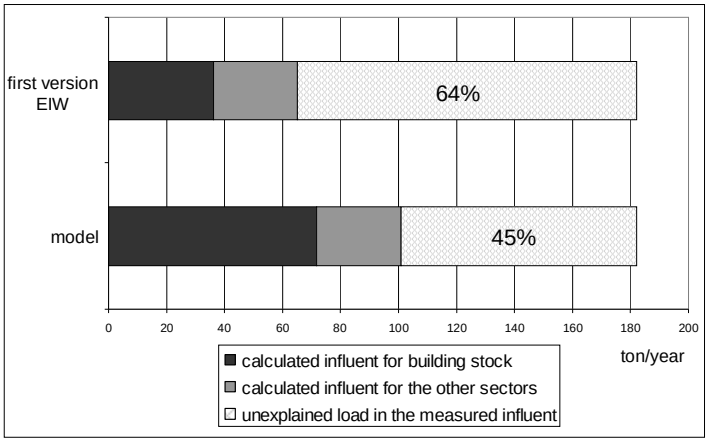


Figure 6. Validation at the inlet of the UWWTPs, confronting calculations with measurements for zinc.

4. THE EIW-PSS: PLANNING AND POLICY APPLICATIONS

The EIW-PSS provides the VMM with an important planning support system enabling a quantification of the pressures and impacts of point and diffuse emissions caused by different sectors and stakeholders. The algorithms implemented conserve the true geographical nature of the processes at high spatial resolutions, thus enabling an analysis and establishment of spatially explicit mitigation measures. Such spatial approach is most relevant for regions like Flanders, typified by high population densities and associated environmental pressures.

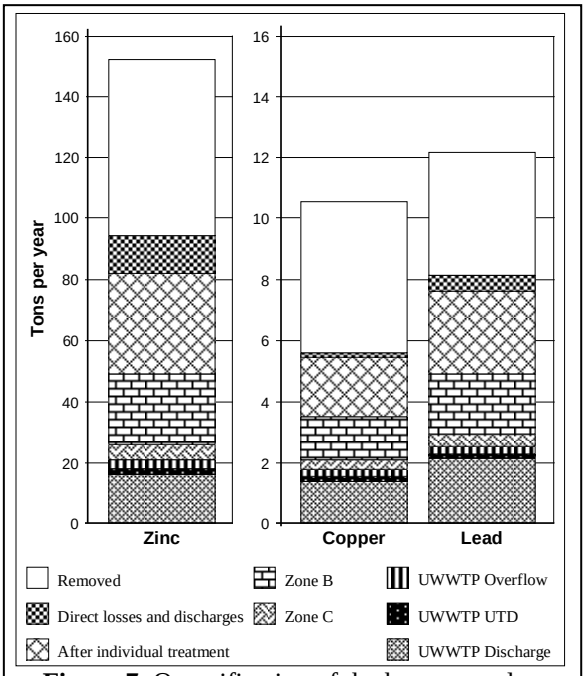


Figure 7. Quantification of the heavy metals in the nodes of the material flow scheme.

The study of the building stock revealed the benefit of a generic system enabling the straightforward assessment of scenarios. One such exercise analysed the evolution relative to heavy metal emissions in the period 1998-2005. It analysed improvements in the sewage system and treatment facilities (more area of zone A, more UWWTPs,, higher purification rates, less UTD …) as well as changes in the mix of buildings, their location in space and changes in the building materials used. The latter is especially interesting for historical analyses but also to explore the influence of land use change in the future (indicated by external land use change models, calculating 30 years ahead for instance). The sensitivity of several policy relevant parameters was also examined, among which the efficiency of septic tanks and the obligatory installation of individual treatment plants servicing remotely located buildings. Figure 7 demonstrates among other things the reduction in total emissions that could be realized if zone B would be connected to an UWWTP. Alternatively, similar reductions of the emissions could be attained in the 'individual treatment' zone with individual WWTPs featuring higher purification rates.

5. CONCLUSIONS

The EIW-PSS is an innovative tool supporting the Flemish Environmental Agency in their monitoring and reporting obligations as well as their scenario, planning and policy assessments. The conceptual framework and methodology have been elaborated with a generic usage in mind, and, their application to a first sector, namely the emission of heavy metals caused by corrosion in the building stock, showed satisfactory results. Among others, the study reduced the unexplained emission loads of Cu, Pb and Zn at the inlet of UWWTPs substantially. The results of the analysis carried out as well as the scenario capabilities of the first prototype were an impetus to continue the development.

Clearly, the deployment of the full framework in all its details will take time and effort. Effort will go into the development, integration and calibration of the various algorithms as part of the modules required to complete the scheme of Figure 2. Knowledge and modelling components to that effect will be derived from the literature, but also from state of the art applications. This should result in a system as generic as possible, yet, enabling the implementation of all the sectors and pollutants of interest. A system also which is sufficiently straightforward in its application so that it can be used for a wide range of policy relevant exercises by its envisaged end users. Specific studies will need to be carried out sector-by-sector and pollutant-by-pollutant. This will typically involve exhaustive literature research supplemented with measuring campaigns. Hence, at any particular point in time, the availability of information and resources will determine largely the level of depth that can be attained in representing sectors and their emissions. However, such type of development will not withstand the usage of intermediate versions of the system for practical purposes as the methodology allows for the representation of different sectors at different levels of detail. In the course of time the EIW-PSS can be upgraded with more and better information to evolve into an increasingly powerful planning support system.

ACKNOWLEDGEMENTS

The authors wish to thank Ir. K. De Cuyper and Ir. K. Dinne from the Belgian Building Research Institute for their work resulting in the estimation of the EF and the EEV.

REFERENCES

Crouzet, Ph., and J. Bogestrand, A European Inventory of Emissions to Inland Waters – A First Proposal, *Technical Report nr 8*, EEA, 1999.

Eicher C.L., and C.A. Brewer, Dasymetric mapping and areal interpolation: implementation and evaluation, *Cartography and Geographic Information Science, 28*(2), 125-138, 2001.

EmissieRegistratie Nederland, Spatial allocation documents, Retrieved November 5th, 2007, http://www.emissieregistratie.nl/erpubliek/misc/Documenten.aspx.

Engelen, G., L. Van Esch, L. Janssen, M. Van Holderbeke, J. Provoost, K. De Cuyper, and K. Dinne, Kwantificering en verfijning van het aandeel van de diffuse bronnen op de emissies naar oppervlaktewater in het stedelijk (gerioleerd) gebied. Deel: emissies van metalen te wijten aan corrosie van bouwmaterialen, Report 2006/IMS/R397, VITO, Mol, 2006.

Engelen, G., and L. Van Esch, Evolutie van de emissies in water uit corrosie van bouwmaterialen aan de hand van de referentiejaren 1998, 2002 en 2005, Report 2007/IMS/R/0429, VITO, Mol, 2007.

Mennis, J., Generating surface model of population using dasymetric mapping, *The Professional Geographer*, *55*(1), 31-42, 2003.

Syncera Water B.V., Emissie-Inventaris Water, Syncera Water B.V., Delft, 2005.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Predicting Winners and Losers: Translating Air Pollution Policy Initiatives into Indicators of Habitat Quality Change

E.C. Rowe[a], C.D. Evans[a], S. Smart[b] and B.A. Emmett[a]

[a] *Centre for Ecology and Hydrology, Environment Centre Wales, Deiniol Road, Bangor, Gwynedd, LL57 2UW, UK. (ecro@ceh.ac.uk)*
[b] *Centre for Ecology and Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, LA1 4AP, UK.*

Abstract: Damage to semi-natural habitats due to air pollution has hitherto been quantified using biogeochemical indicators, but predictions of species change are necessary to relate impacts on ecosystems to the habitat quality metrics used in conservation and policy communities. Ongoing pollution by reactive nitrogen (N) is delaying the recovery of acid-sensitive ecosystems from sulphur (S) pollution, and causing additional damage through eutrophication. Dynamic soil chemistry models are currently used to set national emissions thresholds for acid pollutants in a UNECE process, using simple targets based on soil pH or acid-base balance. These indicators are however poorly related to conservation targets based on habitat quality, which are defined in terms of the presence and abundance of indicator species. We demonstrate an approach to predicting habitat quality change under different N and S pollution scenarios, which brings together biogeochemists, biodiversity specialists and policymakers in an open modelling framework.

***Keywords*:** Acidification; Eutrophication; Indicator; Open modelling; Niche.

1. INTRODUCTION

The Convention on Long-Range Transboundary Air Pollution (CLRTAP) of the UNECE is an effective international policy instrument which uses evidence of actual or potential damage from pollutants to set emissions limits. The CLRTAP has been particularly successful in regulating emissions of the acidifying pollutants nitrogen (N) and sulphur (S). Damage indicators and thresholds are currently defined in terms of soil chemical criteria such as pH or Ca/Al ratio. This approach is being reconsidered due to two pressures. Firstly, a large reduction in S emissions has shifted the emphasis from acidification damage towards eutrophication, i.e. a damaging increase in ecosystem productivity driven by N pollution. Secondly, EU conservation legislation uses habitat quality indicators that are largely based on species occurrence, not soil properties (Figure 1). Hence, we have developed a model chain which predicts the effects of pollution emissions scenarios on species-based habitat quality indicators, by linking models of pollutant deposition, soil chemistry, species occurrence and habitat quality. This requires much transdisciplinary communication, so there is a clear requirement for transparency in model description, in particular the definition of model inputs and outputs.

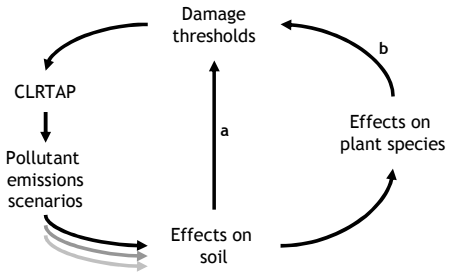


Figure 1. Schema for policy development within the Convention on Long-Range Transboundary Air Pollution (CLRTAP), illustrating the change from (a) soil-based damage indicators to (b) plant-species-based damage indicators.

2. SIMULATING SPECIES OCCURRENCE

2. 1 Introduction

A two-stage process was used to simulate the effects of different N and S deposition scenarios on plant species. Firstly, a dynamic model which is currently used by the UNECE Coordination Centre for Effects to simulate effects of pollution on soil [de Vries et al., 2007] was used to predict changes in soil pH and carbon (C) to N ratio. These variables were used as inputs into a set of static regression models [Smart et al., 2005] which predict species' probabilities of occurrence under a given set of environmental conditions.

2.2 Soil chemistry

The MAGIC model focuses on acid-base dynamics, solving at each timestep a system of linear equations which describe ionic competition for exchange sites. This model also simulates N dynamics using a simple saturation function. Nitrogen pollution affects habitats mainly through effects on competition. With a decrease in N limitation, low-growing, light-demanding species are shaded out by increased growth of faster-growing and taller species. However, the soil microflora competes strongly with plants for N additions, and initial inputs of N into a pristine ecosystem are mainly immobilised by interaction with the soil carbon (C) pool, resulting in a decrease in soil C/N ratio [Gundersen et al., 1998]. Because of the large organic matter pool in most soils, C/N ratio changes only slowly with typical levels of atmospheric inputs. In the absence of historic soil measurements, N addition experiments allow predictions of soil C/N change to be assessed (Figure 2).

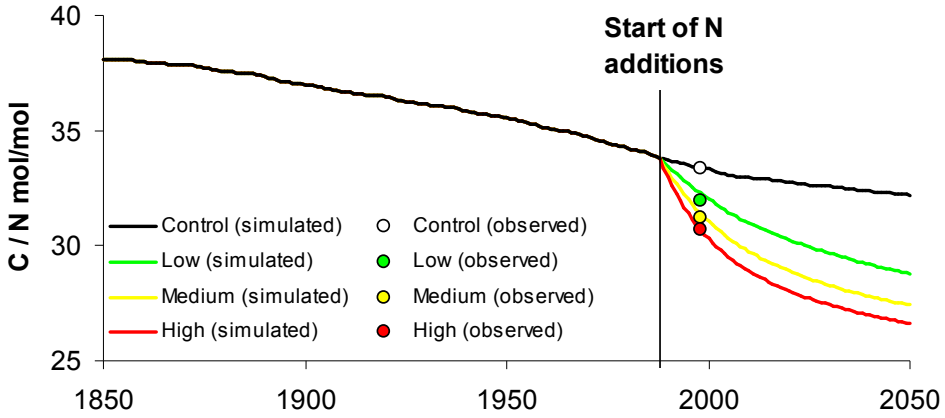


Figure 2. Measured soil C/N ratio in an upland dry heathland (Ruabon, N Wales, UK) receiving 0, 40, 80 and 120 kg N ha^{-1} y^{-1} since 1988, and MAGIC simulations of this site [Evans et al., 2006]. Initial C/N was obtained by calibrating the model to measured C/N in the zero additions treatment.

2.3 Plant niches

Plant species occur within specific ranges of environmental conditions, determined by their autecology, competition with other plants, and susceptibility to pests and diseases. This 'realised niche' is a hypervolume within the space defined by a set of environmental factors. Maximum probabilities of occurrence exist for many species in relation to these factors, although some species have broad tolerance or bimodal distributions when plotted against individual factors. Ellenberg [1992] and subsequent authors defined optima scores for European plant species on arbitrary scales related to the availability of nutrients, water and light, and acid reactivity. We modelled the occurrence of plant species in two stages, to make use of datasets where environmental factors were measured alongside occurrence as well as the far larger datasets where only species were recorded. For the former dataset, multiple regression was used to calibrate mean scores for present species against five soil variables: total C and N content, pH, % soil moisture, and bicarbonate-extractable phosphorus. Minimum adequate models were determined by first rejecting factors with no significant effect on mean Ellenberg score and then using deviance reduction tests to accept or reject interactions and quadratic terms for the remaining factors. Next, multiple logistic regression was used to relate probability of occurrence within the larger dataset to the mean Ellenberg score, by stepwise elimination from the set of significant effects, interactions and quadratic effects. This results in a set of per-species models (Figure 3), collectively known as GBMOVE, which can predict the probability of a species occurring under a given set of environmental conditions [Smart et al., 2005].

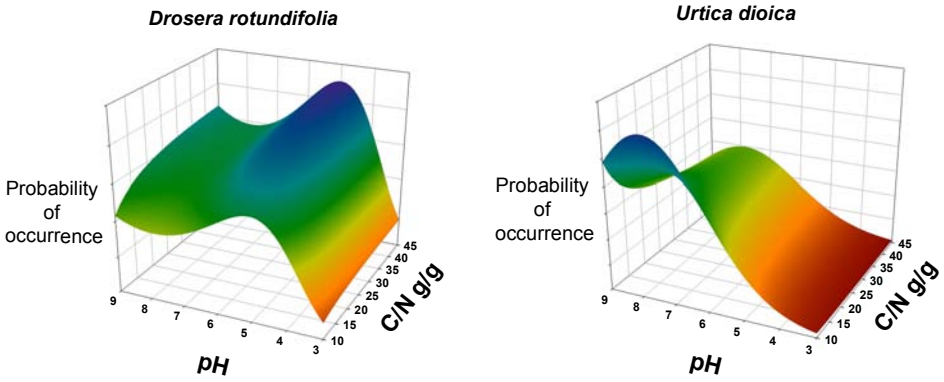


Figure 3 GBMOVE niche models for *Drosera rotundifolia* and *Urtica dioica* (plants typical of infertile and fertile habitats, respectively) in relation to soil pH and soil carbon/nitrogen ratio, at fixed levels of canopy height (0.8 m), soil moisture content (35%) and soil carbon content (5%).

2.4 Species dynamics

Coupling the MAGIC dynamic soil chemistry model with the GBMOVE static species niche model allows the dynamics of plant species occurrence to be simulated. Testing this model chain is difficult due to the scarcity of long-term datasets covering both soil chemistry and plant species occurrence. However, the model chain proved fairly successful in simulating the rates of change in probability of occurrence in a floristic dataset extending from 1973 to 2001 (Figure 4).

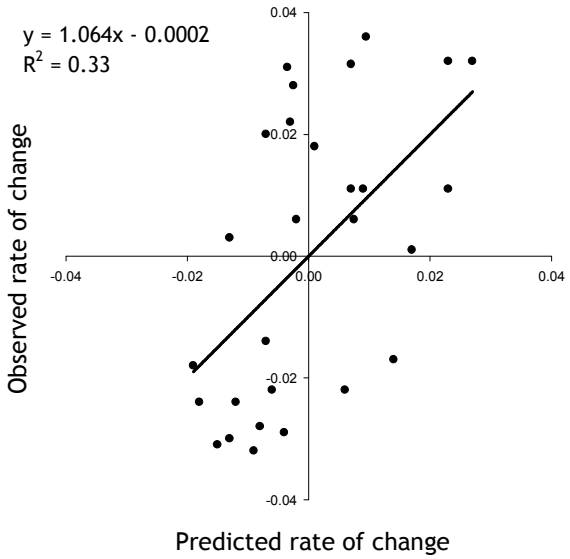


Figure 4. Predicted versus observed change for individual species in blanket bog (Moor House long-term monitoring site, Cumbria, UK). Predicted change is the slope coefficient of a linear regression on occurrence probabilities predicted by the MAGIC-GBMOVE model chain for each year between 1973 and 2001. Observed change is the slope coefficient of a linear regression on % frequency in sample plots in each survey year. Pearson correlation coefficient = 0.568, p=0.002 [Smart et al., 2005].

3. FROM SPECIES OCCURRENCE TO BIODIVERSITY INDICATORS

Predictions of changes in probability of occurrence for a large number of individual species are of limited value in themselves for assessing habitat quality, which is assessed according to criteria such as rarity, typicality and fragility [Ratcliffe, 1977]. The Habitats Directive of the European Union [EEC, 1992] has lead in the UK to the development of Common Standards Monitoring (CSM) protocols [JNCC, 2006] which list species for each habitat, grouped into positive and negative indicators. We have used this categorisation to interpret predictions of species change, by calculating an overall habitat-specific quality score Q as

$$Q = \sum_{i=1}^{p} \frac{\left(\frac{P_i}{P\max_i} \right)}{p} - \sum_{j=1}^{n} \frac{\left(\frac{P_j}{P\max_j} \right)}{n} \qquad (1)$$

where P_i and P_j are the probabilities of occurrence of the p positive indicators and n negative indicators for a habitat, respectively, and Pmax is the maximum probability of occurrence for the species within the parameter space. Re-scaling to Pmax is necessary because the large variation in abundance of individual species within large survey datasets and in maximum occupancy of suitable sites means that absolute probabilities of occurrence vary widely. Rescaling to p and n allows for the variable numbers of positive and negative indicator species listed for different habitats. Species not included in the CSM indicator lists are not used in this calculation.

The approach is illustrated using simulations of C/N and pH change due to atmospheric S and N deposition (Figure 5). Whilst with more N pollution the probability of occurrence increases for some positive indicators, and decreases for some negative indicators, the overall pattern of response is for N pollution to favour negative over positive indicators and so there is a clear decline in the overall quality score.

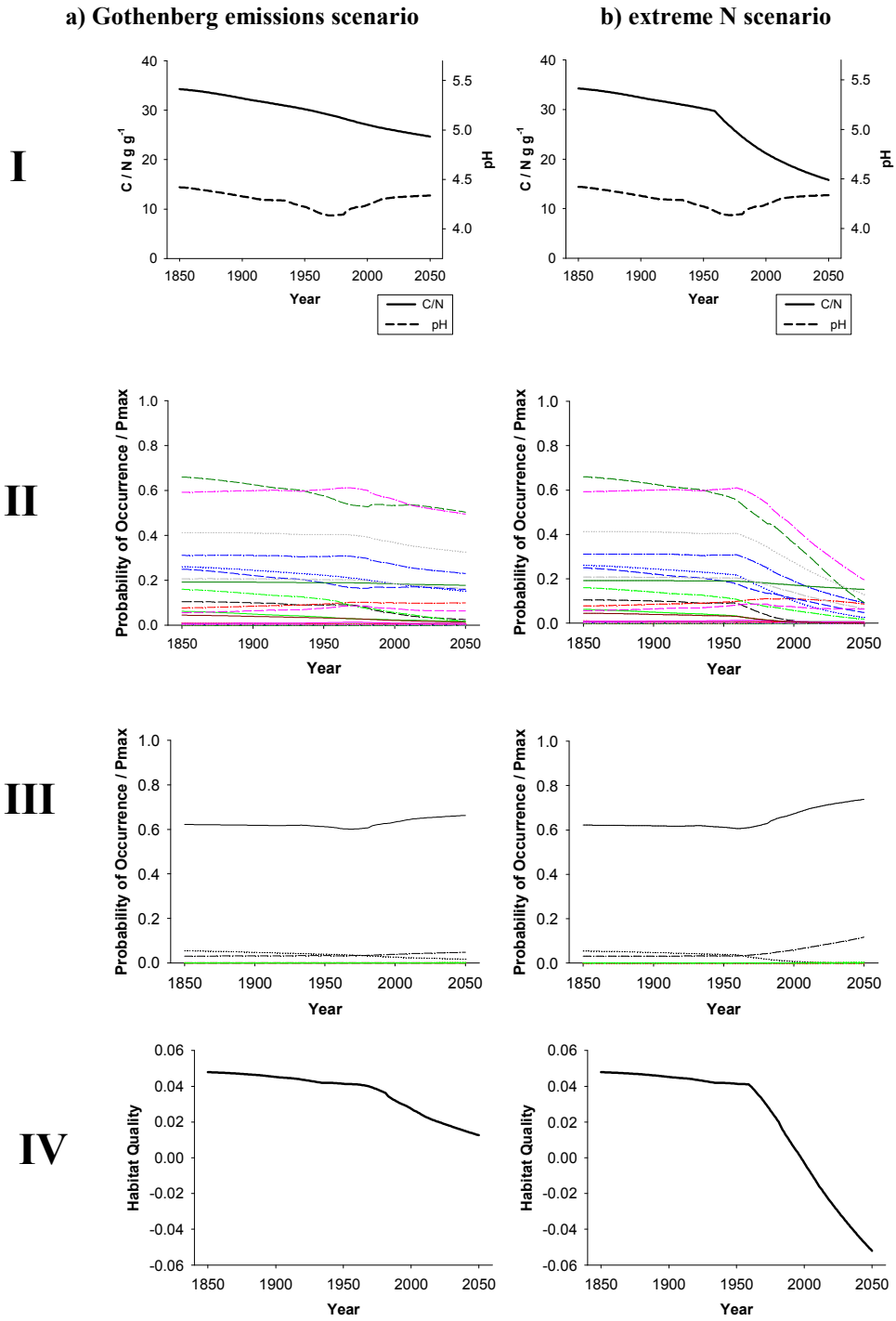


Figure 5. Simulated changes in blanket bog at Moor House long-term monitoring site, Cumbria, UK, under a) Gothenberg emission scenario, and b) an extreme N addition scenario with an additional 50 kg N ha^{-1} yr^{-1} from 1960. (I) Soil pH and C/N ratio simulated using the MAGIC soil chemistry model. Soil water content and canopy height were assumed to be constant. (II) Probabilities of occurrence of positive Common Standards Monitoring indicator species for blanket bog, rescaled to Pmax. (III) Probabilities of occurrence of negative CSM indicator species for blanket bog, rescaled to Pmax. (IV) Overall habitat quality Q (see equation 1).

4. MODEL IMPROVEMENT

4. 1 Uncertainty

Linking models increases uncertainty related to both parameterisation and model structure. A major source of uncertainty in the MAGIC-GBMOVE model chain relates to the prediction of mean Ellenberg fertility score from soil biophysical measurements. Between-habitat variation in the inert proportion of soil organic matter [Rowe et al., 2006] results in a poor correlation between total soil C and N pools and mean Ellenberg fertility score. We are currently measuring a set of soil properties including mineralisable N alongside floristic data in a large national survey, to identify more accurate predictors. This reduction in parameter-uncertainty will however lead to a requirement for more structural complexity in the soil organic matter model.

4. 2 Additional drivers

The inclusion of multiple environmental gradients in the GBMOVE niche models offers the potential to link to other dynamic biophysical models. Canopy height is used in the niche models as a surrogate for ground-level light availability, which has a profound effect on environmental suitability for plant species. We are currently using this mechanism to link GBMOVE to the SUMO vegetation model [Wamelink, 2007] to simulate the effects of vegetation succession on plant species occurrence. The effects of changes to rainfall pattern could be represented via effects on mean soil moisture content. We have also developed an extended version of GBMOVE that includes other climatic variables, allowing simulations of climate change. Since the niche models are based on empirical data that includes the effects of multiple drivers, we have some confidence that this approach can be applied to simulations of several interacting biophysical drivers.

4. 3 Open model development

The model development process has historically been monopolised by the subset of scientists who are also programmers. We are using several methods to open this process to wider scrutiny and discussion. Static niche models related to measurable factors are easily understood and checked by plant ecologists, compared with dynamic models of species populations or of cover proportion. Dynamic soil chemistry models written in FORTRAN are being re-implemented in SIMILE graphical modelling software [Muetzelfeldt and Massheder, 2003], which can be understood and assessed by non-programming modellers. This also makes it easy to see submodel boundaries, eliminating the danger of double representation of processes when combining models. Lastly, the use of CSM indicators to assess habitat quality allows discussion with habitat specialists in the UK nature conservation agencies using familiar terms.

5. CONCLUSIONS

By linking a dynamic soil chemistry model to static plant niche models, we have obtained positive correlations between predicted and measured change in the abundance of individual plant species. This approach could be applied to a variety of other problems in nature conservation, which urgently requires methods for predicting the effects of a changing environment on species abundance. The proposed one-dimensional habitat quality summary statistic is based on indicators used by conservationists and policymakers, yet allows definition of a simple damage threshold that could be used to define acceptable pollutant loads.

REFERENCES

Cosby, B.J., Ferrier, R.C., Jenkins, A., Wright, R.F., Modelling the effects of acid deposition: refinements, adjustments and inclusion of nitrogen dynamics in the MAGIC model. *Hydrology and Earth System Sciences* 5 (3) 499–517, 2001.

Cosby, B.J., Wright, R.F., Hornberger, G.M., Galloway, J.N., Modelling the effects of acid deposition: assessment of a lumped parameter model of soil water and streamwater chemistry. *Water Resources Research* 21 51–63, 1985.

de Vries, W., Kros, H., Reinds, G.J., Wamelink, G.W.W., Mol, J., Van Dobben, H., Bobbink, R., Emmett, B.A., Smart, S.M., Evans, C.D., Schutlow, A., Kraft, P., Belyazid, S., Sverdrup, H., Van Hinsberg, A., Posch, M., Hettelingh, J.-P., Developments in deriving critical limits and modeling critical loads of nitrogen for terrestrial ecosystems in Europe. Alterra, Wageningen, The Netherlands, 2007.

EEC, Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. Official Journal of the EEC, L 206 , 22/07/1992.

Ellenberg, H., Weber, H.E., Dull, R., Wirth, V., Werner, W., Paulissen, D., Zeigerwerte von Pflanzen in Mitteleuropa. 2nd Ed. *Scripta Geobotanica* 18 1–258, 1992.

Evans, C.D., Caporn, S.J.M., Carroll, J.A., Pilkington, M.G., Wilson, D.B., Ray, N., Cresswell, N., Modelling nitrogen saturation and carbon accumulation in heathland soils under elevated nitrogen deposition. *Environmental Pollution* 143 (3) 468–478, 2006.

Gundersen, P., Callesen, I., de Vries, W., Nitrate leaching in forest ecosystems is related to forest floor C/N ratios. *Environmental Pollution* 102 403–407, 1998.

JNCC, Common Standards Monitoring Guidance for Upland habitats. Version October 2006. URL: http://www.jncc.gov.uk/page-2237, 2006.

Muetzelfeldt, R., Massheder, J., The Simile visual modelling environment. *European Journal of Agronomy* 18 (3-4) 345–358, 2003.

Ratcliffe, D. (Ed.), A nature conservation review: The selection of biological sites of national importance to nature conservation in Britain. Cambridge University Press, 1977.

Rowe, E.C., Evans, C.D., Emmett, B.A., Reynolds, B., Helliwell, R.C., Coull, M.C., Curtis, C.J., Vegetation type affects the relationship between soil carbon to nitrogen ratio and nitrogen leaching. *Water Air and Soil Pollution* 177 (1-4) 335–347, 2006.

Smart, S., Evans, C., Rowe, E., Wamelink, W., Wright, S., Scott, A., Roy, D., Preston, C., Hill, M., Rothery, P., Bullock, J., Moy, I., Emmett, B., Maskell, L., Atmospheric nitrogen pollution impacts on biodiversity: Phase 1 - Model development and testing. Centre for Ecology and Hydrology, 2005. URL: http://www.defra.gov.uk/science/project_data/DocumentLibrary/WC02008/WC02008_2866_FRP.pdf

Wamelink, G.W.W., Simulation of vegetation dynamics as affected by nitrogen deposition. Ph.D. thesis, Wageningen University, 2007.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Monitoring the effects of human activities and environmental change on a population system

M. Gámez[a], I. López[a], J. Garay[b] and Z. Varga[c]

[a] *Departamento de Estadística y Matemática Aplicada, Universidad de Almería, La Cañada de San Urbano, 04120 Almería, España (mgamez@ual.es)*
[b] *Ecological Modelling Research Group of the Hungarian Academy of Science and Department of Plant Taxonomy and Ecology, Eötvös Lorand University, Budapest, Hungary.(garayj@ludens.elte.hu)*
[c] *Institute of Mathematics and Informatics, Szent István University, Páter K. u. 1., H-2103 Godollo, Hungary. (Varga.Zoltan@gek.szie.hu)*

Abstract: The monitoring and analysis of the processes taking place in an ecosystem is a key issue for a sustainable human activity. A system of populations, as the biotic component of a complex ecosystem is usually affected by the variation of its abiotic environment. Even in nearly natural ecosystems an abiotic effect like climatic implications of global warming may cause important changes in the dynamics of the population system. In ecosystems involving field cultivation or any industrial activity, the abiotic parameter in question may be the concentration of a substance, changing e.g. as a result of pollution, application of a pesticide or a fertilizer, etc. In many cases the observation of the densities of each population may be technically complicated or expensive, therefore the question arises whether from the observation of the densities of certain (indicator) populations, the whole state process of the population system can be uniquely recovered. The paper is aimed at a methodological development of state monitoring, under the conditions of a changing environment. It is shown, how the technique of mathematical systems theory can be applied not only for the approximate calculation of the state process on the basis of the observed data, even under the effect of an exogene abiotic change with known dynamics; but in certain cases, also for the estimation of the unknown biological effect of the change of an abiotic parameter. The proposed methodology is applied to simple illustrative examples concerning a three-species predator-prey system.

Keywords: environmental monitoring, observer system, ecosystem

1. INTRODUCTION

Detection of the effect of human activities (e.g. pollution) and environmental change such as climatic variation, is an important part of the analysis of a *composite system, including ecological, human and environmental subsystems*. The effect of such factors on the parameters of population system models has been studied by several authors, see e.g. Gragnani [2002], Xia [2007]. The problem we consider is the following: Suppose we know that an abiotic change has an effect on the parameters of a population system, we want to estimate, at the same time both the whole state process and the change of the mentioned parameters, observing the densities of certain indicator populations.

For the solution of this problem a constructive method is proposed. Mathematical systems theory offers appropriate concepts and technique for the development such a methodology. The methodological foundations of the application of controllability and observability to frequency-dependent population models (described by systems with invariant manifold),

have been set in Varga [1989] and Varga [1992], see also Scarelli and Varga [2002]. The original problem of state monitoring of a population system as formulated in Varga et al. [2003] is that, from the observation of the time-dependent densities of certain species, the whole state process of the population system is to be recovered. An important concept for the solution of this problem is *observability*. The latter in this context means that *from the observation of one or several (but not all) state variables, it is possible to recover the whole state process of the populations system, in a unique way* (without determining, however, a constructive method to obtain this process). Observability has been analysed in different population system models in Varga et al. [2002, 2003] and Shamandy [2005], see also López [2003], López et al. [2004]. In order to prove observability, we use a general sufficient condition for local observability of nonlinear observation systems, published in Lee and Markus [1971]. For a general review on the application of mathematical systems theory in population biology, see Varga [2008].

Once local observability near an existing equilibrium is proved, we also need a constructive method for the estimation of the state process. To this end an auxiliary system, the so-called *observer system* can be applied. The latter is constructed from the observed data, and asymptotically estimates the complete state process with exponential speed of convergence, see Sundarapandian [2002] and López et al. [2007a, b].

Now, if by human activities (e.g. pollution) any environmental disturbance takes place, the parameters of the population system model may change, which can be expressed in terms of unknown (constant or time-dependent) parameters, added to the known system parameters. In the paper we deal with two cases. First we suppose there is an unknown additive constant in certain biological parameters and, from the observation, we estimate both these unknown constants and the unknown solution of the population system model. In the second case, certain biological parameters change according to a known dynamics described by an additional system called *exosystem*. In Section 2, from nonlinear systems theory we briefly recall a concept and a theorem concerning observability. Section 3 is devoted to the observer design, and contains illustrative numerical examples for the application of the proposed methodology to both cases mentioned above. If we apply the above state process estimation to the composite system including both the population system model and the exosystem, we obtain asymptotic estimation not only for the state process of the population system, but at the same time also for the unknown parameters indicating the abiotic disturbances. Using computer simulation, the above construction will be illustrated by a three-species Lotka-Volterra model, where the Malthus parameters are changed as result of a human (environmental) intervention.

2. CONCEPT OF OBSERVABILITY. APPLICATION TO A PREDATOR-PREY MODEL

Given positive integers *m*, *n*, let

$$f : \mathbf{R}^n \to \mathbf{R}^n, \quad h : \mathbf{R}^n \to \mathbf{R}^m$$

be continuously differentiable functions and suppose for some $x^* \in \mathbf{R}^n$ we have that $f(x^*) = 0$ and $h(x^*) = 0$.

We consider the following observation system

$$\dot{x} = f(x) \tag{1}$$

$$y = h(x) \ , \tag{2}$$

where y is called the *observed function.*

Definition 2.1 Observation system (1)-(2) is called *locally observable* near equilibrium x^*, over a given time interval $[0,T]$, if there exists $\varepsilon > 0$, such that for any two different solutions x and $\bar{x}$ of system (1) with $| x(t) - x^* | < \varepsilon$ and $| \bar{x}(t) - x^* | < \varepsilon$ $(t \in [0,T])$,

the observed functions $h \circ x$ and $h \circ \bar{x}$ are different. ($\circ$ denotes the composition of functions. For brevity, the reference to $[0,T]$ will be suppressed).

For the formulation of a sufficient condition for local observability, consider the linearization of the observation system (1)-(2), consisting in the calculation of the Jacobians $A := f'(x^*)$ and $C := h'(x^*)$.

Theorem 2.2 (Lee and Markus [1971]). Suppose that

$$rank[C \mid CA \mid CA^2 \mid ... \mid CA^{n-1}]^T = n. \tag{3}$$

Then the observation system (1)-(2) is locally observable near equilibrium x^*.

We consider now the biological model of 2 preys and 1 predator of the form $\dot{x} = f(x)$, determined by the following system of differential equations:

$$\begin{aligned} \dot{x}_1 &= x_1(a_1 - b_{11}x_1 - b_{12}x_2) \\ \dot{x}_2 &= x_2(-a_2 + b_{21}x_1 - b_{22}x_2 + b_{23}x_3) \\ \dot{x}_3 &= x_3(a_3 - b_{32}x_2 - b_{33}x_3) \end{aligned} \tag{4}$$

with $a_i, b_{ij} > 0$ for all $i, j = 1,2,3.$

It is easy to provide a simple algebraic condition for the existence of an equilibrium in mathematical sense, however its positivity depends on the model parameters. Throughout the paper we shall suppose that there exists an equilibrium $x^* > 0$ for the considered model.

We assume that we observe the densities of the prey populations, i.e., the observation equation is $y = h(x) := (x_1 - x_1^*, x_3 - x_3^*).$ Then

$$C = \frac{\partial h}{\partial x}(x^*) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Linearizing the Lotka-Volterra system (4) around x^* we obtain

$$A = \frac{\partial f}{\partial x}(x^*) = \begin{bmatrix} -b_{11}x_1^* & -b_{12}x_1^* & 0 \\ b_{21}x_1^* & -b_{22}x_2^* & -b_{23}x_2^* \\ 0 & -b_{32}x_3^* & -b_{33}x_3^* \end{bmatrix}.$$

It is easy to check that $rank[C \mid CA \mid CA^2]^T = 3.$ Thus, by Theorem 2.2 the system is locally observable near the equilibrium, and the whole system state can be monitored observing only the densities of the prey populations.

3. CONSTRUCTION OF AN OBSERVER SYSTEM FOR STATIC AND DYNAMIC ENVIRONMENT

Now, the construction of an observer system will be based on Sundarapandian (2002). Let us consider observation system (1)-(2).

Definition 3.1. Given a continuously differentiable function $G : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n$, system

$$\dot{z} = G(z, y) \tag{5}$$

is called a *local asymptotic (respectively, exponential) observer for observation system* (1)-(2) if the composite system (1)-(2),(5) satisfies the following two requirements.

i) If $x(0) = z(0)$, then $x(t) = z(t)$, for all $t \geq 0$.

ii) There exists a neighbourhood V of the equilibrium x^* in $\mathbf{R}^n$ such that for all $x(0), z(0) \in V$, the estimation error $z(t) - x(t)$ decays asymptotically (respectively, exponentially) to zero.

Theorem 3.2 (Sundarapandian [2002]). Suppose that equilibrium x^* is Lyapunov stable for system (2), and that there exists a matrix K such that matrix $A - KC$ is Hurwitz (i.e. its eigenvalues have negative real parts), where $A = f'(x^*)$ and $C = h'(x^*)$. Then the dynamic system defined by

$$\dot{z} = f(z) + K[y - h(z)] \tag{6}$$

is a local exponential observer for observation system (1)-(2).

Now, for the estimation of a change in the dynamical parameters of an ecosystem, we recall that Sundarapandian [2002] also considered the possibility of an "input generator" determined by an external system called *exosystem*, $w' = s(w)$, in terms of which we can form a composite (nonlinear) system with inputs of the form

$$\begin{aligned} \dot{x} &= F(x, u(w)) \\ \dot{w} &= s(w) \\ y &= h(x), \end{aligned} \tag{7}$$

where we suppose that $F : \mathbf{R}^n \times \mathbf{R}^k \to \mathbf{R}^n$, $s : \mathbf{R}^k \to \mathbf{R}^k$ are continuously differentiable and $F(x^*, 0) = 0$, $u(w^*) = 0$, $s(w^*) = 0$. Here $u(w(t))$ is interpreted as a time-dependent vector of system parameters of the original system (1), corresponding to right-hand side f in the sense that $F(x, 0) = f(x)$. For the construction of an observer for the composite system we can apply the following

Theorem 3.3 (Sundarapandian, [2002]. Suppose that Suppose that equilibrium x^* is Lyapunov stable for system (7), the linearization of system (7) is observable, and that there exists a matrix K such that matrix $A - KC$ is Hurwitz, where $A = F'(x^*, w^*)$ and $C = h'(x^*)$. Then dynamic system defined by

$$\dot{z} = F(z, u(w)) + K[y - h(z)] \tag{8}$$

is a local exponential observer for observation system (7).

Applying this methodology, we can not only asymptotically recover the state process x from the observation, but in certain cases also estimate the unknown change of the model parameters of the population system.

To apply this construction to our population system (4), we remind that in López et al. [2007] equilibrium $x^* = (x_1^*, x_2^*, x_3^*)$ was proved to be asymptotically stable for system (4). In fact, function

$$V(x) := \alpha H x_1 / x_1^* + \beta H x_2 / x_2^* + \gamma H x_3 / x_3^* - (\alpha + \beta + \gamma) \quad (x \in \mathbf{R}_+^3),$$

with $\alpha = b_{21} / x_2^* x_3^* b_{12}$, $\beta = 1 / x_1^* x_3^*$, $\gamma = b_{23} / x_1^* x_3^* b_{23}$, and $H(v) := v - \ln v$ ($v \in \mathbf{R}_+$) was shown to be a Lyapunov function, implying asymptotic stability of equilibrium x^* for system (4), globally with respect to the positive octant of $\mathbf{R}^3$.

3.1. Observer for a system with unknown environmental change

Assume that the considered ecosystem consists, on the one hand, of a system of several interacting populations living in the given habitat, and the abiotic environment on the other.

The latter may also be exposed to climatic (e.g. seasonal) changes and/or human intervention, such as e.g. pollution, described by certain abiotic parameters (e.g. temperature or concentration). In this sections considering the predator-prey model (4), we suppose that the reference value of certain abiotic parameters change to an unknown constant value. The effect of this change will be described by a small additive term (disturbance) $w \in \mathbf{R}$ in certain model parameters. In our illustrative numerical examples we will show how we can recover the whole state process of the population system and estimate the unknown disturbance at the same time, by constructing and solving the corresponding observer system. In Example 1 we shall admit disturbance in all Malthus parameters.

Consider first our population system, completed with a trivial equation for the abiotic parameter w:

$$\begin{aligned} \dot{x}_1 &= x_1(a_1 + c_1 w - b_{11}x_1 - b_{12}x_2) \\ \dot{x}_2 &= x_2(a_2 + c_2 w + b_{21}x_1 - b_{22}x_2 + b_{23}x_3) \\ \dot{x}_3 &= x_3(a_3 + c_3 w - b_{32}x_2 - b_{33}x_3) \\ \dot{w} &= 0 \end{aligned} \tag{9}$$

with $a_i, b_{ij}, c_i > 0$ for all $i, j = 1, 2, 3.$

It is clear that with the equilibrium $x^* > 0$ of the previous section, $(x^*, 0)$ is an equilibrium for system (9).

Now, for the estimation of both components (x, w) of system (9), i.e. both the state process and the unknown parameter, we suppose that the densities of both preys are observed:

$$y = h(x) = (x_1 - x_1^*, x_3 - x_3^*).$$

Then

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \; A = \begin{pmatrix} -b_{11}x_1^* & -b_{12}x_1^* & 0 & c_1 x_1^* \\ b_{21}x_2^* & -b_{22}x_2^* & b_{23}x_2^* & c_2 x_2^* \\ 0 & -b_{32}x_3^* & -b_{33}x_3^* & c_3 x_3^* \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{10}$$

Now, we easily obtain that $rank[C \,|\, CA \,|\, CA^2 \,|\, CA^3]^T = 4$. Therefore, by Theorem 2.2, the system is locally observable near the equilibrium, and applying the technique of Sundarapandian [2002] we can construct a corresponding observer system, according to Theorem 3.2.

Example 1. As a numerical example, we consider the following two-prey one-predator model with the presence of an unknown environmental disturbance w:

$$\begin{aligned} \dot{x}_1 &= x_1(2 + 0.1w - 1.1x_1 - 0.1x_2) \\ \dot{x}_2 &= x_2(-1 - 0.3w + x_1 - 0.2x_2 + 0.5x_3) \\ \dot{x}_3 &= x_3(3.6 + w - 0.8x_2 - 0.7x_3) \\ \dot{w} &= 0. \end{aligned} \tag{11}$$

System (11) has an equilibrium: $x^* = (1.4608, 3.9307, 0.6506, 0)$, and with

$$K := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.1 \end{pmatrix}^T$$

matrix $A - KC$ is Hurwitz, therefore by Theorem 3.2 we can construct the following observer system

$$\begin{aligned} \dot{z}_1 &= z_1(2 + 0.1z_4 - 1.1z_1 - 0.1z_2) \\ \dot{z}_2 &= z_2(-1 - 0.3z_4 + z_1 - 0.2z_2 + 0.5z_3) \\ \dot{z}_3 &= z_3(3.6 + z_4 - 0.8z_2 - 0.7z_3) \\ \dot{z}_4 &= [y - (z_1 - x_1^*, z_3 - x_3^*)](1, 0.1). \end{aligned} \tag{12}$$

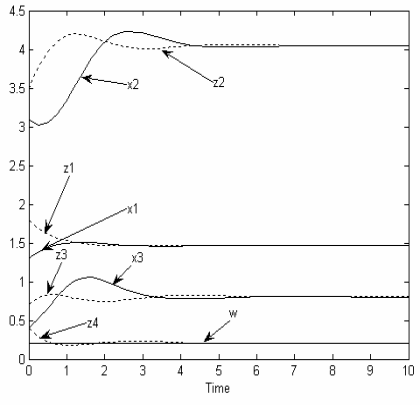


Figure 1. Solutions of systems (11) and (12)

We suppose that environmental perturbation corresponds to the value $w = 0.2$, and take an initial condition $(x_0, w_0) := (1.3, 3.1, 0.4, 0.2)$ near the equilibrium of system (11), and similarly, we consider another nearby initial condition, $z_0 := (1.8, 3.5, 0.7, 0.4)$ for the observer system (12). Figure 1 shows that the corresponding solution z tends to the solution x of the original system.

3.2. Observer for a system with exosystem describing environmental change

In this section we will design observer for a population system, supposing that in the abiotic environment there is a continuous change obeying a known dynamic law, which affects certain parameters of the population system. This abiotic process may be e.g. pollution produced by an industrial plant, a periodical (seasonal) change of temperature, or a monotonous increase of yearly average temperature due to global warming etc. In our illustrative examples below, this dynamics will be described by a system of differential equations. The situation can be handled in terms of a composite observation system of the form (7), with an abiotic dynamics as exosystem.

First we assume that, with known positive constants α and δ, the external system

$$\begin{aligned} \dot{w}_1 &= \alpha w_2 \\ \dot{w}_2 &= -\delta w_1 \end{aligned} \tag{13}$$

describes a periodic change that affects the coefficients of the interaction between predator and preys in the following form

$$\begin{aligned} \dot{x}_1 &= x_1(a_1 - b_{11}x_1 - (b_{12} + c_{12}w_1)x_2) \\ \dot{x}_2 &= x_2(-a_2 + (b_{21} + c_{21}w_1)x_1 - b_{22}x_2 + (b_{23} + c_{23}w_1)x_3) \\ \dot{x}_3 &= x_3(a_3 - (b_{32} + c_{32}w_1)x_2 - b_{33}x_3) \\ \dot{w}_1 &= \alpha w_2 \\ \dot{w}_2 &= -\delta w_1, \end{aligned} \tag{14}$$

where all c_{ij} -s are positive. Since zero is a Lyapunov stable equilibrium of system (13), by a known result (see e.g. Isidori, [1995]), equilibrium $(x^*, 0)$ of system (14) is Lyapunov stable, and the linearization of (14) is observable with the same observation

$y = h(x) = (x_1 - x_1^*, x_3 - x_3^*)$ as in the previous example. Again, by Theorem 3.3 we can construct an observer for this system:

Example 2. We consider the system

$$\begin{aligned} \dot{x}_1 &= x_1(2 - 1.1x_1 - (0.1 + 0.3w_1)x_2) \\ \dot{x}_2 &= x_2(-1 + (1 + 0.2w_1)x_1 - 0.2x_2 + (0.5 + 0.3w_1)x_3) \\ \dot{x}_3 &= x_3(3.6 - (0.8 + 0.8w_1)x_2 - 0.7x_3) \\ \dot{w}_1 &= 1.3w_2 \\ \dot{w}_2 &= -2.1w_1 \end{aligned} \tag{15}$$

Now using matrix $K := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}^T$

we obtain the following observer:

$$\begin{aligned} \dot{z}_1 &= z_1(2 - 1.1z_1 - (0.1 + 0.3w_1)z_2) + [y - (z_1 - x_1^*, z_3 - x_3^*)](1, 0) \\ \dot{z}_2 &= z_2(-1 + (1 + 0.2w_1)z_1 - 0.2z_2 + (0.5 + 0.3w_1)z_3) \\ \dot{z}_3 &= z_3(3.6 - (0.8 + 0.8w_1)z_2 - 0.7z_3) \\ \dot{w}_1 &= 1.3w_2 \\ \dot{w}_2 &= -2.1w_1 \end{aligned} \tag{16}$$

For the solution of the composite system (15), we set the initial value $(x_0, w_0) := (1.3, 3.1, 0.4, 0.1, 0.1)$, and calculate the solution of the observer system (16) with initial value $z_0 := (1.8, 3.5, 0.7)$. Figure 2 shows how the state process is approximated by the solution of the observer even in case of variable environment.

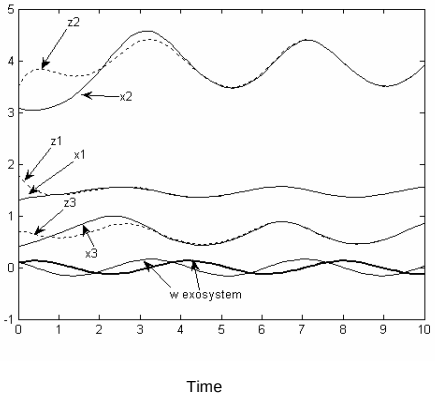


Figure 2. Solutions of systems (15) and (16)

4. DISCUSSION

As a continuation of our recent research line, the monitoring methodology has been further developed towards models with changing environment. With the applied methodology, monitoring is not restricted to the estimation of the population state process or the detection of a change in a constant intrinsic growth rate (or predation rate) of the population system model, but also makes it possible to trace the biological consequences of a continuous change in certain abiotic environmental factors, including industrial pollution and climate change.

Recently, Waldherr [2005] obtained considerable theoretical results concerning observer design for the *2-dimensional* Lotka-Volterra system. The novelty of our work consists in the application of exosystems and the estimation of the environmental effects, and our approach can also be extended from the presented simple illustrative ecosystem model to multispecies and non Lotka-Volterra type case, e.g. to trophic chains with structure *resource – producer - primary consumer*, see Shamandy [2005], without difficulty.

For the application to *real field data*, we will realize the following steps: 1. Extend our approach to discrete-time ecosystem models, based on nonlinear discrete-time observer

design, see e.g. Sundarapandian [2005]. 2. Regarding the general complexity of a real ecosystem, a first application of our approach will deal with the relatively closed and simple system of a greenhouse. 3. In a recent project proposal, we have planned to set up a monitoring system to be applied in *biological pest control in greenhouse crops*. In this case, however, the dynamic model of the involved insect populations must also include developmental stages of the insects and the dependence of the individual development on abiotic parameters like temperature and humidity. 4. Based on this model, a protocol will be obtained for the timing of the application of a biological agent (predator or parasitoid), *supporting the decision making process in biological pest control.*

ACKNOWLEDGEMENTS

The authors wish to thank the Ministry of Education and Science of Spain for the financial support of the project TIN2007-67418-C03-02, which has partially supported this work. The research was also supported by the Hungarian National Scientific Research Fund (OTKA K 62000).

REFERENCES

Gragnani, A. The role of toxicants on predator-prey dynamics. *5th ESMTB Conference on Mathematical Modelling and Computing in Biology and Medicin*. 2002, Milan, Italy. Abstracts, 179, 2002.

Isidori, A. Non linear control systems, Springer, New York, 1995.

López, I.. *Observabilidad y controlabilidad en modelos de evolución* (Tesis doctoral). Almería: Servicio de Publicaciones de la Universidad de Almería. ISBN: 84-8240-683-3. 2003.

López, I., Gámez, M. and Carreño, R. Observability in dynamic evolutionary models. *Biosystems*, 73, 99-109, 2004.

López I, Gámez M, Molnár. Observability and observers in a food web. *Applied Mathematics Letters* 20 (8): 951-957. 2007a.

López, I., Gámez, M., Garay, J. and Varga, Z. Monitoring in a Lotka-Volterra model. *Biosystems*, 83, 68-74, 2007b.

Lee, E.B. and Markus, L. Foundations of Optimal Control Theory. Wiley New York. 1971.

Scarelli, A., Varga, Z. Controllability of selection-mutation systems. *BioSystems* **65**, No 2-3, 113-121, 2002.

Shamandy, A. Monitoring of trophic chains. *Biosystems*, Vol. 81, Issue 1, 43-48. 2005.

Sundarapandian, V. Local observer design for nonlinear systems. *Mathematical and computer modelling* 35, 25-36. 2002.

Sundarapandian, V. Nonlinear observer design for a general class of discrete-time nonlinear systems with real parameter disturbance. C*omputers and Mathematics with applications* 49 1177-1194, 2005.

Varga, Z., On controllability of Fisher's model of selection. In "*Differential Equations*" (Eds. C. M. Dafermos, G. Ladas, G. Papanicolau) Marcel Dekker, New York (1989) 717-723.

Varga, Z. On observability of Fisher's model of selection. *Pure Math. and Appl*. Ser. B, 1, 15-25. 1992.

Varga, Z. Applications of mathematical systems theory in population biology. *Periodica Mathematica Hungarica* , Vol. 56 (1), 157-168, 2008.

Varga, Z., Scarelli, A. and Shamandy, A. An observability problem of population ecology. In: V. Capasso (Ed.), *Mathematical Modelling and Computing in Biology and Medicine*, 5th ECMTB Conference 2002 (Abstracts, p.189). MIRIAM, Milan, 2002.

Varga, Z., Scarelli, A. and Shamandy, A. State monitoring of a population system in changing environment. *Community Ecology* 4 (1), 73-78, 2003.

Waldherr, S. Observability Analysis and Observer Design for Controlled Population Dynamics, PhD. Thesis, University of Stuttgart, 2005

Xia, Y. Positive periodic solutions for a neutral impulsive delayed Lotka–Volterra competition system with the effect of toxic substance_Nonlinear Analysis: Real World Applications 8 204 – 221, 2007.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Creating robust sensor networks - architecture and infrastructure

G. Schimak[a], Th. Usländer[b], D. Havlik[a] and R. Argent[c]

[a] Austrian Research Center GmbH – ARC, A-2444 Seibersdorf, Austria
(gerald.schimak@arcs.ac.at; denis.havlik@arcs.ac.at)
[b] Fraunhofer IITB, Germany (uslaender@iitb.fraunhofer.de)
[c] The University of Melbourne, Australia (r.argent@unimelb.edu.au)

Abstract: Significant unregulated development is occurring in sensor devices, protocols and sensor networks especially in the areas of air and water quality monitoring and management where the focus has been on human health and exposure impacts. Satellite technology is well established as a means of deploying sensors. However, for monitoring of air and water we lack the interconnectivity of both in-situ and remote sensing devices that link to interoperable information systems. One reason for this deficiency is the lack of a sensor network infrastructure generic enough to support this interoperability and interconnectivity.

This workshop will explore the architectural aspects of creating robust sensor networks including the aspects of the processing and visualisation of sensor-related information. The discussion will take into account topics like interoperability, self-organisation, network management and sensor service architectures. Candidate open geospatial architectures specified by European initiatives and research projects are taken as examples.

Keywords: Sensor Networks; Service Infrastructure, Architecture.

1. INTRODUCTION

Sensors provide some of the basic input data for environmental monitoring as well as for risk management of natural and man-made hazards. In this paper the word ‘sensors’ covers everything from remote sensing satellites, which provide valuable images of large regions, to single instruments providing highly-detailed point-based information installed on the Earth’s surface.

Because data from sensors play such an important part in improving understanding of our environment considerable investment has been made in the development and maintenance of highly-sophisticated sensor networks.

In spite of the ubiquitous need for information from sensor networks, the use of the data gleaned from such networks is hampered in many ways.

(1) Information about the presence and capabilities of sensor networks operating in a region is often difficult to obtain due to a lack of easily available and usable meta-information.

(2) Once sensor networks are identified their data is often not straightforward to access due to a lack of interoperability between data acquisition and dissemination systems.

(3) Incompatibilities between sensors and the systems that support them often limit the ability to transfer and share data. The current situation can lead to a lack of efficiency and under-use of available data which can severely compromise important applications such as risk mitigation.

2. CREATING ROBUST SENSOR NETWORKS - CHALLENGES

2.1 Generic (Architectural) Requirements

One major challenge when specifying and developing new sensor networks is to select the right architecture at the beginning of the project. The European research project ORCHESTRA (http://www.eu-orchestra.org) has provided a generic toolbox of utilities and a knowledge base for generic risk management applications upon which future-oriented service- oriented architectures can be developed.

The architectural principles that guided the specification of the Reference Model for the ORCHESTRA Architecture (see Usländer [RM-OA, 2007]) have to be adapted and extended for applications that involve service networks, especially for those that include the access to sensors and sensor-related information. We propose to consider these principles when discussing about robust sensor network architectures and infrastructures:

- **Rigorous Definition and Use of Concepts and Standards:** A Sensor Service Architecture should make rigorous use of proven concepts and standards in order to decrease the dependence on vendor-specific solutions, help ensure the openness of service network and support the evolutionary development process of an architecture.

- **Loosely Coupled Components:** The components involved in a service network should be loosely coupled, - where loose coupling implies the use of mediation to permit existing components to be interconnected without changes to the basic system architecture.

- **Technology Independence:** A Sensor Service Architecture should be independent of technologies, their life-cycles and changes in sensor design as far as practically feasible. It should be possible to accommodate changes in technology (e.g. life-cycle of middleware technology) without changing the Sensor Service Architecture itself. The Sensor Service Architecture should be independent of specific implementation technologies (e.g. middleware, programming language, operating system). If possible, the Sensor Service Architecture should not be influenced by or deal with limitations of specific implementation technologies.

- **Evolutionary Development - Design for Change:** A Sensor Service Architecture should be adaptive and designed to evolve over time, i.e., it should be possible to develop and deploy the system in an evolutionary way. A Sensor Service Architecture should be able to cope with changes of user requirements, system requirements, organisational structures, information flows and information types in the source systems.

- **Component Architecture Independence:** A Sensor Service Architecture should be designed such that a service network and source systems (i.e. existing information systems, sensors and sensor networks) are architecturally decoupled. This means that a Sensor Service Architecture should not impose any architectural patterns on source systems for the purpose of having them collaborate in a Service Network, and no source system should impose architectural constraints on a Sensor Service Architecture.

- **Generic Infrastructure:** Services should be independent of the application domain. This means that Services should be designed in such a flexible and adaptable way so that the Services can be used across different thematic domains and in different organisational contexts. Also that the update of integrated

components (e.g. sensors, applications, systems, ontologies) have minimal impacts on the users of the Services.

In the following, the architectural approach of the European research project SANY (http://www.sany-ip.eu) is described as an example how to address such architectural principles in an open, geospatial service-oriented architecture.

2.2 SANY Sensor Service Architecture

The SANY project argues that the development of robust sensor networks encounters obstacles and challenges not only in the sensor domain but also in higher-level functional domains in order to further process sensor-related information and present it according to the end-user's needs. These functional domains are illustrated in Figure 1 by Usländer [SANY, 2007].

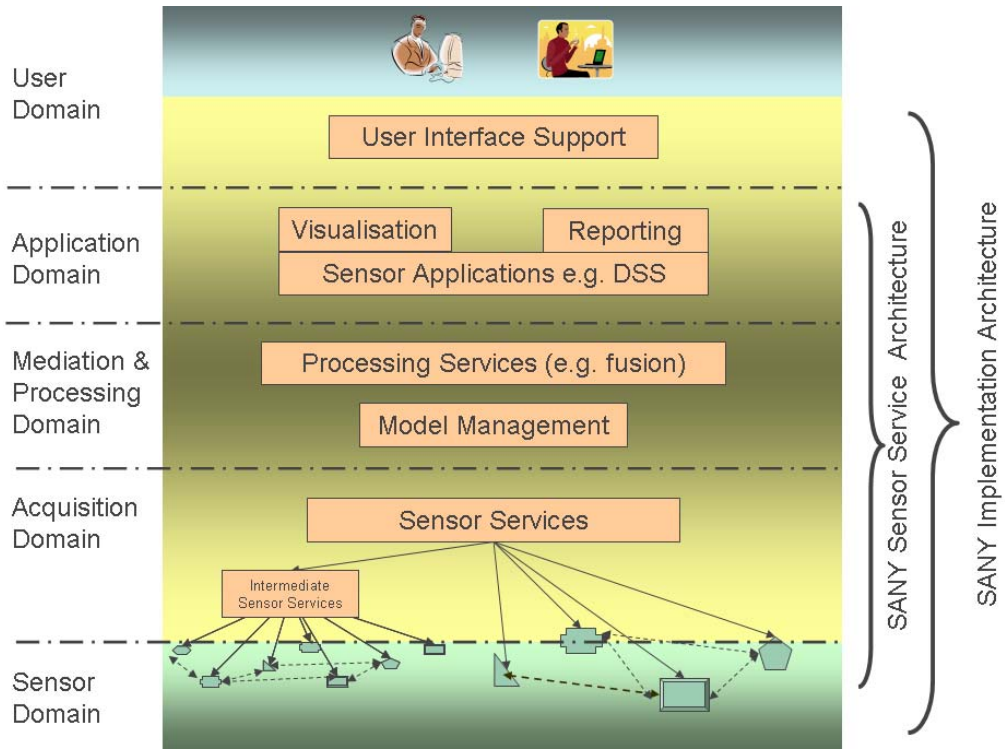


Figure 1: Functional Domains of the SANY Sensor Service Architecture

The major component of the SANY Sensor Service Architecture is the ORCHESTRA Architecture which is described in Usländer [RM-OA, 2007] as an "open architecture that comprises the combined generic and platform-neutral specification of the information and service viewpoint as part of the ORCHESTRA Reference Model" (RM-OA). The SANY Sensor Service Architecture builds upon some essential features of the ORCHESTRA Architecture. These comprise:

- Rules and guidance on specifying information and service models in the context of international standards.
- Basic re-usable specification units for information models (e.g. pre-defined feature types) and service models (e.g. re-usable interfaces).
- A series of textual descriptions and formal specifications of generic ORCHESTRA Architecture Services that are application and technology independent.

The SANY project intends to re-use and apply these specific components in the context of the SANY Sensor Service Architecture to the extent that they enable the development of service networks based on sensor information and higher-level sensor information processing.

2.3 Functional domains

SANY services are designed to serve the needs of applications and users. Although there is no prescribed hierarchy of services, SANY services may be grouped into functional domains that match their design purpose. The SANY Sensor Service Architecture distinguishes between the following functional domains of SANY services (see Figure 1).

- **Sensor Domain**: SANY services in the sensor domain cope with the configuration and the management of individual sensor devices and their organization into sensor networks. Examples are services that support communication between the sensors themselves, e.g. a take-over service in case of an impending sensor battery failure. Services in this domain are abstractions from proprietary mechanisms and protocols of sensor networks.

- **Acquisition Domain:** SANY services in the acquisition domain deal with access to observations gathered by sensors. This includes other components in a SANY Sensor Network (e.g. a database or a model) that may offer their information in the same way (as observations) as sensors do. They explicitly deal with the gathering and management of information coming from the source system of type “sensor”. The information acquisition process may be organized in a hierarchical fashion by means of intermediate sensor service instances (e.g. using data loggers).

- **Mediation and Processing Domain:** SANY services in the mediation and processing domain are not specific to the sensor service architecture. They are specified independently of the fact that the information may stem from the source system of type “sensor”. They mediate access from the application domain (see below) to the underlying information sources. They provide generic or thematic processing capabilities such as fusion of information (from sensors and other information sources), the management of models and the access to model results. In addition, service support for the discovery of sensors, data and services, naming resolution or service chaining are grouped in the mediation and processing domain.

- **Application Domain:** Based on services within the acquisition and processing domain, SANY services in the application domain support the rendering of information in the form of maps, diagrams and reports directly to the end-user in the user domain.

- **User Domain:** The functionality of the user domain is to support the system interface to the end user. SANY will not specify dedicated services to support the user interface. However, when building real systems and applications, such functionality is essential. SANY considers this functionality to be specified in a dedicated implementation architecture that also may take proprietary components and products into account.

3. USER REQUIREMENTS

Beyond the architectural requirements of SANY - the workshop will also foster discussion concerning user requirements within various environmental domain(s). In the SANY project users from different environmental monitoring domains (air pollution risks, marine risks, geo-hazards) have identified their most important needs. An analysis of the various use cases has led to a set of requirements that have been grouped into functional blocks as illustrated in Figure 2 and listed below [Watson/SANY 2007].

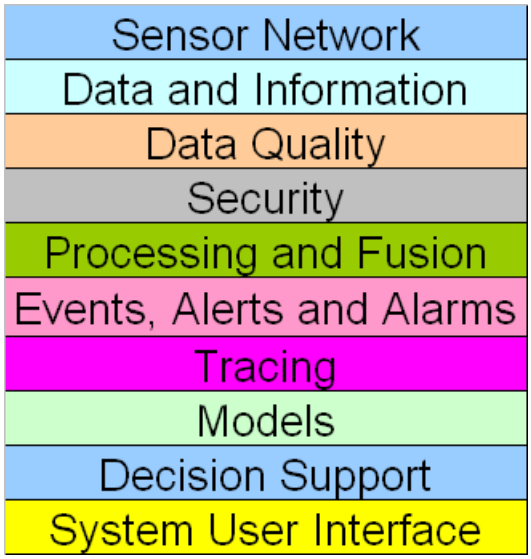


Figure 2: User requirement groups for sensor (service) networks

Here, the term "system" stands for the entirety of a sensor service network including its architecture, its hardware and software components.

3.1 Sensor Network

- **Plug & measure** refers to the capability of being able to add a new sensor node without a manual re-configuration of the sensor network or sensor node.
- **Dependability** is required to provide data access and management services, in order to cope with the dynamic availability of possibly redundant sensor data sources, especially in the case of mobile sensors.
- **Sensor network management**. Of particular interest here is the localisation of sensor nodes, e.g. for the planning and management of their deployment or the configuration of the measurement frequency in order to optimise network and battery load.
- **Deployment of mobile ad hoc sensor clusters**. Especially in the case of biological and chemical hazards, the responsible administrative authority needs to monitor air pollution or water quality in order to quickly assess environmental risk. Unfortunately, appropriate sensors are often not available at the sites they are needed.
- **Self-validation of sensor nodes** with regard to its residual battery life and measurement capability (needed for re-calibration or maintenance) is important for the assessment of the node deployment and data quality.
- **Battery life-time optimisation** through selective data transmission is a useful function for assessing information via an interface to sensor nodes, often with self-diagnosis capability. It should support the ability to automatically select between alternate data transmission routes and/or choose the frequency of data transmissions, if the residual battery level of a sensor is too low.

3.2 Data and Information

- **Data sources** not only include sensors and databases of archived data, but also data obtained from a laboratory analysis of samples, or data entered manually by humans. Data sources may also be results of fusion services.
- **Spatial and temporal metadata.** The system should be capable of adding spatial metadata (sensor location information) and temporal metadata (sensor measurement time information) to any measurement data with/without this metadata from outset.
- **Rapid access to data.** The system should be capable of accessing data in sufficiently short time.
- Data type. The system should operate with different data types (e.g. fields, coverages) associated with measurement data like:
    - a single sensor measurement observation represented by a single value and unit located at a fixed location and detected at a particular time (time, date and year). For critical evaluation situations it has to be distinguished between a measurement result with a sensor internal clock time stamp or the nearly adequate particular time when the result has been received by the system.
    - time series of sensor measurement observations represented by a time series of triples {value and unit, feature of interest, time}. The time representation may be either absolute time (time, date and year) or a relative time representation (e.g. every day at midday starting from a certain date).
- **Geographic objects of several types**. The system should be capable of handling geographic objects of several types (e.g. the types specified by the ISO 19107 and ISO 19123 such as MultiPoint, MultiLineString, MultiPolygon, LineStrings, Polygons, MultiCurve, MultiSurface).
- **3D fields.** The system should be capable of operating with three-dimensional measurement data (e.g. components of vector fields etc.) of an observed property (e.g. wind speed, water current etc.)
- **Images (e. g. IR/UV, SLAR, INSAR).** As an example, such image data may be delivered by an earth observation sensor mounted on an aircraft or on a satellite with a certain spatial resolution.
- **Maps.** The system should be capable of handling a wide spectrum of maps (topography, roads, land usage etc).

3.3 Data Quality

- **Data quality.** To assess the quality (e.g. uncertainty) of measured data information about the quality of its measurement (e.g. accuracy, tolerance, resolution, drift) should be available. The system should be capable of accessing this sensor data quality assurance information.
- **Sensor level spatial and temporal uncertainty.** To ensure a previously defined user data quality target level, the system should provide details of data quality levels according to:
    - available sensor data sources which may vary by location ,
    - estimated data statistical uncertainties derived from past data (e. g. model temperature predictions at a given location depends on the distance to available sensor values and on topological conditions e.g. located on mountain), and
    - the time period and frequency of measurement observations (limited e.g. by a certain availability frequency of EO satellite images of an area)
- **Certification of data and its propagation.** The system should be capable of supporting the process to formally certify the data quality, e.g. by instructing a certified laboratory per email with an investigation of new microbial sampling at a specific beach.

- **Associate data quality with the measurement context (validation).** The system should be capable of associating and storing data quality together with measurement context as meta-information. The system should be capable of visualizing this meta-information in its spatial and temporal context. This feature may be used as a tool for quality validation and as an aid to proper placement of sensors.

3.4 Security

- **Confidentiality of data** - UAA (User Management, Authentication, Authorisation). The system should provide several security related services:
  - Authentication Service (verification of user identities, support for multiple authentication mechanisms)
  - Authorisation Service (support of roles, authorisation for services authorisation on sensor level, etc)
  - User Management Service (storage and management of user profiles)
- **Monitoring Service** (checks the availability of services and sensors)
- **Billing, encryption, track data ownership.** The provision of sensor data requires a significant investment in equipment and in supporting services. These services normally cannot be offered free of charge, therefore a billing service is necessary. Another security topic is to protect measurement data against manipulation. In order to ensure data authenticity encryption may be necessary at different operational levels (e.g. during transmission, storage in databases etc.). It may be a security problem to allow certain users of the system to track sensor data over a period of time (e.g. a shipping company does not want competitors to know the exact routes of their ships).

3.5 Processing and Fusion

- **Interfaces for data processing services development.** The system should provide general data processing services e.g. merging data, extracting relevant data for reports. In general, the system should be capable of handling meta-information. This includes the storage of intermediate information in order that overall services (and service chains) can execute with acceptable speed.
- **Image analysis and feature extraction.** The system should be capable of processing images and extracting features (e.g. such as a road or a watercourse).
- **Different time steps, resolution.** The system should be capable of handling and adjusting different temporal and / or spatial resolutions of sensor data, depending on the sensors used and if values are missing.
- **Improving spatial resolution** if the spatial resolution of different sensor data is too heterogeneous or if values are missing.
- **Fusion of measurements.** The system should be capable of processing and merging measurements of the same phenomenon using different sensorplatforms. The information about the source should be stored in meta-information. Data gathered at different times but generated by the same equipment should be processed using a fusion service.
- **Library of algorithms as ((statistical) processes) services.** The system should be capable of providing statistics libraries e.g. for interpolation and merging services.
- **Workflow.** The system should be capable of performing service chaining according to workflows. The output produced by a model service may be connected to the input of sequential process.
- **Visualisation.** In many situations several forms of visualisation are prescribed, e.g. maps, diagrams. To display them in a flexible manner styled layers should be capable of being output separately..

3.6 Events, Alerts and Alarms

- **Threshold surpasses detection.** The system should be capable of detecting threshold surpassing. For instance in the field of air quality measurement, pollution of the environment can be asserted if data values obtained from sensors reach defined threshold levels.
- **Alert algorithms.** Typically, an alert will result if a number of data observations for pre-selected parameters no longer fall within a prescribed range ((data plots outside the desirable operating range of values). Alerts cause application level procedures and / or workflows to be executed in response to the event.
- **Interfaces for alarm management.** Alarms provide communication procedures associated with the current environment to warn about an imminent hazard and to initiate emergency procedures (such as evacuation).

3.7 Tracing

- **Tracing requirements** address the need to document what information sources were used as a basis for the decisions taken and the decision making process itself. The purpose is to be able to provide a retrospective justification for decisions made, these decisions may later be contested by affected parties to the decision.

3.8 Models

- The system should provide in general
    - a model service catalogue, where all available services may be selected by the user
    - an execution management service that handles input/output and, (optionally if expedient for system performance) data by reference
    - the capability of using the output of models and fusion algorithms as sensor values without changing the system.

  In addition to basic support for models and model wrapping (the process of integrating the models into the SANY service network), model-specific functionality is required in the following areas:
    - Gather applicable source impact models
    - Domain skills compilation
    - Library of models as processes
    - Library of geo-statistical analysis as processes
    - Predictive models for adaptive sampling
    - Spill advection & dispersion modelling
    - Forecasting risks (water quality, bathing water, beach closed)
    - Model long-term degradation of ecosystems
    - Improved soil models

3.9 Decision Support

- **Providing supporting information for decisions**. The types of decisions to be taken will dictate the need for modelling, data fusion and visualization of model results and other information. Auxiliary methods may be required to compute utility functions and to undertake multi-criteria analyses.

3.10 System User Interface

- **User registration**. The system should offer the user a dialog window to login normally with user name and password. When entering the password the input window should not display the characters in a readable form.

- **Message to the user.** The system should be capable of sending a message to the user in a separate window. This window must be visible to the user (e.g. it pops into the foreground).
- **Sending form(s) to the user.** The system should be capable of sending a form or several forms to the user to and to encourage them fill them out with the required information.
- **Sending message that authorisation failed.** If an authorisation has failed (e.g. due to insufficient access rights), the user should be informed by a message.
- **Sending message that authorisation failed and waits for a new input.** If an authorisation has failed (e.g. due to a wrong password), the user should be informed by a message and is requested to repeat the login procedure.

4. CURRENT EXAMPLES FROM ONGOING PROJECTS

4.1 ORCHESTRA and SANY

The European Commission (EC) is funding a number of Integrated Projects within the Sixth Framework Programme concerned with improving the accessibility of data and services for risk management. Two of these projects, 'Open Architecture and Spatial Data Infrastructure for Risk Management' (ORCHESTRA, http://www.eu-orchestra.org/) and 'Sensors Anywhere' (SANY, http://sany-ip.eu/), were introduced earlier and some pertinent features of these projects that relate to sensor networks were briefly discussed.

These projects have developed an open distributed information technology architecture and have implemented web services for the access and use of the data emanating, for example, from sensor networks. These developments are based on existing data and service specifications proposed by international standardisation organisations such as ISO and the Open Geospatial Consortium (OGC)..

The projects seek to develop the principal features of the EC INSPIRE initiative (http://inspire.jrc.it), which was launched in 2001 and began this year to implement elements relevant to risk management.

Thanks to the open nature of the architecture and the services being developed within these projects, the technology can be implemented by any interested party and can be accessed via the web by all potential users of the software. The architecture is based on a service-oriented approach that makes use of Internet-based applications (web services) whose inputs and outputs conform to recognized standards.

The benefit of this philosophy is that it is expected to favour the emergence of an operational market for risk management services in Europe. It eliminates the need to replace or radically alter the multitude of already operational IT systems in Europe, drastically lowering costs for users [Klopfer/Kannellopoulos 2008]. Furthermore, it allows users and stakeholders to achieve interoperability while using the system most adequate to their needs, budgets, culture etc. (i.e. it has flexibility).

4.2 Architectural Requirements put on Sensor Networks from initiatives like GMES

The 'Global Monitoring for Environment and Security' (GMES) represents a concerted effort to bring data and information providers together with users, so they can better understand each other and make environmental and security-related information available to the people who need it through enhanced or new services.

For example, the GMES initiative (http://www.gmes.info) as illustrated in Figure 3 below focuses on four major components:

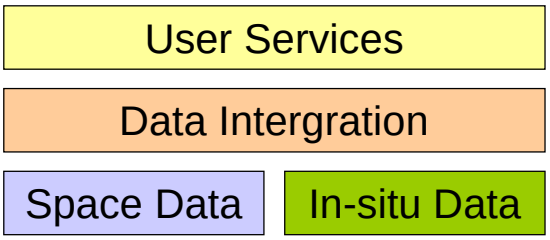


Figure 3: Major GMES components

The final report for the GMES initial period [GMES FR, 2004] provides the following key architectural requirements on pages 25-26:

"For GMES to become a success, the architecture needs to facilitate the integration of standalone data and information elements. It should allow to the selection and aggregation of information from heterogeneous sources and should provide the capability to translate data and information between the various sources in real time. This applies as much to the incorporation of socio-economic data and information, as well as products derived from the space and in situ observing networks.

GMES must therefore provide a structured framework for data integration and information management, i.e., a European shared information capacity. The following key architectural and user-oriented requirements will therefore drive the implementation of GMES:

- *Openness, based on agreed open standards, facilitating seamless communication and interoperability, i.e. the ability of different devices or systems (usually from different vendors) to work together, as well as enabling user service autonomy;*
- *Federated architecture, enabling systems to grow and evolve;*
- *Simplicity of architecture (e.g. modularity of components), to break the complexity barrier, systems must be made easier to design, administer and use;*
- *Self-configuration, programmability, scalability (e.g. to handle various levels of operational load and external conditions);*
- *Dependability, i.e. the system's resilience to security threats or breakdown;*
- *User-friendliness of services and interfaces, e.g. in the handling of user request services, access control, workflow management, delivery management, visualisation, data extraction (e.g. "multilinguality"), multiuser sessions, administration;*
- *Data security, protection of provider and user data against alteration, theft and misuse;*
- *Quality of service;*
- *Ubiquity of access, including global reach."*

In a overview paper prepared by DG [GMES, 2005] on Data Integration and Information Management Capacity, provisional and important functional requirements are listed. The paper suggests that the approach needed to create an efficient data integration and information management component is to use a "system of systems" design as explained below:

"The GMES information and services infrastructure goes well beyond a simple exchange of data. It addresses the need to integrate business and workflow processes which span across boundaries, be they political, administrative or thematic. It addresses the main issues of heterogeneity and fragmentation of the information: "heterogeneity" meaning that the same information is often represented differently, "fragmentation" meaning that needed information is spread over multiple locations.

The approach is to seamlessly integrate existing systems into a "system of systems" perspective. System of systems is an emerging design and development method of complex systems build from large scale component systems. The subsystems that comprise the system of systems are generally built by different organisations, having different goals, are very often built to different quality standards, and are managed independently.

Reusing existing legacy systems in a dependable fashion without the need for extensive re-engineering is a key problem currently faced by industry. System of systems can be seen as new systems linking data, services and workflows to produce new data, new services together with metadata (information about the information products generated[1]).

This approach, which was a research topic in the last decade, is now becoming sufficiently mature. In the short term, robust system architectures will be developed and tested allowing the exchange of data and services that are well identified: This is the syntactic interoperability phase. In the longer term, we expect to achieve significant semantic interoperability, which will allow cross-system search for data and services. It could be dubbed the "GMES Google" since it will work for data and services in the way web search engine works for web pages.

A number of non technical obstacles should however be addressed to eventually reduce the complexity of the implementation, such as trusted electronic billing principles between data providers and services providers to partly overcome the hurdles created by different data policies. Mechanism should also be agreed upon to manage access privilege across institutional borders in a practical, transparent and secure way.

This approach will allow sharing and efficient management of information that is consistent across organisations, borders and thematic domains such as from land use and mapping to risk management and security."

The suggested system of systems approach is still regarded today as the best way forward as indicated in the following excerpt of the report of Joint Operability Workshop held in April 2007:

"There was broad agreement that any proposed solution for a single information space for the environment must allow for multiple architectures and hence be a "system of systems". Further comments on this topic included:

- *Hierarchical architecture is required to cope with services at various levels: EO ground segment services, GMES core services, GMES downstream services etc.*
- *SOA is well suited for deployment of system of systems.*
- *Significant investments have been made already by ESA, ECMWF[2]etc. Therefore it is unrealistic that everybody will adopt a single architecture. This leads to a system of interoperable systems which may be based (internally) on heterogeneous technologies.*

[1] In the SANY Service Architecture therefore called "meta-information".
[2] European Centre for Medium range Weather Forecasting

- *An exhaustive list of use cases is unfeasible – the number of possibilities is too large and is changing rapidly. Thus, a flexible architecture is needed which also leads to the concept of a system of systems. The fully top-down approach and pan-European architecture that was anticipated as recently as five years ago appear now to not be feasible. The concept is being replaced by pan-European interoperability in a system of systems-like architecture.*

- *Many systems are legacy systems in this field as satellites have a relatively short life and scientific advances in EO are rapid. It is totally unfeasible to convert all existing data so they all have the same data and metadata formats. Any viable solution must be based on fully utilising legacy systems. Harmonised, "on the fly" access to legacy datasets is required.*

- *Access to data from heterogeneous missions must be seamless and transparent."*

As recommended in the following extract of the overview paper [GMES, 2005] on Data Integration and Information Management Capacity, the technology of choice to be used to implement the GMES architecture is SOA: "The underlying architecture for ensuring interoperable GMES services is the Service Oriented Architecture (SOA)".

Modern system architecture for building environmental sensor networks are based on a less tightly coupled collaboration of distributed services. For this, the term SOA is becoming widely used, but there is little agreement in the manner by which it is used. The World Wide Web Consortium (W3C) for example refers to SOA as 'A set of components which can be invoked, and whose interface descriptions can be published and discovered'. This broad definition highlights a key aspect of an SOA: components (e.g. functionalities) can be discovered and invoked dynamically. This type of architecture is different from "hard wiring" of business processes, which make changes according to the recognition of new circumstances very difficult. SOAs are inherently more flexible and adaptable than most alternate approaches.

A service-oriented architecture is also proposed for the following reasons:

- A SOA can represent business functionality remaining implementation- and vendor- neutral while providing standards-based shared services.
- Existing legacy systems can be extended with service interfaces, and, in that way, become part of a SOA.
- By using inherent service discovery and access flexibility (see above), a SOA enables GMES/INSPIRE service providers to be more agile and to respond more quickly to changing business opportunities and changing requirements.
- A set of common generic services can provide standard, domain-independent functionality (for discovery, search, navigation, data access, authorisation etc.) which only needs to be implemented once.
- Sharing of services — no need to "re-invent the wheel"
- Loose coupling — ability to update applications with minimal effect on services that invoke them
- Location transparency — ability to re-host applications with minimal effect on services that invoke them

GMES applications are built on top of a joint, adopted infrastructure that will be interoperable and much easier to integrate into multi-purpose, cross-application operations."

The adoption of SOA architecture based on open standards for its services is a good path to take in order to provide flexibility, stability, and durability while preventing vendor lock-in. Example of open standards are the OGC standards for Web services.

The relevance of these two major European initiatives, INSPIRE and GMES, for the set-up and design of robust sensor networks and higher-level service infrastructures will be a topic for the discussion at the workshop, too.

ACKNOWLEDGEMENTS

The authors wish to thank the SANY and ORCHESTRA Consortium for their valuable input. Special thanks to the architectural working group (SP2) of SANY for providing most of the requirements input listed above.

SANY and ORCHESTRA are Integrated Projects co-funded by the Information Society and Media DG of the European Commission within the RTD activities of the Thematic Priority Information Society Technologies.

REFERENCES

GMES FR, 2004 Global Monitoring for Environment and Security (GMES): Final Report for the GMES Initial Period (2001-2003) http://www.gmes.info/action_plan/index.html

GMES, GMES reflection paper on Data Integration and Information Management Capacity, DG-INFSO, Draft 6, July 2005

Klopfer, M. and I. Kannellopoulos (eds.), ORCHESTRA – An Open Service Architecture for Risk Management, http://www.eu-orchestra.org/docs/ORCHESTRA-Book.pdf, ISBN 978-3-00-024284-7, 2008.

Usländer, T. and SANY-Consortium, SANY-Deliverable, D2.3.1 Sensor Service Architecture V0 (D2.3.1 Sensor Service Architecture V0 _Doc. V1.0_.pdf), 2007 (http://sany-ip.eu/publications/1957).

Usländer, T. (Ed.). Reference Model for the ORCHESTRA Architecture (RM-OA) Version 2 (Rev.2.1). OGC Best Practices Document 07-097. http://portal.opengeospatial.org/files/?artifact_id=23286, October 2007

Watson, K. and SANY-Consortium, SANY Project Deliverable D2.2.2 Sensor Scenarios and Requirements Version 1.2, (to be published).

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A taxonomy of robustness in sensor service networks

Kym S. Watson and Thomas Usländer[a]

[a]*Fraunhofer Institute IITB, Fraunhoferstr. 1, 76131 Karlsruhe, Germany (kym.watson| thomas.uslaender@iitb.fraunhofer.de)*

Abstract: We consider which aspects of a sensor service network are important in achieving robustness, i.e. the ability to adapt or react to changes and external influences. The sensor service network is viewed as a collection of functional domains ranging from the Sensor Domain up to the Application and User Domains. The characteristics of robustness relate not only to resilience to communication problems, but also to dynamic sensors and services and to the overall system architectural design.

Keywords: Robustness, sensor, service network

1 INTRODUCTION

The word *robust* stems from the Latin *robustus* meaning oaken (like oak wood) or hardy and strong. A widespread oak tree is the *Quercus robur*, commonly called English oak. *Robustness* implies resilience and the ability to adapt to changes and external influences. The degree of robustness is determined by the types of changes and influences which can be coped with. The attribute robust is broadly applicable to practically all entities, including living beings, natural and man-made objects, processes and technical systems.

Ideally, a robust object or system has its own inherent ability to react and / or adapt to changes and external influences, even to those for which no explicit reaction has been foreseen.

The workshop contribution will propose a classification of factors characterizing robustness in a sensor service network. The essential issues of what should be made robust to which changes and influences are summarized in the sections below.

2 FUNCTIONAL SCOPE OR WHAT MUST BE ROBUST

The complete functionality of a sensor service network can be structured into domains as shown in Figure 1. Layer-like interfaces are not implied between the domains, instead the domains should be considered as a loose functional structure.

The Sensor Domain covers the actual sensor or transducer technology, and the communication within the sensor network (wired or wireless such as ZigBee) and to gateway nodes with access to higher level networks, which are typically IP based. The sensor nodes may be mobile and belong to varying sub-networks. Information sources such as a database of archived data or a model-based algorithm can both be considered as sensors as defined in the Observation & Measurement Model of OGC, cf. Cox [2007].

The overlying functional domains serve to find, integrate, generate and present information originating in the Sensor Domain. The overall goal is of course to achieve robustness in the User

Domain where decisions are made. Robustness must be built in to each functional layer for the overall system to be robust. This is illustrated below for several aspects of the lower layers.

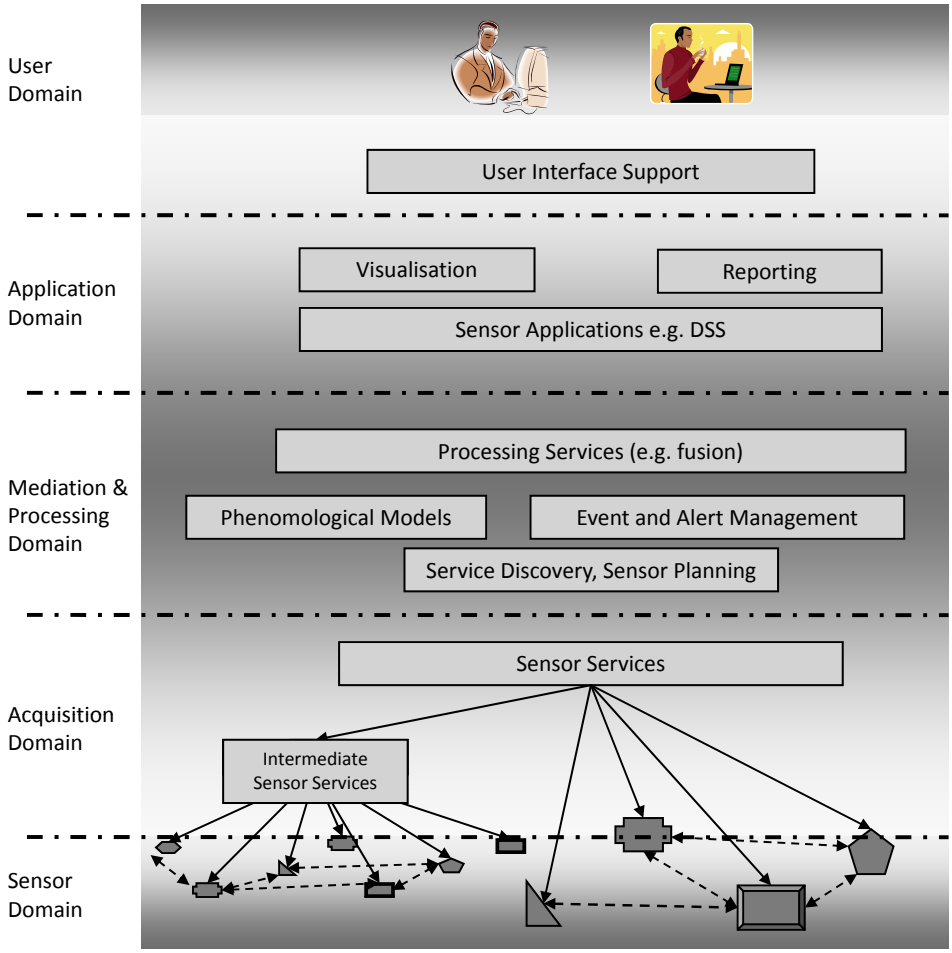


Figure 1: Functional Structure of a Sensor Service Network as defined in the project SANY

3 CHARACTERISTICS OF ROBUSTNESS

3.1 Architectural Robustness

The architecture shall make use of proven concepts and standards in order to decrease the dependence on vendor-specific solutions, to ensure openness of the service network and to support the evolutionary development of the system. The main architectural characteristics in this respect are (cf. Usländer [2007]):

- The system components shall be *loosely coupled*, implying the use of mediation to allow components to be interconnected without changes.
- The architecture shall be *independent of technologies* such as sensor communication protocols, middleware and programming languages as far as possible, and also *independent of the structure or format of existing information sources*.

- The architecture shall be *designed for evolution* in order that new user requirements and deployment forms can be handled easily.
- The service infrastructure shall be *generic* in the sense that it is independent of a particular application domain.
- The architecture shall follow *recognized standards* (such as from ISO, OGC, OASIS, W3C) and *strategic initiatives* such as GMES, GEOSS and INSPIRE.

3.2 Robustness in the Sensor domain

A wireless ad-hoc sensor network, probably with energy autarkic or battery powered nodes, is susceptible to disturbances and must be able to dynamically reconfigure itself when a communication link or sensor fails. Sensors can be mobile and belong to changing network configurations.

3.3 Robustness in the Acquisition Domain and Mediation and Processing Domain

Resilience to uncertain data: Data uncertainty in the Sensor Domain is caused not only by measurement inaccuracy, but also by missing sensor values which have to be replaced by other sensor values or approximated using a spatial-temporal fusion of available data and, possibly, models of the phenomenon of interest. In the latter case the system reaction is in the Mediation and Processing Domain. A sensor planning service can help to deploy sensors where most needed. For critical applications, redundancy in the Sensor Domain is required, too.

Resilience to changing sensors and sensor services: The system must be able to cope with changing, possibly very voluminous sources of information (sensors, models) and services. These sources need to be discovered by registration or harvesting and integrated into the system. The ideal is *plug & measure*, implying a minimum of manual adaptation to new sources.

Resilience to heterogeneous data formats and property names: This refers to the ability of services to map data provided with different names for the same physical concepts and in different languages or formats. Semantic and ontological techniques are required here.

4 CONCLUSIONS AND RECOMMENDATIONS

Robustness of a sensor service network can only be achieved by exercising a bundle of measures across the functional domains of the overall system. The reward for this work is expressed in terms of longer system life, lower life cycle costs and greater user benefit.

ACKNOWLEDGMENTS

The authors wish to thank the Information Society and Media Directorate–General of the European Commission (DG–INSFO) for co-funding the integrated projects SANY and ORCHESTRA within the area Information Communication Technologies for Environmental Risk Management. The results of these projects were the primary sources for this paper.

REFERENCES

Cox, S. Observation & Measurements - Part 1: Observation Schema, OGC Document 07-022r1. approved as OpenGIS specification, October 2007. http://portal.opengeospatial.org/files/?artifact_id=22466&version=1.

Usländer, T. Reference Model for the ORCHESTRA Architecture Version 2 (Rev. 2.1). OGC Best Practices Document 07-097, October 2007. http://portal.opengeospatial.org/files/?artifact_id=23286.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Advanced Cascading Sensor Observation Service

D. Havlik, G. Schimak, R. Barta

[a] *Austrian Research Centers GmbH. – ARC, A-2444 Seibersdorf, Austria*
(denis.havlik@arcs.ac.at)

Keywords: Open Geospatial Consortium (OGC), OGC Sensor Web Enablement (OGC-SWE), SANY IP, ORCHESTRA IP, Sensor Observation Service (SOS), cascading SOS, translating SOS, UWEDAT, Global Earth Observation System of Systems (GEOSS), Global Monitoring for Environment and Security (GMES)

Main goals of "Global Monitoring for Environment and Security" (GMES, see http://www.gmes.info/), and "Global Earth Observation System of Systems" (GEOSS, see http://www.earthobservations.org) can be summarized as: "how to remove the obstacles for sharing environmental information?".

Austrian Research Centres (ARC) participates in development of the IT-infrastructure, architecture and services capable of addressing this challenge in the context of risk management through ORCHESTRA IP (http://www.eu-orchestra.org/), and in the context of environmental monitoring networks through SANY IP (http://sany-ip.eu/).

In order to minimize the cost associated with "bridging the technology gaps" between the networks, ORCHESTRA proposes an abstract architecture and a methodology for mapping this abstract architecture onto the technical platform of choice. At the level of information exchange, ORCHESTRA demands self-description of all services and data offered in a network. In particular: "*all meta-information MUST be provided at least in a form suitable for interpretation by humans*"; "*syntactic meta-information MUST also be provided in a form suitable for interpretation by machines*"; and "*providing semantic meta-information in a form suitable for interpretation by machines (e.g. by means of an ontology) is highly encouraged*" [Schimak (Ed.), 2007].

An analysis of the service specifications developed by Open Geospatial Consortium "Sensor Web Enablement" working group (OGC SWE; see http://www.opengeospatial.org/projects/groups/sensorweb) performed by SANY IP [Havlik et. al 2007; Bartha,(ed.), 2006] has shown that OGC SWE services are both very generic and compatible with the ORCHESTRA architectural model [Usländer, 2007]. Consequently, SANY IP decided to embrace the OGC SWE service interfaces as they are, and concentrate on implementing the advanced "GEOSS (software) building blocks", extending the functionality and optimizing the overall sensor network performance. One such "building block", which will be firstly discussed at this workshop, is the advanced cascading SOS service (SOS-X).

As its name suggests, the cascading SOS is a client to underlying SOS service(s) [Na, 2007] and provides alternative access routes to users (or services) interested in accessing the data. In its simplest form, cascading SOS *replicates* part of the data offered by underlying SOS service with no changes to the data model (Figure 1a). This kind of service replication can be e.g. used to assure controlled access only to a subset of data offered by lower-level SOS, without relying on complex security mechanisms. In fact, many of the organisations ARC collaborates with use this type of gateway services as a reliable

mechanism for separating the information that may be seen only within their own network from those offered to a wider public.
Second example of the cascading SOS deployment is the *simple aggregation* of data from several underlying services; e.g. aggregation of offerings from several data loggers into complete sensor network offering, or aggregation of offerings from EU states on a single service (Figure 1b).

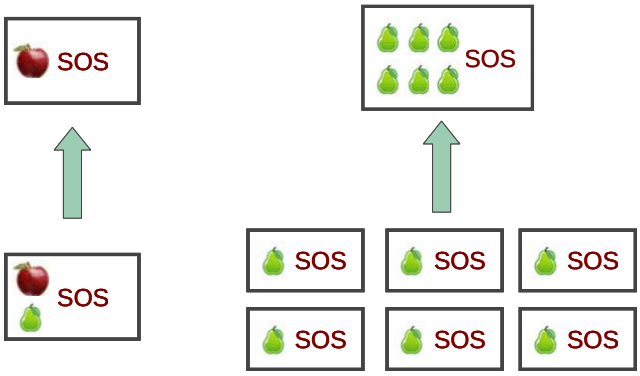


Figure 1: Two forms of simple cascading: (a) replication; and (b) aggregation

Simple SOS replication and *simple SOS aggregation* described above are only feasible in the context of a single environmental domain with a single data model accepted by all involved parties. Unfortunately, this is rarely a case in the context of GEOSS or GMES, because (a) information models often aren't standardized above the level of a single state; and (b) different tasks require different views to information.

With this in mind, we decided to develop a service capable of transforming the SOS information model (Fig. 2), similar in spirit to ORCHESTRA Translating Feature Access Service (FAS-X) [Anders, 2008]. Following the ORCHESTRA naming convention, the advanced form of the cascading SOS service may be called "Translating Sensor Observation Service" (SOS-X) in later versions of the SANY architecture.

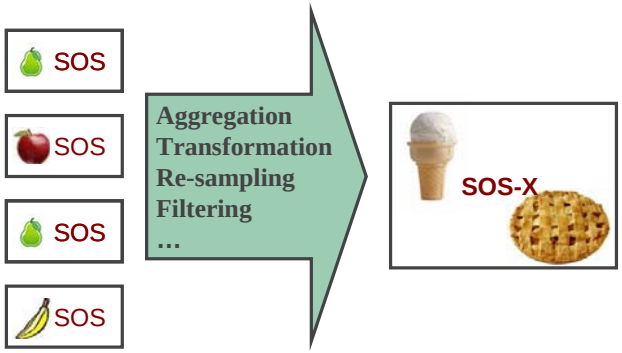


Figure 2: Advanced SOS cascading with SOS-X

Most of the SOS servers in SANY IP organise information in time series. Consequently, a decision was made to integrate an enhanced version of the UWEDAT "formula server" [Schimak, 2003] called "Formula 3" (F3) into SOS-X prototype. F3 shall be described in details in a separate publication. For the sake of current discussion, it is important to understand that F3 works independently from UWEDAT monitoring system, and that it allows arbitrary algebra operations on time series, annotation, and re-mapping of the meta-information.

As a result SOS-X shall be capable of sophisticated pre-processing of the data from underlying SOS servers, including on the fly information model transformation; unit conversions, temporal and spatial interpolation; temporal and spatial re-sampling

(interpolation and aggregation/averaging); and filtering (e.g. finding the pollution episodes; merging of offerings; cutting the offerings in smaller sub-offerings; etc.).

At the time of writing this article, the cascading SOS is in the early prototype phase, and some parts of the functionality are still under discussion. Current prototype is capable of simultaneously accessing the data from one or more SOS servers and from the ARC's UWEDAT monitoring system, and exposing them over SOS interfaces with no changes to the information model. F3 pre-processing engine itself is developed in parallel, with the intention to integrate the two pieces of software into first SOS-X prototype and evaluate its usability in SANY architecture and pilot applications (v1) by the end of 2008. This leaves enough time to development the improved version of the SOS-X (v2) before the end of the SANY IP (Q4 2009). Depending on the feedback from v1 evaluation, SOS-X v2 development may either concentrate on stability, interface standardization and user friendliness, or on additional features.

In the long run, ARC is interested in expanding the SOS-X into universal "bridging" service that would minimize the price of sharing the information in the context of GEOSS and GMES. Therefore, the SOS-X implementation architecture foresees a plug-in mechanism for extending the cascading mechanism to other sources of information, such as data files, proprietary legacy systems, ORCHESTRA Feature Access Service, or OGC Web Coverage Service. As a proof of concept, SOS-X can be already used as a front end to UWEDAT monitoring system (Figure 3).

Another interesting feature which may be integrated in for the v2 of the SOS-X service is pre-fetching and caching of information. This could significantly diminish the time needed to query the archived information, especially in the case when caching server is positioned on a fast LAN, near to the data consumers (Figure 3). This could significantly enlarge the usability of OGC SOS service for handling large data sets, as compared with the more common situation where data needs to be transferred over (slow) Internet lines. In addition, the caching server could also improve the robustness and scalability of the data access, e.g. in the case the original server itself or the Internet connection is temporarily inaccessible or overloaded.

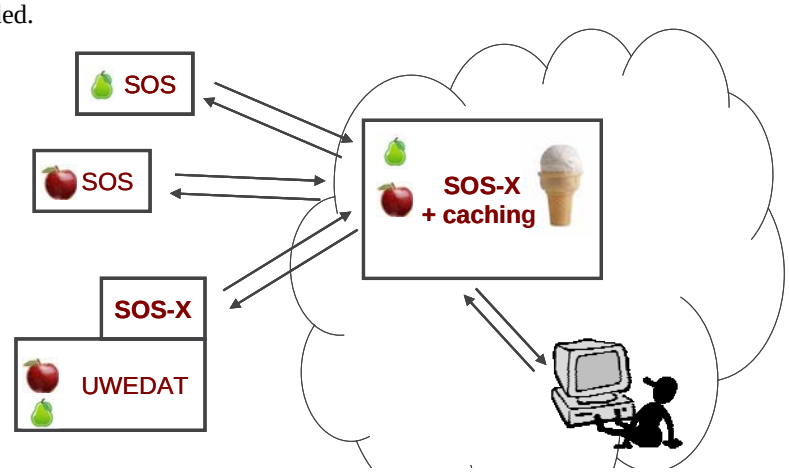


Figure 3: Caching SOS-X as a gateway to UWEDAT system and as SOS proxy

However, caching opens up a number of challenging questions. In fact, the efficient work of the cache (with SOS specifications v1.0.0) can only be assured if two important rules are followed: (1) all data offered by underlying SOS must carry a publication date; and (2) no data is ever deleted. A number of other measures can be taken to further improve the efficiency of the caching, such as (1) adding the information on expected frequency of updates to the offerings; (2) designing the offerings in a way that minimizes the duplication; (3) offering a subscription mechanism to notify when updates are available; or (4) even push complete updates to subscribers.

ACKNOWLEDGEMENTS

The authors wish to thank the Information Society and Media DG of the European Commission (DG-INSFO) for co-founding of the SANY and ORCHESTRA integrated projects within the RTD activities of the Thematic Priority Information Society Technologies

REFERENCES

Anders, F-S (Ed.) (2008): "Translating Feature Access Service (FAS-X)" v0.4, ORCHESTRA –Open Architecture and Spatial Data Infrastructure for Risk Management, Integrated Project, http://www.eu-orchestra.org/publications.shtml#OAspecs

Bartha, M (Ed.) (2006): "D2. 1.2 – Report on Technology review"; SANY IP deliverable: <http://sany-ip.eu/>

Havlik, D., Schimak, G., Denzer, R., and Stevenot B. (2006): "Introduction to SANY (Sensors Anywhere) Integrated Project" in Klaus Tochtermann, Arno Scharl (eds). Proceedings of the 20-th international conference "Informatics for environmental protection", Graz, Austria, September 6-8 2006 ISBN: 3-8322-5321-1

Na, A.; Priest, M (Eds.) (2007): "OpenGIS Sensor Observation Service (v.1.0.0) ; OGC "Implementation Standard" No. 06-009r6; https://portal.opengeospatial.org/files/?artifact_id=26667

Schimak, G. (2003): "Environmental data management and monitoring system UWEDAT" in Environmental Modelling and Software, Volume 18, Number 6, July 2003, pp. 573-580(8)

Schimak, G. (Ed.) 2007: "Rules for ORCHESTRA Application Schemas for Meta-Information", Annex B1 of the Reference Model for the ORCHESTRA Architecture (RM-OA Version 2, Deliverable of the ORCHESTRA Consortium), ORCHESTRA – Open Architecture and Spatial Data Infrastructure for Risk Management, Integrated Project; <http://www.eu-orchestra.org/documents.shtml>

Usländer, Th. (Ed.). (2007): "Reference Model for the ORCHESTRA Architecture v.2.1", OGC "Best Practices" paper No. 07-097, http://www.opengeospatial.org/standards/bp

Usländer, Th. (Ed.). (2008): "Specification of the Sensor Service Architecture v0", SANY – Sensors Anywhere FP6 Integrated Project deliverable D2.3.1, http://sany-ip.eu/publications/1957

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Coupled Modelling of Forest Fires and Floods in Multi-risk Scenarios under SCIER System

T. Metelka[a] and D. Caballero[b]

[a] DHI HIF a.s. Na vrsich 5, 100 00 Prague 10, Czech Republic (t.metelka@dhi.cz)
[b] Tecnoma, Isla del Hierro 7, San Sebastián de los Reyes, 28700 Madrid, Spain, (davidcaballero@tecnoma.es)

Keywords: forest fire; flood; flash-flood; numerical model; multi-risk; scier; sensor network; wildland-urban interface.

Floods and fires represent one of most devastating natural phenomena in the territory of Europe. Moreover, the combination of natural hazards is recorded in many European countries such as flood followed by land slide, fire followed by flood causing large devastation in complex environment of human lives, health, property and nature. The protection against such combination of natural events requires involvement of sophisticated approaches and advanced technology.

The SCIER research project belongs to the 6th Framework Programme of the European Commission, and aims to contribute in coping with environmental risks for prediction of occurrence and impact of these natural hazards with ultimate goal in protection of human health, property and environment. SCIER focuses on design, development and demonstration of an integrated system of sensors, networking and computing infrastructure, aimed for detecting, monitoring, predicting and assisting in crisis management of flood and fire natural hazards or accidents in the wildland-urban interface (W-UI), that is the areas where forests and rural lands interface with homes, other buildings and infrastructures.

SCIER system integrates networks of sensors, which monitorise environmental factors triggering and holding forest fires and floods, communication technologies, computation and processing systems, prediction models for forest fires and floods, GRID architecture for the computing of several possible futures and a GIS platform onto which the main application is installed. Small towns, settlements, housing areas and industrial installations are the targeted protection units around and within which a number of sensors are installed, such as temperature, air humidity and cameras, and which readings are collected and pre-processed by data sinks and sent to an alert-unit which is in charge to launch the information processes and communicate results to the end-users. Four study areas have been set-up in Czech Republic, France, Greece and Spain and a set of field and lab experimentations took place in 2007 and 2008 in Portugal and Greece to test and prove the concepts and the systems proposed in the project.

The purpose of the Sensing Subsystem in SCIER is the monitoring of the environmental parameters (e.g. temperature, humidity, wind-speed, etc.) that are relevant to the assessment of a hazardous situation. The measurements are done by sensor nodes, deployed in the W-UI, and are provided with wireless communication capabilities. They are organized in a wireless sensor network (WSN), which is linked to the communication infrastructure servicing the homes in the W-UI. Emphassis is given to sensor localisation, energy efficiency, incident driven sensor management, dynamic deployment of sensor networks and self-configured sensor network communications, and data collection and pre-processing of the sensor readings.

The SCIER sensing components deployed in the W-UI zones are the following:

- In-field sensor nodes, bearing two or three sensors, and responsible for measuring and transmitting raw parameter value
- In-field sensor nodes that act as cluster heads. They are similar to the sensor nodes except that they are more powerful and can bear additional, more complex sensors
- Out-of-field vision sensors based on an optical camera. They are capable of both reporting raw measurements corresponding to the various parts of the scene they focus on, and of notifying on high-risk levels of the measurements
- The data sinks, that relay the information to and from the sensor nodes towards the SCIER Computing System or the alerting control unit.

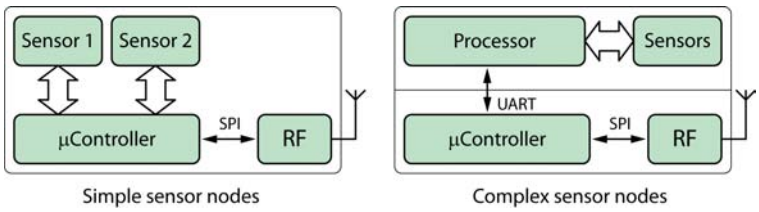


Figure 1. Block diagram of a sensor node

The network is made of nodes and a gateway. The nodes and the gateway use WiseMAC as the communication protocol. The Gateway is interconnected to the systeme that receives the data via a serial interface on which the HCI protocol.

The **L**ocal **A**lerting **C**ontrol **U**nit (LACU) controls a network of variable type sensors, receives input from them, and executes fusion algorithms on the received data. LACUs are positioned all over the area prone to environmental risks, such as forest fires or floods. LACUs by themselves form a network with two distinct hierarchy/domain levels.depending on the authority responsible for deployment and maintenance: LACUs belonging τo the public infrastructure (public LACUs), and private LACUs, bought and installed by property owneres.

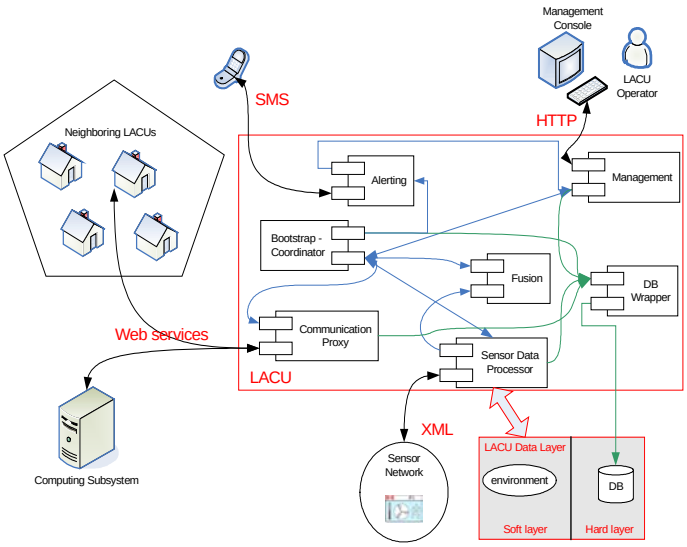


Figure 2. LACU internal software architecture

Each node hosts an application (applicative software), which is the "main" of the software running on the node micro-controller. The complexity of this application is of course limited by memory space and time-related requirements
Three modes of operations or "states" are designed for the LACUs:

1) NORMAL state: The LACU operates under normal conditions. No indication of an emergency situation has been detected.
2) NOTIFIED state: The LACU has already received a notification for a possible emergency situation (input source for this notification may be either the CS or a neighbouring LACU).
3) ALERTED state: An emergency situation has been acknowledged, either from the LACU itself or the CS.

This paper presents mainly the aspect, methods and results on coupling forest fire and flood models under the SCIER system in real multi-risk scenarios, and the use of the input of network of sensors. The basic approach followed is based on common understanding of soil characteristics and their change under circumstances of fire.

The use of CN curves together with land recovery curves constitutes the fundamental methodology used to couple forest fire and flood models. The hydrological response of the W-UI catchment to heavy rainfall differs depending on the different soil conditions, which are represented by CN curve numbers. Several if/then scenarios are presented and discussed in the paper showing the impact of burned area on the way floods are forming and routing in a particular river catchment. Sensitive areas are then detected by coupling fire/flood modelling, out of which particular restoration measures are proposed.

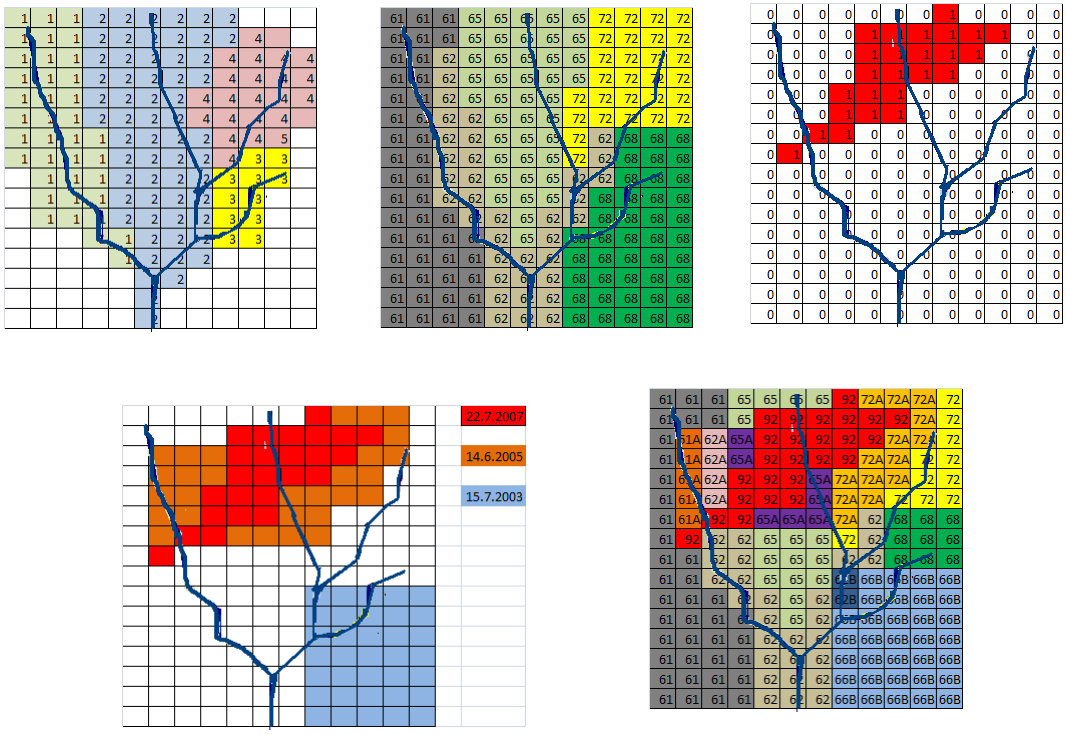


Figure 3. Process of coupling forest fire effect in the soil and hydrological models through the CN curve numbers

This multi-risk approach has already been followed in ORCHESTRA project [Caballero, 2007] in a similar area, also in Catalonia and also tested in MEDIGRID project in Extremadura region, in Spain [Bovolo, 2007] In ORCHESTRA pilot study case, the main

goal was to improve the inter-operability among the involved information and simulation systems through the implementation of OA services. This experience has helped in the development of the pilot study presented here, although in SCIER the focus has been given to the integration of sensed data for the monitoring of fire and floods conditions and the simulation of possible futures in a GRID infrastructure.

ACKNOWLEDGEMENTS

The presented work has been developed under the auspices of the SCIER research project, partially funded by the European Commission under the 6th Framework Programme, and the authors present the paper in behalf of all the members of SCIER consortium.

REFERENCES AND BIBLIOGRAPHY

Bovolo, C.I., Abele S.J; Bathurst, J.C., Caballero, D., Ciglan, M., Eftichidis, G., Simo, B. (2007) *A Distributed Framework for Multi-risk Assessment of Natural Hazards Used to Model the Effects of Forest Fire on Hydrology and Sediment Yield*. Ed. Computers & Geosciences

Caballero, D., Esteban, J., Izquierdo, B. (2007) *ORCHESTRA: a unified and open architecture for risk management applications*. European Geosciences Union General Assembly 2007 Vienna, Austria, 15 – 20 April. In Geophysical Research Abstracts, Vol. 9, 08557, 2007 SRef-ID: 1607-7962/gra/EGU2007-A-08557

Chen, T., Wu, P. And Chiou Y, *An early fire detection method based on image processing*, In Proc. IEEE Intern. Conf. on Image Processing, ICIP '04, pp. 1707-1710.

Fastcom Technology Sa, (2003). *Boulevard de Grancy 19A, CH-1006 Lausanne, Switzerland, Method and Device for Detecting Fires Based on Image Analysis*. Patent Cooperation Treaty Application No. PCT/CH02/00118, PCT Publication No. WO02/069292.

Healy, G., Slater D., Lin T., Drda B., Goedeke, A.D. (1993) *A system for real time fire detection*, In Proc. IEEE Computer Vision and Pattern Recognition Conference (CVPR 93), pp. 605-606.

Liu C.B., Ahuja N. (2004) *Vision based fire detection*, In. Proc. Intern. Conf. on Pattern Recognition (ICPR, 2004) vol. 4, pp. 134-137.

Philips W., Shah M., Lobo N. (2002) *Flame recognition in video*, Pattern Recognition Lett. 23 (1-3), 319-327

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Towards Process Oriented Environmental Monitoring Systems for Intelligent Risk Management

Oliver Faust [a], **Bernhard H.C. Sputh**[a] **and Alastair R. Allen**[a]

[a]*School of Engineering, University of Aberdeen, Aberdeen AB24 3UE, UK*
{o.faust, b.sputh, a.allen}@abdn.ac.uk

Abstract: Water monitoring and associated risk management requires a large number of in-situ monitoring stations. These stations measure water quality parameters and relay the measurements to a data centre. From a data processing point of view, such systems constitute distributed processing networks. In present water monitoring systems and in a wide range of other environmental monitoring systems the network aspect stops at the node or in-situ monitoring station (IMS) level. This paper extends the networking idea from the system level to the IMS design level. We model an in-situ monitoring station as a network of communicating sequential processes. For this network we establish the fundamental system properties such as: security, stability and functionality. In a second step we translate this functional model into an implementation model with the same properties. The designer has the freedom to translate particular network processes into either hardware or software components. These components communicate over well defined interfaces, this ensures composability. In the practical part we follow the two step design approach to integrate a new sensor into an existing IMS. First we create a process network with the desired properties. Then this process network is translated into hardware and software components for a hybrid processor. Finally, the functionality of the resulting component network is tested.

The process oriented design approach leads to more secure, reliable, and functional systems. Furthermore, this design method simplifies the integration of In-situ Monitoring Stations into the water risk management system, because the complete system functionality is designed according to networking principles.

Keywords: Remote sensing; Water risk management; Process oriented system design; Process networks; Decision support systems; Hardware Software co-Design; In-situ sensors; Formal models; Communicating Sequential Processes

1 INTRODUCTION

Clean drinking water is one of the most important, if not the most important food for humans and animals alike. Furthermore, according to Gunatilaka [2006], it is in constant danger of being polluted by environmental threats. This is the reason why 9 institutions from 6 European countries formed a consortium to carry out the WARMER (WAter Risk Management in EuRope) project. The driving idea behind WARMER is to provide an intelligent water risk management system. Such intelligent systems have been identified by Sànchez-Marrè et al. [2006] as a future direction for environmental monitoring systems. They argue that only decision making systems can tackle the complex problems posed by environmental pollution and habitat destruction. According to Haykin [1994] all intelligent systems are feedback structures which are conveniently described as networks. Therefore, the network aspect plays an important role in the design of intelligent decision making systems. For example, Smiatek [2006] proposes environmental modelling in an

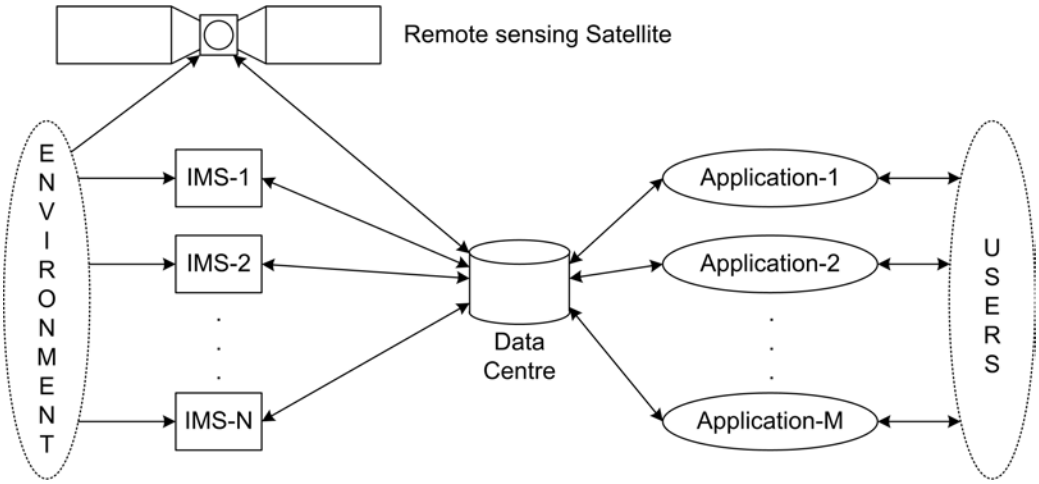


Figure 1: Overall System proposed by the WARMER project

event-driven multitasking network environment. He argues that such a network is the solution to the problem of linking data which is stored in different locations. This particular network based solution provides the means to create and deploy multitasking applications. The WARMER system can be described as a network of independent in-situ monitoring stations (IMS) and remote sensing satellites, which measure environmental parameters. The acquired data is sent to a data centre where it is processed and stored. Furthermore, the data centre grants independent user applications access to the data. Figure 1 illustrates the overall system setup.

From a data processing point of view the WARMER system and indeed most remote sensing systems are distributed processing networks. Remote sensing satellite, IMS, data centre, and user application constitute independent processing entities. These entities perform local data processing and exchange data via some sort of communication network. Unfortunately, in current water monitoring systems, and in a wide range of other environmental monitoring systems, the networking aspect stops at this abstract system level. That means, the underlying structure of individual entities is not designed and described in terms of networked components. This inconsistency breaks the system design flow. Such breaks are undesirable because they make it hard to understand the underlying principles. These underlying principles are very important for the design of intelligent risk management systems. For example, there is the network specialist who understands the overall system structure, but who does not understand the underlying systems. Similarly, there is the device specialist who understands, say, the application, but has only a vague understanding of the overall system. This lack of understanding leads to the fact that neither of these two specialists can introduce a great deal of intelligence in their respective system design. Apart from this philosophical or knowledge sharing problem, the break in the design flow leads, in many cases, to underperforming systems. Such systems tend to underutilise the available resources of distributed processing systems. At a first glace, there are two important properties of distributed processing systems which are especially beneficial for environmental monitoring. The first property is parallel processing. Executing more than one operation at the same time speeds up the processing. Another, largely underestimated, property is the fact that a specific task can be executed anywhere in the network. This leads to the following corollary property: In a distributed processing system the functionality is mapped onto the available processing resources. Clearly, it is impossible to make use of this property in an environmental monitoring system, if the individual components are not designed according to network principles.

In this brief we propose to extend the networking idea from the system to the node design level. This will increase flexibility and performance of the resulting environmental monitoring systems. Furthermore, this overarching system concept makes it possible to reason about security, stability and functionality of the complete system. These properties are the fundamental requirements for intelligent decision making systems, because such systems are designed as networked components

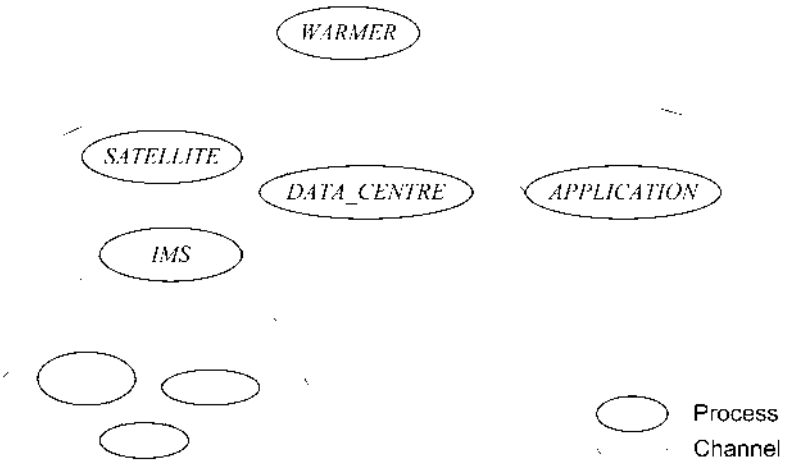


Figure 2: *WARMER* process network

which work cooperatively on the decision making task. To achieve the goal of extending the network aspect from the system to the design level we adopt the theory of communicating sequential processes (CSP).

2 PROCESS ORIENTED SOLUTION

According to the CSP theory each process is either atomic or composed of a network of processes. Therefore, in a CSP model the meaning of process of depends on the way a particular argument is constructed. For example, the system shown in Figure 1 can be modelled as a *WARMER* process. This process is composed of a process network, as shown in Figure 2. One part of this network is the *IMS* process. In this argument the *WARMER* process represents the system and the *IMS* process is a system component. Now, following our idea of extending the network aspect towards the IMS design, we model the *IMS* process itself as a process network. The processes which make up this process network gather information, manipulate and distribute data. In this argument the *IMS* process represents a system which is composed of multiple components. The process concept constitutes a recurring pattern throughout the complete design. This is an elegant solution to the problem of the broken design flow.

REFERENCES

Gunatilaka, A. Can EU directives show Asia the Way. *Asian Water*, pages 14–17, December 2006.

Haykin, S. *Neural networks: a comprehensive foundation*. Prentice Hall, 1994.

Sànchez-Marrè, M., K. Gibert, R. S. Sojda, J. P. Steyerc, P. Struss, and I. Rodrìguez-Roda. Uncertainty Management, Spatial and Temporal Reasoning, and Validation of Intelligent Environmental Decision Support Systems. In *Proceedings of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software"*. International, 2006.

Smiatek, G. Environmental modeling in an event-driven multitasking network environment. In *Proceedings of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software"*. International, 2006.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Evaluating the information content of data for uncertainty reduction in hydrological modelling

B. Croke[a], T. Wagener[b], D. Post[c], J. Freer[d], I. Littlewood[e]

[a] *Department of Mathematics, and The Fenner School of Environment and Society, The Australian National University, Canberra ACT 0200, Australia*

[b] *Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA16802, USA*

[c] *CSIRO Land and Water, Canberra, Australia*

[d] *University of Lancaster, Lancaster, UK*

[e] *Didcot, Oxfordshire, UK*

Abstract: The inclusion of additional information in a model should improve the model's performance and reduce associated uncertainties. The additional information may take the form of higher temporal and/or spatial resolution of data already used, or additional types of data (including soft data). However, additional data is not necessarily equivalent with additional information as a model may only be able to make use of a fraction of the information in the data, or the approach chosen to extract data might be inefficient. Consideration of the information-to-noise ratio of the data is needed in order to evaluate whether the information will indeed improve the model's performance or reduce uncertainty. Evaluating the information-to-noise ratio and the information content of data in general are often difficult tasks. One simple but effective approach to this problem is to compare the performance and uncertainty of a model with and without additional information. This workshop will focus on the theory and use of additional information in improving the performance of and reducing the uncertainty in hydrological models. This includes consideration of the uncertainty in datasets as well as techniques to evaluate the effectiveness of the additional information in improving models.

Keywords: Rainfall-runoff models; uncertainty; performance metrics; information.

1. INTRODUCTION

When mathematical models are applied to estimate hydrological variables, e.g. streamflow from rainfall and other hydroclimatalogical data, it is necessary to consider the impact of uncertainties in the input data, in model structures, and in the values of a model's parameters. The Prediction in Ungauged Basins (PUB) Decade (2003-2012) of the International Association of Hydrological Sciences (IAHS) aims to reduce predictive uncertainty in model outputs, particularly when dealing with ungauged catchments [Sivapalan et al., 2003; Wagener et al., 2004]. Several PUB Working Groups (WGs), each with different but sometimes overlapping objectives, are working towards making progress with the aims of PUB. The Top-Down modelling WG (TDWG, WG1) and the Uncertainty Estimation for Hydrological Modelling WG (WG7) have convened this workshop to review progress to date and discuss possible future PUB research. In order to set manageable boundaries for the workshop the paper introduces only some aspects of uncertainty and its reduction. A thorough review of the topic has not been attempted here. Nevertheless, by outlining some recent and ongoing work related to uncertainty in hydrological modelling, with appropriate citations, the paper sets the scene for the workshop.

A focus of the workshop is to address the information content of data, particularly additional data, in terms of whether progress towards reducing estimation uncertainty can be, or is likely to be, made. The purpose of including additional information in model inputs is to improve the performance of the model. The additional data may take a number of forms, including higher temporal and spatial resolution of existing data being used (e.g. spatial distribution and intensity of rainfall), or adding data that were not previously included [e.g. Seibert and McDonnell, 2002; Freer *et al.* 2004; Son and Sivapalan, 2007]. Note that additional information may also be in the form of constraints. For example, bounds on the steady-state gain of an autoregressive-moving average exogenous (ARMAX) model yield linear inequality constraints on the ARMAX parameters which can be used to reduce uncertainty in their estimates [Norton, 1979]. The inclusion of additional information will improve the model's performance only as long as the information content in the data (or more accurately, the information that the model is able to use) is greater than the noise that is introduced. Sorooshian *et al.* [1983] suggested that it is the quality of information contained with a dataset, rather than the quantity (i.e. length of record), which may be the more important characteristic for a data set as beyond a certain point (this will depend on the nature of the system, and the data as well as the uncertainty in the data), the additional data does not significantly improve the parameter estimates. Unfortunately, datasets often have significant uncertainty, resulting in occasions when more noise than information is added by the inclusion of the data. Therefore, the information-to-noise ratio (or more broadly the net information gain) of any datasets needs to be considered before inclusion of the data within a model.

There have been a number of studies investigating ways of addressing the question of the impact model errors have on model performance and calibration. Yapo *et al*., [1996] conclude that "the factor currently limiting model performance is the unavailability of strategies that explicitly account for model error during calibration". The BATEA calibration method [Kavetski *et al*. 2002] is an example of allowing for model error, though there can be confusion between model error and error in the model inputs [Thyer *et al.* 2007]. Marshall *et al.* [2006] proposed the use of the hierarchical mixtures of experts (HME) framework for reducing the predictive uncertainty due to errors in the model structure by combining the predictions of two or more models. In calibration of an in-river phosphorus model, McIntyre and Wheater [2004] found that error in the pollutant load data, model structure as well as the inherent model parameter equifinality compromised the value of calibration data, concluding that expectations of model performance should be reviewed in light of the resource limitations, and that there is a need for more robust estimation of model uncertainty. In a study of the application of two rainfall-runoff models (GR4J and TOPMO) on twelve catchments in the USA (basin sizes ranged from 1021 to 4421 km^2, with the ratio of mean annual rainfall rainfall to potential evaporation ranging from 0.5 to 1.7), Perrin *et al*. [2007] found that if non-continuous calibration periods were used, in general approximately 350 calibration days selected to cover a range of conditions were sufficient to obtain robust parameter values, though for drier catchments, it was more difficult to reach stable parameter values.

The structure of this workshop position paper is based on the following sections:

- Review of recent uncertainty research
- Estimation and propagation of uncertainty
- How to estimate the reduction in uncertainty
- How to maximize information extraction from available data
- Example studies
- Discussion and conclusions

2. REVIEW OF RECENT UNCERTAINTY RESEARCH

There has been a considerable volume of work studying uncertainty and the impact on model outputs. Oudin *et al.* [2006] studied the impact of corrupted (both random and systematic) inputs (precipitation - P and potential evapotranspiration - PET) on model performance, concluding that random errors in P have potentially the largest effect on

model performance. Systematic errors in P may be accounted for in the model calibration if the model structure permits this. However, from the viewpoint of predicting flows at ungauged locations, systematic errors in rainfall are more problematic unless the error does not change significantly between the gauged and ungauged sites. The influence of temporal resolution of rainfall data on generation of infiltration excess runoff has been explored by Bronstert and Bárdossy [2003], with model outputs being most sensitive to resolution for events with rainfall intensity near the infiltration capacity.

It is important to remember that estimation of uncertainty in observed and modelled quantities may not be an end in itself. When uncertainties are used to draw conclusions, the confidence in the uncertainty estimates needs to be assessed in order to be able to define the confidence in the conclusions. In the view of Montanari [2007], this is possible when sufficient information is available. Generally, this requires consideration (e.g order of magnitude estimation) of all potential sources of uncertainty, though only significant contributors to the uncertainty need to be estimated with greater care. In hydrology, some of the sources of uncertainty that need to be considered are:

- observation errors;
- rating curves, and when these are being extrapolated;
- spatial interpolation and aggregation, including the spatial resolution of the available data, and the spatial scales of variations in the fields being interpolated.
- influence of temporal resolution (data time-step)

Only by careful consideration of the sources of uncertainty can valid conclusions be drawn regarding the utility of the data, and the impact the data have on modelled outputs and subsequent conclusions.

2.1 Uncertainty due to modelling data time-step

Recent work [Littlewood, 2007; Littlewood and Croke, submitted; Littlewood, in prep.a,b] has focussed on a discrete-time modelling artefact whereby IHACRES (and probably many other similar models') parameter values are dependent upon the data time-step used for calibrating the model. As more information was added to an IHACRES model for one of the Plynlimon research basins in the UK, by reducing the data time-step, selected model-fit statistics and precision (and accuracy) on the model parameters improved. This is not only because a smaller time step implies more information per unit time; use of a smaller time step also reduces systematic distortion of the estimated unit hydrograph by (i) averaging over a time step and (ii) aliasing, if instantaneous flow is sampled. It might be expected, therefore, that a model calibrated using a small data time-step (e.g. hourly), when applied in simulation-mode followed by aggregation of the results to a larger data time-step (e.g. daily), will give better estimates of streamflow at that larger time-step. However, care is needed since the distortions due to the sampling depend on how the sampling is done (e.g. mean values over a sampling interval, or instantaneous value for each sampling interval).

As modelling data time-step decreased from daily to hourly for the Plynlimon catchment in question, an IHACRES-derived slow flow index ($0\% < SFI < 100\%$) increased by about 20 percentage points. The shape of the curve described by the plotted changes invited extrapolation to a data time-step of effectively zero, giving an SFI* independent of data time-step. Corresponding values of an analogous baseflow index, which has been widely applied in regionalisation studies, increased by about 40 percentage points. Furthermore, the results exhibited no way of deriving a data time-step-independent value of this baseflow index. Slow flow and baseflow indices estimated from discrete-time data are unlikely to be consistently characteristic between catchments. Therefore, regionalisation studies that have assumed that SFI, the baseflow index, and rainfall–streamflow model parameters generally are robust when derived using discrete-time data for all gauged basins employed, may not have generated the best possible regionalisation equations. Better regionalisation equations, e.g. using SFI* and other data time-step-independent model parameters would lead to reduced uncertainty in hydrographs estimated at ungauged (flow) sites.

Further work is required to investigate this important aspect of regionalisation and catchment-scale modelling using parsimoniously parameterised rainfall–streamflow models. The TDWG TRUMPER (Towards Reducing Uncertainty in rainfall-streamflow Model Parameter Regionalisation) initiative has been established for that purpose. Amongst other things, TRUMPER will provide an opportunity to compare the results from Plynlimon outlined above with corresponding results from application of a data-based mechanistic (DBM) approach [Young and Romanowicz, 2004; Young and Garnier, 2006] whereby continuous-time parameter values, which should be independent of data time-step, are estimated using discrete-time data.

3. ESTIMATION AND PROPAGATION OF UNCERTAINTY

Data contain both information that will improve a model's performance and error that will degrade the model's performance. Consequently, evaluation of the impact a dataset has on the modelled output requires propagating the uncertainty through the model, as well as the net effect of information able to be used by the model. Andréassian *et al.* [2007] found that for a catchment in France, there was no difference in model performance between a lumped approach using areal rainfall estimate for the entire catchment, or a semi-distibuted approach where separate areal rainfall estimates for individual subcatchments were used.

While Monte Carlo techniques can easily be used to estimate the uncertainty in model outputs resulting from uncertainty in model inputs (both data and parameters) so long as the input uncertainties can be well described, the resulting computational load can be prohibitive. It is possible to use analytical techniques to propagate uncertainties through a model, though this is straightforward only if the model is linear and the errors additive and Gaussian. For analytical assessment of uncertainty in models that are sufficiently linear, only a measure of the uncertainty is needed (e.g. standard deviation, or 95% confidence limits). Note that the Bayesian Monte Carlo (BMC) approach also needs to have the distribution of the uncertainties defined; a weakness of the BMC method compared to the Markov Chain Monte Carlo (MCMC) method. BMC is generally inefficient compared with MCMC through poor sampling distribution [Qian *et al*., 2003]. For likelihood-based methods [e.g. GLUE -Beven and Binley, 1992], a non-informative prior density is adopted (usually uniform), with the assumption that the information extracted from the posterior distribution is not significantly influenced by the choice of prior. However, the information introduced by a flat prior depends on how the model is parameterized; for example, a flat prior for parameter a in an exponential $exp(-at)$ is not the same as a flat prior for an equivalent parameter, the time constant $\tau = 1/a$ in $exp(-t/\tau)$. Shrestha *et al.* [2007] used fuzzy set methods to study the impact of uncertainty in the stage-discharge relationship on a one-dimensional hydrodynamic model, giving a method of assessing uncertainty in a non-probabilistic but subjective framework.

In some cases that are important within a hydrological context, computationally intensive Monte-carlo-based approaches to uncertainty analysis are not necessary since analytical solutions exist. The most important example in this regard the case of linear or near linear system models, as used in flow routing, where the Kalman filter [Kalman, 1960] provides for uncertainty propagation associated with the state and output variables, while recursive parameter estimation for transfer function models [e.g. Young, 1984, 2002, 2005] allows for uncertainty propagation in relation to the model parameters. Also, an important nonlinear example, which reduces the need for Monte Carlo-based analysis, is the 'Hammerstein' model, where the model consists of an input nonlinearity in series with a linear system. Such models are appropriate for characterizing rainfall-flow processes (e.g. IHACRES, HYMOD, DBM models), where the nonlinearity is associated with the generation of 'effective' rainfall, and these have been exploited [e.g. Young 2002, Romanowicz *et al*, 2006] within a flood forecasting context where uncertainty analysis is, of course, an important prerequisite. In this situation, state and parameter estimation for the linear component, as well as the associated uncertainty propagation, can be accomplished using slightly modified Kalman filter algorithms. The only requirement for numerical methods is for uncertainty propagation in relation to parameters of the serial nonlinearity.

In the case when the model output is a simple function of one input:

$$y = f(x) \tag{1}$$

Croke [2007] showed that, through Taylor series expansion of *f*, the standard deviation of *y* can be obtained analytically providing sufficient information is known about the distribution of the uncertainty in *x*. If the function *f* is sufficiently close to linear over the range of uncertainty in *x*, only the first term in the Taylor series expansion is needed:

$$\sigma_y^2 \approx \left(\left. \frac{df}{dx} \right|_{x=\bar{x}} \right)^2 \sigma_x^2 \tag{2}$$

However, if the function *f* is sufficiently non-linear (e.g. the function $y=x^2$ near $x=0$) , then the second derivative at least will be needed:

$$\sigma_y^2 \approx \left(\left. \frac{df}{dx} \right|_{\bar{x}} \right)^2 \sigma_x^2 + \left. \frac{df}{dx} \right|_{\bar{x}} \left. \frac{d^2 f}{dx^2} \right|_{\bar{x}} {}^3M_x - \frac{1}{4} \left(\left. \frac{d^2 f}{dx^2} \right|_{\bar{x}} \right)^2 \left\{ \sigma_x^4 - {}^4M_x \right\} \tag{3}$$

$$ {}^nM_x = \sum_{i=1}^{N} \left\{ \frac{(x_i - \bar{x})^n}{N-1} \right\}$$

This result can be generalised to consider the distribution of uncertainty. Taking

$$x = \bar{x} + \varepsilon_x, \qquad y = \bar{y} + \varepsilon_y \tag{4}$$

the deviation ε_y of *y* from its mean is approximated by:

$$\begin{aligned} \varepsilon_y = y - \bar{y} &= f(\bar{x}) + \sum_{i=1}^{N} \frac{\varepsilon_x^{\,i}}{i!} \left. \frac{d^i f}{dx^i} \right|_{x=\bar{x}} - f(\bar{x}) - \sum_{i=2}^{N} \frac{\text{mean}(\varepsilon_x^i)}{i!} \left. \frac{d^i f}{dx^i} \right|_{x=\bar{x}} \\ &= \sum_{i=1}^{N} \frac{\varepsilon_x^{\,i}}{i!} \left. \frac{d^i f}{dx^i} \right|_{x=\bar{x}} - \sum_{i=2}^{N} \frac{\text{mean}(\varepsilon_x^i)}{i!} \left. \frac{d^i f}{dx^i} \right|_{x=\bar{x}} \end{aligned} \tag{5}$$

for some finite *N*. For highly non-linear models, additional information on the distribution of the uncertainty is needed (e.g uniform, or normally distributed). Croke [2007] showed that if the second-order term in the Taylor series expansion of the model is needed, then the 3rd and 4th moment about the mean of the distribution is needed in order to calculate the standard deviation of the model output, therefore requiring more information regarding the distribution of the uncertainty in the input.

4. HOW TO ESTIMATE THE REDUCTION IN UNCERTAINTY?

The relative uncertainty before and after new data have been assimilated into the model provides a quantification of the information content in the new data that the model is able to utilise. This uncertainty reduction can be measured using multiple approaches, partially depending on the approach to uncertainty estimation chosen, namely statistical versus set-theoretic approach. Another question is: Which uncertainty is reduced? This is mainly the uncertainty in the model predictions, but it could also be the uncertainty in the model parameters or in some other item which the model is required to reproduce well, e.g. flow-duration curve for a runoff model.

A simple approach that works if the distributions of parameters or individual output variables are not too complex is to compare the prior (before new data are assimilated) and posterior (after new data are assimilated) standard deviations or variances of individual distributions. The ratio of before to after standard deviations provides a direct quantification of learning [Vrugt *et al.*, 2005].

An alternative approach can be used to assess reduction in parameter uncertainty if a set-theoretic approach to parameter estimation has been utilized [Freer *et al.* 2004]. The result of a set-theoretic approach is a population of behavioural parameter values, i.e. values

causing the model to meet one or more requirements.. Adding new data should lead to a reduction in the size of the behavioural parameter population. Wagener *et al*. [2002; 2003] quantified this reduction by comparing the 90% percent confidence limits of before and after behavioural distributions of individual parameters.

If the reduction in the uncertainty in ensemble predictions of some output variable time series is of interest, then the improvement in sharpness can be calculated as suggested by Yadav et al. [2007]. The unconditioned parameter ranges for a model can be used to produce prior ensemble predictions. For each time step these predictions represent 0% sharpness. Assimilating data into the model, which reduces the parameter ranges and therefore usually the subsequent ensemble prediction uncertainty should lead to an increase in sharpness. This increase of sharpness is related to a reduction of the ensemble uncertainty range and 100% sharpness would mean the ensemble prediction has been reduced to a single trace.

The approaches mentioned in this section enable a quantification of the uncertainty reduction in a relative sense, before versus after assimilation of new data. How valuable new data are for this purpose largely depends on how successful a chosen approach is in extracting this information. The next section will discuss the role of objective functions or performance metrics or likelihood functions in this context.

5. HOW TO MAXIMIZE INFORMATION EXTRACTION FROM DATA?

5.1 Performance metrics

A number of metrics are used to evaluate the performance of models, with many based on the sum of squared errors (SSE). In hydrology, a commonly used objective function is the Nash-Sutcliffe efficiency [NSE, Nash and Sutcliffe, 1970]. Garrick *et al*. [1978] suggested two modifications to the NSE:

- replacing the mean flow as the "no model" forecast with the mean value at the current date (and time) for each sampling interval calculated from the period being considered (averaging could be at a different resolution than that of the modelled and observed values; for example, using a seasonal average flow). Note that this requires a sufficiently long period to permit calculation of the mean, which will potentially be a problem for selection of a calibration period). Ideally, the current timestep should not be used in calculating the mean.
- including an updating procedure for real-time forecasting applications

Such objective functions do not routinely account for the uncertainty in the observed and modelled values being compared. Rather, they assume particular behaviour of the errors (most assume homoscedasticity) and as a result, their validity as performance indicators depends on the validity of this assumption. In the case of hydrological models, the assumption of homoscedasticity is usually not valid. A simple approach for correcting for the heteroscedasticity in the model residuals is through the use of the Box-Cox transformation [Box and Cox, 1964]:

$$\tau(Y;\lambda,\alpha) = \begin{cases} \left((Y+\alpha)^{\lambda} - 1\right) & if\ \lambda \neq 0 \\ \ln(Y+\alpha) & if\ \lambda = 0 \end{cases} \tag{6}$$

where α is a shift parameter (often set to zero). While this approach can significantly reduce the heteroscedasticity, the uncertainty is assumed to depend only on the current value of Y. If there is significant correlation between the uncertainties and other quatities (e.g. when there is hysteresis) the Box-Cox transformation may not be adequate. The result obtained from applying the Box-Cox transformation is also model dependent as the influence of model structure uncertainty is also removed. Consequently, different models will employ different values of λ, and comparing the performance of models will be difficult.

Another approach involves weighting the model residuals by their uncertainty [Croke, 2007]. This ensures that the variance in the weighted residuals is constant, as long as the estimates of the variability in the uncertainty are sufficiently accurate. This has similarity to optimal (inverse-covariance) weighting in minimum-covariance, linear, unbiased estimation (generalized least-squares regression) [e.g. Aitken, 1935 and Plackett, 1949].

5.1.1 Modified objective functions

Croke [2007] showed that objective functions can be modified using an optimal weighted average approach to explicitly include the influence of uncertainties in the modelled and measured quantities. Here, the following definitions will be used:

- observed value at timestep *i*: $x_{o,i}$
- modelled value at timestep *i*: $x_{m,i}$
- model residual at timestep *i*: $e_i = x_{o,i} - x_{m,i}$
- observed deviation from mean: $d_{o,i} = x_{o,i} - \bar{x}_o$

For example, the Nash-Sutcliffe model efficiency R^2 [Nash and Sutcliffe, 1970] is:

$$R^2 = 1 - \frac{\sum_i^n e_i^2}{\sum_i^n d_{o,i}^2} \tag{7}$$

The modified Nash-Sutcliffe efficiency is given by:

$$R_m^2 = 1 - \left[\frac{\sum_i^n w_{e,i} e_i^2 \Big/ \sum_{i=1}^n w_{e,i}}{\sum_i^n w_{d,i} d_{o,i}^2 \Big/ \sum_{i=1}^n w_{d,i}} \right] \tag{8}$$

where the weights $w_{d,I}$ and $w_{e,I}$ are:

$$w_{d,i} = \frac{1}{\lambda_{o,i} \gamma_{f_{o,i}^2}}, \qquad w_{e,i} = \frac{1}{\alpha_i \gamma_{e_i^2}} \tag{9}$$

and α_i and $\lambda_{o,I}$ are:

$$\alpha_i = (\Delta e_i)^2 = (\Delta x_{o,i})^2 + (\Delta x_{m,i})^2 \tag{10}$$

$$\lambda_{o,i} = (\Delta d_{o,i})^2 = (\Delta x_{o,i})^2 + (\Delta \bar{x}_o)^2 \tag{11}$$

where Δ represents the uncertainty in each quantity, and γ is the correction for non-linearity. For example, for a second order correction, γ for estimation of the standard deviation (based on equation 3) is:

$$\gamma_{x_i^2} = 1 + \frac{\left.\frac{df}{dx}\right|_{\bar{x}} \left.\frac{d^2 f}{dx^2}\right|_{\bar{x}} {}^3M_x + \frac{1}{4}\left(\left.\frac{d^2 f}{dx^2}\right|_{\bar{x}}\right)^2 \left\{\sigma_x^4 - {}^4M_x\right\}}{2x\Delta x} \tag{12}$$

5.1.2 Example

The modified NSE described above was tested using a Monte Carlo approach [Croke, 2007], employing synthetic data generated by the IHACRES rainfall-runoff modelling software [Croke *et al.* 2006]. Random errors in the "observed" flow were added by introducing random errors in the parameters of the rating curve, thereby introducing highly heteroscedastic errors with significant serial correlation. Only error in observed flow was considered; propagation of uncertainty in the rainfall through the model (e.g. uncertainty framework of Huard and Mailhot, [2006]) is still under development (Mungkasi *et al.* in prep). For this test, the parameters of the linear module (unit hydrograph) were fixed, with only 2 parameters in the non-linear module (τ_w and c) optimised.

Figures 1 and 2 show the resulting optimised parameter values for 100 trials, and clearly demonstrate that, given a good estimate of the uncertainty in each timestep, the modified NSE gives a much smaller uncertainty in the calibrated parameter value. This shows the importance of a good estimate in the uncertainty in the model inputs.

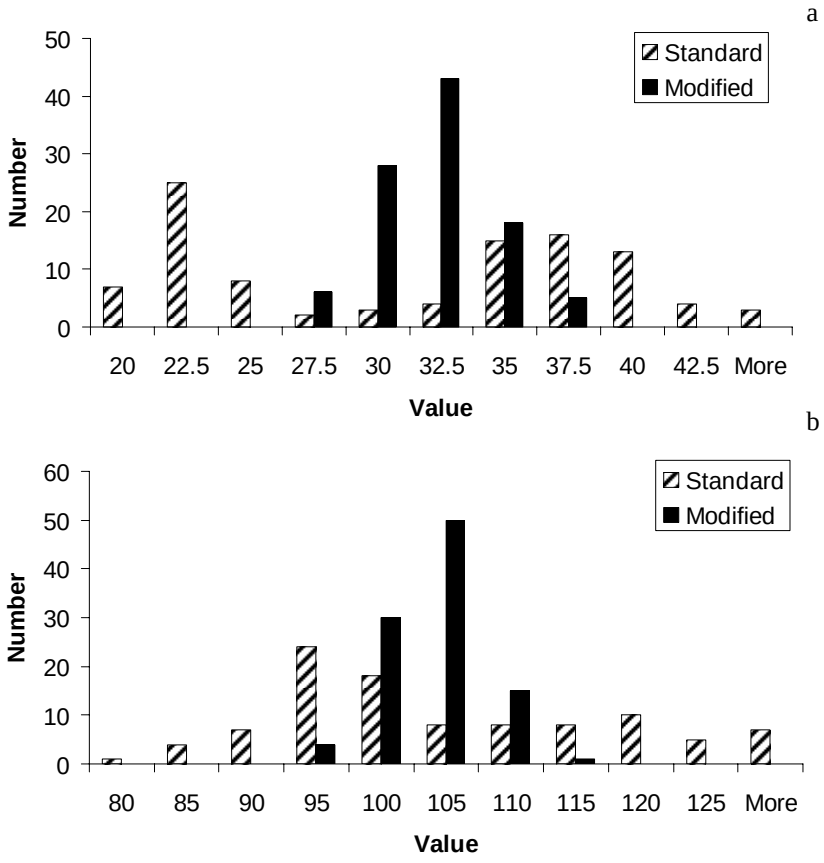


Figure 1. Calibrated parameter values for (a) the catchment drying rate parameter (τ_w, nominal value 30) and (b) the catchment storage parameter (c, nominal value 100).

Weighted ordinary least squares (OLS) estimation techniques have a long history [e.g Aitken, 1935, Plackett, 1949, 1950]. Williams and Yeh showed that generalised least squares (GLS) is a better technique for estimating parameter values in rainfall-runoff models than ordinary least squares (OLS), with GLS permitting weighting of inputs to the model (though they note that the effects of noise can "render the procedure inaccurate or inoperative"). There is also a long history of dealing with noise in recursive GLS or the equivalent; for example by adjoining a noise structure model [Plackett, 1950; Kalman, 1960].

There are a number of approaches that have been developed to address the influence of input uncertainty on model calibration [e.g. Beven and Binley, 1992; Kavetski *et al.* 2002; Huard and Mailhot, 2006]. However, to date the assessment of model performance has ignored uncertainties in the observed and modelled values.

5.1.3 Inclusion of Significance of Model output in Performance Metric

An extension to the modified objective functions can be made by considering the significance of different parts of the model output, given the question the model is attempting to answer. This can be done either with a binary switch (1 for time steps to be considered, 0 for time steps to be ignored) or a weight set for each time step. This is a minor correction to the modified NSE given above:

$$R_m^2 = 1 - \left[\frac{\sum_i^n s_i w_{e,i} e_i^2}{\sum_i^n s_i w_{d,i} d_{o,i}^2} \right] \left[\frac{\sum_{i=1}^n s_i w_{d,i}}{\sum_{i=1}^n s_i w_{e,i}} \right] \tag{13}$$

where s_i is the significance of each time step. The significance of each time step can be selected based on any desired criteria (e.g. discharge rate, slope of hydrograph etc). This gives the modeller the ability to focus the calibration and testing of the model to specific parts of the hydrograph. The result is that the original Nash-Sutcliffe efficiency assumes that the significance is inversely proportional to the weighting applied. As the weights are themselves inversely proportional to the square of the uncertainty, the conclusion is that the original Nash-Sutcliffe efficiency adopts a significance for each point that is proportional to the square of the uncertainty in the corresponding values for that point. Further, the weights for the numerator and denominator are the same, implying negligible uncertainty in the modelled flow (and mean of the observed flow).

This approach allows for considerable flexibility in the objective function at the cost of difficulty with comparing results for different studies. As a guideline, it is recommended that the objective function be evaluated using a uniform significance. While this will permit direct comparison of model performance, care should be exercised in this comparison as model parameters will not necessarily have been optimised for the case of uniform significance.

5.2 Model structure and complexity

Dividing the residuals by their estimated uncertainty will enable assessment of the suitability of the model structure and parameterisation providing that the uncertainties are sufficiently well known. There are two possible tests that can be performed: a check for heteroscedasticity in the scaled residuals (testing the model structure) and assessment of the magnitude of the scatter in the scaled residuals (testing model complexity).

If the scaled residuals (each residual divided by its estimated uncertainty) show significant heteroscedasticity, this indicates either that the model is introducing heteroscedasticity (e.g. by not adequately representing the impact of rainfall intensity, or antecedent conditions), or that there is a systematic error in the estimated relative uncertainties. As with the modification to objective functions discussed in section 4.1, this test only requires sufficiently accurate estimation of the relative uncertainty (i.e. the variations in the uncertainty) in the residual at each time step. Statistical approaches such as the Box-Cox transformation or the BATEA methodology cannot be used to assess the model structure as these approaches combine the influence of such errors with the uncertainty in the modelled and measured flows.

The second test of a model demands more accuracy in the estimation of the uncertainty in the residuals, as the absolute uncertainty rather than the relative uncertainty is needed. By dividing each residual by its estimated uncertainty, the resulting scaled residual has an uncertainty given by the ratio of the actual and estimated uncertainty in the residual. Providing that the magnitude of the estimated uncertainty is sufficiently close to that of the actual uncertainty, the result is a set of scaled residuals, each with an uncertainty of approximately 1. The distribution of the ensemble of scaled residuals can therefore be investigated to test whether the complexity of the model is appropriate for the system being modelled. Assuming, for example, that the uncertainty estimates correspond to the standard deviation, the standard deviation of the scaled residuals should be approximately 1. If the standard deviation is significantly greater than 1, then providing that the estimated

uncertainties are sufficiently close to the actual uncertainties, the model is not adequately representing the response of the system, and is subsequently too simple. Conversely, if the standard deviation of the scaled residuals is significantly less than 1, then the model is over-fitting the data, and the model is too complex. Similar conclusions will result from other definitions of the uncertainty (e.g. 95% confidence bounds). Unfortunately, the uncertainty in the residuals cannot be accurately known, so the utility of this model test is reduced. In order to use this technique to assess model parameterisation, the modeller needs to be able to place confidence limits on the uncertainty estimates. If the uncertainty values are expected to be accurate to within 10%, then an acceptable model would have a value for the standard deviation of the scaled residuals between roughly 0.9 and 1.1.

5.3 Increasing information extraction through breaking –up and transforming streamflow time-series

Two other approaches can be used to increase the amount of information that can be extracted from time series during the calibration process. Breaking up the time series into components [e.g. Boyle *et al*., 2000; Wagener *et al.*, 2001; 2003; Vrugt *et al.*, 2006] and transforming time series to put the emphasis on calibrating the model to different aspects of the system response [e.g. Gupta *et al*., 1998; van Werkhoven *et al.*, 2008]. Both approaches can utilize standard performance metrics as discussed above. In the context of this paper it is assumed that the streamflow is the response variable of interest for simplicity.

Wagener *et al*., [2003] showed that calculating a performance metric like the Root Mean Squared Error (RMSE) on the residuals of the whole time series leads to a loss of information that could otherwise discriminate between different models or parameter sets. This problem can be reduced if the hydrograph is broken up into different periods and the performance metric is calculated for each period separately [e.g. Boyle *et al*., 2000; Wagener *et al*., 2001; 2003; Vrugt *et al*., 2006]. The break-up can be based on: [1] experience with a specific model so that different periods are likely controlled by different parameters, [2] based on hydrological understanding so that different periods represent different system response modes, or [3] based on an analysis of the information content so that periods with high information content for different parameters are separated out.

An alternative approach is the use of data transformations before model residuals are calculated. One method which has already been mentioned is the use of Box-Cox transformation [Gupta *et al.*, 1998]. Calculating the RMSE on the untransformed data leads to a focus on high flows (where the uncertainty is greatest), while doing the same on the transformed data leads to a focus on lower flows. A more process-based approach is the transformation of the observed and simulated streamflow data into hydrologically relevant signatures, and to calculate the performance metric on these signatures. Van Werkhoven *et al.* [2008] provides an example where performance metrics were calculated on water balance signatures (runoff ratio) and on signatures of the flow duration curve (central slope). The study showed that significantly more information became available for model identification than only using a traditional RMSE measure.

6. EXAMPLES

6.1 Assimilation of soil wetness data

The European project AIMWATER (Analysis, Investigation, and Monitoring of WATER Resources, http://dataserv.cetp.ipsl.fr/AIMWATER/) investigated the assimilation of additional soil wetness information into parsimoniously parameterized rainfall–streamflow models. For a catchment in France Aubert et al. [2003] showed how assimilation of soil wetness information from at-a-point Time Domain Reflectivity (TDR) measurements led to improvements in streamflow forecasting.

However, when a ‘standard’ IHACRES loss module was replaced by a more deterministic loss module, which in turn was calibrated using ‘soil wetness’ derived from satellite observations, unpublished work indicated that there was no clear improvement in model-fit to streamflow for a catchment in the UK. The law of diminishing returns appeared to apply. Given that the ‘standard’ IHACRES model with inputs of only rainfall and air temperature

already gave a reasonably good model-fit to streamflow, it would always be difficult to improve upon that. The fact that 'soil wetness' derived from satellite data, applied in a replacement loss module, did not lead to improved estimation of streamflow could be because the remotely-sensed data did not contain sufficiently good information, or because of the more deterministic loss module, or both. However, from a PUB perspective, remotely-sensed data may assist in setting parameter values of the water balance component of a rainfall runoff model when streamflow data is not available.

6.2 Impact of rainfall data quality on model performance and parameter values

The impact of rainfall and streamflow data quality on model performance has been extensively studied [e.g. Faurés *et al*., 1995; Hansen *et al*., 1996; Chaubey *et al.*, 1999; and Andréassian et al., 2001]. Faurés *et al*., [1995] found that the uncertainty in modelled flow was related to the number of raingauges, while Hansen *et al*. [1996] found that the primary factors affecting model performance were rain gauge density, stream gauge rating quality, catchment response dynamics (particularly the baseflow fraction). In a study based in the Little Washita Basin, Chaubey *et al.* [1999] found that the uncertainty in the AGNPS parameter values were strongly dependent on the spatial variability of the rainfall. Andréassian *et al.* [2001] studied the impact of imperfect rainfall data for three catchments on three hydrological models (GR4J, and derivatives of IHACRES and TOPMODEL), finding that the models used were able to correct for imperfect rainfall input estimates. For predicting flows at ungauged locations however, the variations in the imperfections in the rainfall data between catchments need to be taken into consideration.

Vaze et al. [2008] report the results of a study in which two models, SIMHYD and Sacramento were calibrated to ten catchments in the Murray-Darling Basin (see Figure 1) using rainfall data derived using four different methods:

1. Single raingauge;
2. Arithmetic mean of multiple raingauges;
3. Weighted mean of multiple raingauges (Theissen polygon);
4. 5km gridded rainfall.

Vaze et al. [2008] showed that the quality of the rainfall-runoff models (as measured with Nash-Sutcliffe efficiency, E) improved as the quality of the input rainfall improved (as determined by greater spatial representation of rainfall). They also showed that this improvement was greater for larger catchments than for smaller catchments and hypothesised that this was because individual raingauges are more likely to miss rain events entirely in larger catchments.

The performance of the models was evaluated, both over the calibration and an independent simulation period and the suite of model parameters produced using each method was derived.

The impact of this possible range of model parameters on the ability to regionalise hydrologic response was evaluated by building models with all possible model parameter sets and comparing the modelled streamflow from each of these models with observations over the simulation period.

7. DISCUSSION AND CONCLUSIONS

The uncertainty in the output from a model is a result of uncertainty in model inputs, parameter values, the observed outputs of the system, the suitability of the model structure, and the limitations in the information provided to the model by the input data. The impact additional data has on model uncertainty depends on the information content and uncertainty in the data used. In particular, higher spatial and/or temporal resolution leads to an increase in the information content, but at the cost of a potential increase in the uncertainty. For example, estimates of areal rainfall at a higher spatial resolution without increasing the number of raingauges will result in an increase in the uncertainty of the resulting rainfall estimates due to the reduction in averaging between gauges. Similarly, estimates at higher temporal resolution will have higher uncertainty due to the more variable nature of rainfall at shorter timescales. Introducing additional data (e.g. more

raingauges, or radar estimates of rainfall) has the potential to reduce the uncertainty in the rainfall estimates. The influence of additional data on a model's performance depends on:

1. the model's ability to extract information from the data
2. the sensitivity of the model performance to the uncertainty in the data (it may not be possible to separate this from the previous item).
3. how the model uses the additional data. If the data is used to constrain model parameters, then model performance will be improved providing the constraints become active at some point, or the model is used to bound the output and performance is measured by the width of those bounds. If the model is modified to incorporate the new data, then the model performance may or may not be improved depending on points 1 and 2 above.

The aim of model calibration is not to fit the data; rather it is to use the data to quantify (in terms of the model structure) the catchment response characteristics. This can only be achieved if the assessment of a model's performance takes into account the uncertainty in the model's output and the observations with which they are being compared. All performance criteria make some assumption regarding the distribution of the uncertainty (most assume constant additive uncertainty). If the distribution of the uncertainties is significantly different than what is assumed, then the performance criterion is biased. In the example given here, due to the highly heteroscedastic nature of the uncertainty in flow data, the NSE is biased towards data with highest uncertainty, with the bias proportional to the uncertainty squared. The result is that the NSE is biased towards high flows, as these have the greatest additive uncertainty.

The Box-Cox transformation attempts to correct for this bias, by transforming the quantities being compared (in this case, observed and modelled flow) to significantly reduce the heteroscedasticity in the model residuals. While this is a simple and effective technique to reduce the bias in the performance criteria, the adopted transformation is based on the aggregated impact of all contributors to the heteroscedasticity in the model residuals. Consequently, the significance of each contributor cannot be determined without additional information or assumptions.

An alternative approach is to consider the uncertainties in the observed output(s) of the system, as well as all model inputs and how these propagate through the model. One such approach is BATEA, where the distribution of uncertainties is fitted at the same time as the model parameters are estimated. A different approach to consideration of the uncertainties is to explicitly represent these in the performance criteria. While this is a considerable task, the advantage is that there is considerably more flexibility regarding the nature of the uncertainties, and a measure of the suitability of the model complexity to the application can be derived (it should be noted however, that the conclusions depend on the confidence in the uncertainty estimates). A Monte Carlo trial with synthetic data with error in the rating curve added showed that this approach yields significantly less uncertainty in the parameter estimates, and hence in the model predictions.

There are a number of techniques that have been developed for estimating the reduction in uncertainty resulting from the assimilation of new data, including measuring the reduction in the variances of distributions of model parameters or outputs, reduction in the volume of a behavioural set to estimation of the sharpness of the model predictions. These techniques allow the relative reduction in uncertainty to be quantified.

ACKNOWLEDGEMENTS

The authors would like to thank John Norton and Peter Young for providing very helpful discussions and comments on this paper.

REFERENCES

Aitken, A.C., On Least Squares and Linear Combination of Observations, *Proceedings of the Royal Society of Edinburgh, 55*, 42-48, 1935.

Andréassian, V. M. Bourqui, C. Loumagne, and M. Thibault, Reducing the uncertainty of streamflow simulation by accounting explicitly for the spatial variability of rainfall in a lumped rainfall-runoff model. Presented in Session HW2004, IUGG General Assembly, Perugia, Italy, July 2007.

Andréassian, V., C. Perrin, C. Michel, I. Usart-Sanchez, and J. Lavabre, Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models, *Journal of Hydrology, 250*, 206-223, 2001.

Aubert, D., C. Loumagne, and L. Oudin, Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall–runoff model. *Journal of Hydrology, 280*, 145-161, 2003.

Beven K.J., and A. Binley, The Future of Distributed Models: Model Calibration and Uncertainty Prediction, *Hydrological Processes, 6*, 279-298, 1992.

Box, G. E. P., and D. R. Cox, An analysis of transformations, *Journal of Royal Statistical Society, Series B, 26*, 211-246, 1964.

Boyle, D.P., H.V. Gupta, and S. Sorooshian, Towards improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resources Research, 36* (12), 3663-3674, 2000.

Bronstert, A. and A. Bárdossy, Uncertainty of runoff modelling at the hillslope scale due to temporal variations of rainfall intensity, *Physics and Chemistry of the Earth, 28*, 283–288, 2003.

Chaubey, I., C.T. Haan, S. Grunwald, and J.M. Salisbury, Uncertainty in the model parameters due to spatial variability of rainfall, *Journal of Hydrology, 220*, 48-61, 1999.

Croke, B.F.W., F. Andrews, A.J. Jakeman, S.M. Cuddy, and A. Luddy, IHACRES Classic Plus: A redesign of the IHACRES rainfall-runoff model, *Environmental Modelling and Software, 21*, 426-427, 2006.

Croke, B.F.W., The role of uncertainty in design of objective functions, In Oxley, L. and Kulasiri, D. (eds) *MODSIM 2007 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, December 2007, ISBN: 978-0-9758400-4-7, 2541-2547, 2007.

Di Luzio, M. J.G. Arnold, and R. Srinivasan, Effect of GIS data quality on small watershed stream flow and sediment simulations, *Hydrological Processes, 19*, 629-650, 2005.

Fraurés, J.-M., D.C. Goodrich, D.A. Woolhiser, S. Sorooshian, Impact of small-scale spatial variability on runoff modelling, *Journal of Hydrology, 173*, 309-326, 1995.

Freer, J. E., H. McMillan, J.J. McDonnell, and K.J. Beven, Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures, *Journal of Hydrology, 291*, 254-277, 2004.

Garrick, M., C. Cunnane, and J.E. Nash, A criterion of efficiency for rainfall-runoff models. *Journal of Hydrology, 36*, 375-281, 1978.

Gupta, H.V., S. Sorooshian, P.O. Yapo, Towards improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resources Research, 34*(4), 751-763, 1998.

Hansen, D.P., W. Ye, A.J. Jakeman, R. Cooke, and P. Sharma, Analysis of the effect of rainfall and streamflow data quality and catchment dynamics on streamflow prediction using the rainfall-runoff model IHACRES. *Environmental Software, 11*, 193-202, 1996.

Huard, D., and A. Mailhot, A Bayesian perspective on input uncertainty in model calibration: Application to hydrological model "abc". *Water Resources Research, 42*, W07416, doi:10.1029/2005WR004661, 2006.

Kalman, R.E., A New Approach to Linear Filtering and Prediction Problems, Transactions of the ASME - *Journal of Basic Enginneering, 82*, 35-45, 1960.

Kavetski, D., S.W. Franks, and G. Kuczera, Confronting Input Uncertainty in Environmental Modelling, in "*Calibration of Watershed Models*", by Q. Duan, H. Gupta, S. Sorooshian, A. Rousseau, and R. Tourcotte (eds), Water Science and Application Series 6, American Geophysical Union, Washington DC, 49-68, 2002.

Littlewood, I.G. and Croke, B.F.W., Data time-step dependency of conceptual rainfall-streamflow model parameters: an empirical study with implications for regionalisation (in prep).

Littlewood, I.G. Characterisation of river flow regimes for environmental and engineering hydrology: unit hydrographs for rainfall-streamflow modelling. Invited paper: *Folia Geographica*, (in prep. b).

Littlewood, I.G. Progress with unit hydrograph-based rainfall–streamflow models for engineering and environmental hydrology. Keynote: *Twelfth Biennial Conference of the*

Euromediterranean Network of Experimental and Representative Basins (ERB2008), Cracow, Poland, September 18-20, 2008, (in prep. a)

Littlewood, I.G., Rainfall-streamflow models for ungauged basins: uncertainty due to modelling time step. *Proc of the Eleventh Biennial Conference of the Euromediterranean Network of Experimental and Representative Basins (ERB2006)*, Luxemburg, UNESCO Technical Documents in Hydrology Series, 2007.

Marshall, L., A. Sharma, and D. Nott, Modeling the catchment via mixtures: Issues of model specification and validation. *Water Resources Research*, *42*(11): doi: 10.1029/2005WR004613. issn: 0043-1397, 2006.

McIntyre, N.R., and H.S. Wheater, Calibration of an in-river phosphorus model: prior evaluation of data needs and model uncertainty, *Journal of Hydrology*, *290* (1-2), 100-116, 2004.

Montanari, A., What do we mean by 'uncertainty'? The needs for a consistent wording about uncertainty assessment in hydrology. *Hydrological Processes*, *21*, 841-845, 2007.

Mungkasi, S., B.F.W. Croke, and A.J. Jakeman, Analytical approach for investigating the uncertainty in design of objective functions (in prep).

Nash, J. E. and J. V. Sutcliffe, River flow forecasting through conceptual models part I — A discussion of principles, *Journal of Hydrology*, *10* (3), 282–290, 1970.

Norton, J.P., Recursive linear estimation of linearly constrained parameters, *Electronics Letters*, 15 (22), 715-716, 1979.

Oudin, L., C. Perrin, T. Mathevet, V. Andréassian, and C. Michel, Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models. *Journal of Hydrology*, *320*, 62-83, 2006.

Perrin, C., L. Oudin, V. Andreassian, C. Rojas-Serna, C. Michel and T. Mathevet, 2007, Impact of limited streamflow data on the efficiency and the parameters of rainfall–runoff models, *Hydrological Sciences Journal*, *52*, 131-151, 2007.

Plackett, R.L., A historical note on the method of least squares, *Biometrika*, *36*, 458-460, 1949.

Plackett, R.L., Some theorems in least squares, *Biometrika*, 37, 149-157, 1950.

Qian, S.S., C.A. Stow, and M.E. Borsuk, On Monte Carlo methods for Bayesian inference, *Ecological Modelling*, *159*, 269-277, 2003.

Romanowicz, R.J., P.C. Young, and K.J. Beven, Data assimilation and adaptive forecasting of water levels in the River Severn catchment. *Water Resources Research*, *42* (W06407):doi:10.1029/2005WR004373, 2006.

Seibert, J., and J. J. McDonnell, On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour. Res*., *38*(11), 1241, doi:10.1029/2001WR000978, 2002.

Sivapalan, M., K. Takeuchi, S.W. Franks, V.K. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J.J. McDonnell, E.M. Mendiondo, P.E. O'Connell, T. Oki, J.W. Pomeroy, D. Schertzer, S. Uhlenbrook, E. Zehe, IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences, *Hydrological Sciences Journal*, *48* (6), 857-880, 2003.

Shrestha, R.R., A. Bárdossy, and F. Nestmann, Analysis and propagation of uncertainties due to stage-discharge relationship: a fuzzy set approach. *Hydrological Sciences Journal*, *52* (4), August 2007.

Son, K. and M. Sivapalan, Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. *Water Resources Research*, *43*, W01415, doi:10.1029/2006WR005032, 2007.

Sorooshian, S., V.K. Gupta, and J.L. Fulton, Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: Influence of calibration data variability and length on model credibility. *Water Recourses Research*, *19*, 251–259, 1983.

Tang, Y., P. Reed, and T. Wagener, How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? *Hydrology and Earth System Sciences*, *10*, 289-307, 2006.

Thyer, M.A., B. Renard, D. Kavetski, G. Kuczera, and S. Srikanthan, Bayesian Total Error Analysis For Hydrological Models: Preliminary Evaluation Using Multi-Site Catchment Rainfall Data, In Oxley, L. and Kulasiri, D. (eds) *MODSIM 2007 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, ISBN : 978-0-9758400-4-7, 2459-2465, 2007.

Van Werkhoven, K., T. Wagener, Y. Tang, and P. Reed, Understanding watershed model behavior across hydro-climatic gradients using global sensitivity analysis. *Water Resources Research*, *44*, doi:10.1029/2007WR006271, 2008.

Vaze, J., D.A. Post, F.H.S. Chiew, J.-M. Perraud, J. Pena, N. Viney, and J. Austin, Effect of climate input scale and quality on performance of conceptual rainfall-runoff models. *Water Down Under*, 15-17 April 2008, Adelaide, Australia.

Vrugt, J.A., B.A. Robinson and V.V. Vesselinov, Improved inerse modelling for flow and transport in subsurface media: Combined parameter and state estimation, *Geophysical Research Letters*, *32*, L18408, doi:10.1029/2005GL023940, 2005.

Vrugt, J.A., H.V. Gupta, S.C. Dekker, S. Sorooshian, T. Wagener, and W. Bouten, Confronting parameter uncertainty in hydrologic modeling: Application of the SCEM-UA algorithm to the Sacramento Soil Moisture Accounting model. *Journal of Hydrology*, *325*, 288-307, 2006.

Wagener, T. and H.V. Gupta, Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment*, *19*(6), 378-387. doi: 10.1007/s00477-005-0006-5, 2005.

Wagener, T., D.P. Boyle, M.J. Lees, H.S. Wheater, H.V. Gupta, and S. Sorooshian, A framework for development and application of hydrological models. *Hydrology and Earth System Sciences*, *5*(1), 13-26, 2001.

Wagener, T., M. Sivapalan, J.J. McDonnell, R. Hooper, V. Lakshmi, X. Liang, and P. Kumar, Predictions in Ungauged Basins (PUB) — A catalyst for multi-disciplinary hydrology. *Eos, Trans. AGU*, *85*(44), 451-452, 2004.

Wagener, T., N. McIntyre, M.J. Lees, H.S. Wheater, and H.V. Gupta, Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrological Processes*, *17*(2), 455-476, 2003.

Yadav, M., T. Wagener, H. Gupta, Regionalization of constraints on expected watershed response behaviour for improved predictions in ungauged basins, *Advances in Water Resources*, *30* (8), 1756-1774, 2007.

Yapo, P.O. H. V. Gupta and S. Sorooshian, Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data, *Journal of Hydrology*, *181* (1-4), 23-48, 1996.

Young. P.C., Advances in real-time flood forecasting. *Philosophical Trans. Royal Society, Physical and Engineering Sciences*, *360*(9):1433-1450, 2002.

Young. P.C., Rainfall-runoff modeling: Transfer Function models. In M. G. Anderson, editor, *Encyclopedia of Hydrological Sciences*, *volume 3*, part II, pages 1985-2000, Hoboken, N. J., 2005. John Wiley.

Young. P.C., Recursive Estimation and Time-Series Analysis. Springer-Verlag, Berlin, 1984.

Young, P.C., and H. Garnier, Identification and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems, *Environmental Modelling and Software*, *21*, (8), 1055-1072, 2006.

Young, P.C., and R. Romanowicz, PUB and Data-Based Mechanistic Modelling: the Importance of Parsimonious Continuous-Time Models, in *Complexity and Integrated Resources Management, Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society*, Pahl-Wostl, C., Schmidt, S., Rizzoli, A.E. and Jakeman, A.J. (eds), Volume 3, 1141-1146, iEMSs, 2004.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Effect of Rainfall Data Quality on Rainfall-Runoff Model Calibration and Regionalisation

D. A. Post[1], J. Vaze[2] F. H. S. Chiew[1], J.-M. Perraud[1], J. Teng[2], N. R. Viney[1]

[1]CSIRO Land and Water, Canberra, ACT, Australia

[2]Department of Water and Energy, Queanbeyan, NSW, Australia

E-mail: david.post@csiro.au

Keywords: Rainfall-runoff modelling, regionalisation, parameter values.

The majority of catchment water balance studies use conceptual rainfall-runoff models to estimate total daily runoff at the catchment outlet. In most applications, the observed runoff used to calibrate these rainfall-runoff models is only available at the catchment outlet and thus represents an aggregated response of spatially variable rainfall across the catchment. In many applications, the input climate data (rainfall and PET) used to calibrate the models is available only at limited locations within the catchment or even somewhere close to the catchment. Only in some applications is spatially variable climate data used to drive the models either as a lumped or distributed model.

Many regionalisation studies over the last decade have focussed on the transfer of rainfall-runoff model parameters from gauged to ungauged catchments with the intention of deriving a rainfall-runoff model for an ungauged catchment (i.e. one with no streamflow records). Sometimes this is done through the transfer of calibrated models from 'similar' or nearby catchments, and sometimes it is done through the transfer of model parameters based on their relationships with measurable catchment attributes. Where the latter technique has been used, it is based on the assumption that rainfall-runoff model parameters are representative of the hydrologic response of the catchment and are thus largely independent of each other and of errors in the rainfall inputs.

1. METHODS

To investigate the implications of input rainfall data and the model calibration and simulation methodology on the model calibration and regionalisation, two widely used daily rainfall-runoff models (SIMHYD and Sacramento) were calibrated using the four methods shown in Table 1 on 238 catchments varying in size from ~ 100 km^2 to 2,000 km^2.

Table 1: Calibration methods investigated

Method	Rainfall input	Models
1	Single raingauge closest to catchment centroid	Lumped
2	Weighted (Thiessen polygon) average of all raingauges	Lumped
3	Arithmetic average of 5 km (SILO) rainfall grid	Lumped
4	5 km (SILO) rainfall grid	Distributed

2. RESULTS

The calibration and simulation results for all 238 catchments clearly show that there is substantial improvement in model efficiencies with improved spatial representation of input rainfall data. The results also suggest that the improvements are generally greater in larger catchments than in smaller catchments. This is to be expected as there is a greater likelihood that rainfall will show greater spatial variability in larger catchments and therefore a single raingauge is unlikely to capture all of the rainfall events adequately.

Results from methods 1-3 indicate that the use of different rainfall inputs to calibrate the rainfall-runoff models can produce vastly different parameter values. This appears to be particularly true for the more highly parameterised Sacramento model compared to the less highly parameterised SIMHYD model. The implication of this result is that it is not possible to relate model parameters directly to catchment attributes and use these relationships to produce rainfall-runoff models in ungauged catchments. Also, models calibrated on less spatially distributed rainfall inputs do not perform as well when regionalised as those calibrated on more spatially distributed rainfall inputs. This implies that although we can transfer entire models from gauged to ungauged catchments, if the rainfall used to calibrate the model was poor, the resultant regionalised model will also be poor even if the rainfall in the ungauged catchment is of high quality.

Comparing results from method 4 with those from method 3, there are minor improvements in the efficiencies for both models when comparing a lumped model with a distributed model. However, in general these improvements are not as large as the improvements seen by using an improved representation of the rainfall.

3. DISCUSSION

These results were achieved using the SIMHYD and Sacramento models and may therefore only be applicable to these specific rainfall-runoff models. However, the authors have seen little evidence that the parameters of other rainfall-runoff models are less cross-correlated than these two and this believe these results to be more widely applicable. Similarly, the quality of rainfall inputs in the current study may be considered to be on the poor side (particularly for larger catchments), however calibrating a model on a handful of raingauges (or even just one) for catchments up to around 2,000 km^2 is common practice, at least in Australia where distances between rainfall gauges are often very large.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Prediction of Streamflow and the Effect of Input Parameters in Small Urban Catchments

H. J. Kim, S. J. Noh, C. H. Jang, Y. J. Lee

Hydrology Research Division, Korea Institute of Construction Technology, Goyang-Si, Republic of Korea

E-mail:hjkim@kict.re.kr

Keywords: Urban hydrology; land use; surface water; groundwater

The importance of prediction of streamflow or hydrological cycle in new town development project for ecosystem and human activity along the river front is gradually increasing in Korea. For the streamflow prediction of natural catchments, much kind of conceptual hydrologic models were developed and applied successfully with few parameters in gauged catchments [Perrin *et al.*, 2001] and much of papers are focused on the objective function and optimization method.

In the case of new town development, especially highly developed and ungauged area, the optimized parameters of the conceptual lumped model could not describe the changes of topography, land use, soil type and it is difficult to predict the streamflow changes.

Therefore, the authors use a distributed hydrologic model, WEP (Water and Energy transfer Process model, [Jia *et al.*, 2001] and WEP+ (Pre and Post Processor of WEP, [KICT, 2006]) for assessing surface and groundwater flow variations after new town development.

The water cycle before the construction was simulated and the new town development plan was made into modelling scenario which reflected changes of ground elevation (land filling and cutting), flow direction including urban drainage system, land use, surface soil, aquifer depths and etc.

The most of model parameters and inputs such as Manning's n (overland flow and channel flow), hydraulic conductivity of soil and aquifer were applied from the recommended value or filed investigation. The daily irrigation water supply for the paddy fields was surveyed and the aquifer depths were investigated by the boring wells and spatially interpolated over the catchment.

Simulation results for the scenarios were analysed with hydrographs, flow duration curves and annual water budgets. When comparing the water cycle before new town development, the peak flow is increased and base flow is decreased severely because a lot of pervious area is changed to impervious. The change of unconfined aquifer depth by the land filling and cutting has also significant role to water cycle after the new town development.

REFERENCES

Jia, Y., G. Ni., Y. Kawahara, T. Suetsugi (2001), Development of WEP model and its application to an urban watershed, Hydrological Processes, Vol. 15., pp. 2175-2194

KICT (2006), Development of watershed assessment techniques for healthy water cycle, KICT 2006-071

Perrin C., C. Michel, V. Andreassian (2001), Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, J. Hydrol., 242, pp. 275-301.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Ensemble modelling of runoff across southeast Australia and quantification of uncertainty

Francis H.S. Chiew[1], Neil R. Viney[1], Jai Vaze[2], David A. Post[1],Jean-Michel Perraud[1] and Jin Teng[2]

[1]*CSIRO Land and Water, Canberra, Australia*

[2]*Department of Water and Energy, NSW, Australia*

Email: francis.chiew@csiro.au

Keywords: Rainfall-runoff modelling, uncertainty, ensemble modelling, regionalisation.

Reliable information on current and future water availability are required to properly manage and share limited water resources. This paper describes an ensemble modelling approach used to estimate runoff for 0.05° x 0.05° grid cells (~ 5 km x 5 km) across southeast Australia. There are about 250 'unimpaired' streamflow gauging stations with more than 50 percent data from 1975 to 2006, and they cover less than 1 percent of the over one million square kilometres study region.

Several lumped conceptual daily rainfall-runoff models are used. Interpolated daily rainfall and potential evapotranspiration data for 0.05° grid cells are used to drive the models. For each grid cell, runoff can be estimated using any of the rainfall-runoff models with optimised parameter values from any of the gauged catchments. For the application here, an ensemble modelling approach is used where runoff for each grid cell is estimated as a weighted combination of the runoff from the different models using parameter values from different gauged catchments. The weights and the number of ensembles used are determined from cross-validation analyses of modelling results for the gauged catchments.

The models are calibrated to maximise the Nash-Sutcliffe efficiency of daily runoff with a bias constraint used to ensure that the total modelled runoff is within 5 percent of the total observed runoff. In the cross-validation analyses, each catchment is considered in turn as an 'ungauged' catchment, and runoff for that catchment is estimated using each of the models with optimised parameter values from each of the remaining gauged catchments. The Nash-Sutcliffe efficiencies of daily runoff and biases in the total runoff for each of the simulations are calculated. The results are then analysed to determine the models and characteristics of the donor catchments (whose parameter values are used) that give the better runoff simulations for the 'ungauged' catchments. The models and donor catchment characteristics considered include: the different rainfall-runoff models; quality of model calibration in the donor catchment; and catchment similarity measures (geographical proximity, precipitation, dryness index, terrain, stream density, woody vegetation, effective soil storage, soil transmissivity, etc…).

Based on the cross-validation analyses, weights are used to favour the choice of models and donor catchments to estimate runoff for an ungauged catchment or grid cell. As the runoff simulations come from a weighted ensemble of many simulations, the uncertainty in the runoff simulations can also be determined directly.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Lessons Learned from Experiments in Auto-Calibration of Large Scale Environmental and Hydrological Models

David A. Swayne, Markiyan Sloboda, Simon Hood, Vimal Sharma, William Booty,

David Lam

Computing Research Laboratory for the Environment

Department of Computing and Information Science, University of Guelph,

Guelph, Ontario, Canadaf

E-mail: dswayne@cis.uoguelph.ca

Keywords: model calibration; knowledge discovery; uncertainty management; genetic algorithm; data mining

Recent work by the authors has identified the need to efficiently auto-calibrate models of considerable complexity, in order to investigate future water quality and quality concerns in the Great Lakes' Basin. Initial results in the use of standard models such as AGNPS and SWAT have yielded somewhat disappointing results. The calibration process has been cumbersome, and predictions based on outputs from, say, a watershed model input to a lake circulation model have generally proven problematic: If the flows are correct, the loadings are not, and so forth.

With the use of a supercomputer it is possible to generate a large number of model runs in a very short time, but the inherent sequential nature of a search for a suitable parameter set is at cross-purposes with parallel search strategies. We are, therefore, investigating the incorporation of tools borrowed from the artificial intelligence community for knowledge discovery and uncertainty management. Our experiments have included: genetic algorithm approaches, data mining, and modeling the models themselves with Bayesian Belief Network approximations.

The models we consider consist of: 1) AGNPS and SWAT; 2) lake circulation models of varying complexity; 3) several choices of hydrological drivers of varying complexity. Rainfall and sediment loads, nutrients and phytoplankton are the major measured parameters against which calibration progress is measured. For lake circulation models, long-term simulations require calibration involving wind, temperature observations and solar radiation.

Counterintuitive strategies such as joint calibration of hydrological, erosion / sediment and biological variables (rather than step-wise starting with hydrological) might indicate sources of uncertainty to explain less-than-perfect agreement.

Using the SHARCNet super computer clusters we are conducting series of experiments with the available models to see if we can improve upon results obtained "by the book" which are, to date, disappointing. We will present our findings and conclusions at the workshop.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Quantifying input and model errors in conceptual rainfall-runoff models using Bayesian total error analysis

Dmitri Kavetski, Ben Renard, George Kuczera, Mark Thyer and Stewart Franks

School of Engineering, University of Newcastle, New South Wales

Email: Dmitri.Kavetski@newcastle.edu.au

Keywords: conceptual rainfall-runoff modelling, parameter calibration, input uncertainty, model error, Bayesian hierarchical inference

Calibration and prediction in conceptual rainfall-runoff (CRR) modelling is affected by the sampling and measurement uncertainty in the forcing/response data and by the structural error of the model conceptualisation. This study presents a robust Bayesian Total Error Analysis methodology (BATEA) for dealing with these multiple sources of uncertainty. The core idea is to pose the CRR model calibration as a Bayesian hierarchical model with latent variables describing uncertainties in the data and the CRR model. This provides the opportunity to directly and comprehensively address all sources of uncertainty. A critical challenge is to characterize model error. In the past this has been thwarted by the convenient but indefensible treatment of CRR models as deterministic descriptions of catchment dynamics. Here it is argued that CRR models are fundamentally stochastic because sub-grid variability of catchment processes in time and space cannot be uniquely described by models operating at hillslope or larger scales. Acceptance that CRR models are intrinsically stochastic paves the way for a more rational description of model error. We advance the hypothesis that CRR model error can be characterized by storm-dependent random variation of one or more CRR model parameters. In addition to model error, we describe the treatment of data uncertainty in the observed input and output data using the hierarchical BATEA methodology. A practical implementation of BATEA requires the estimation of a large number of latent variables (realisations of storm-dependent model parameters). This yields high dimensional optimisation and sampling problems to estimate the most likely parameter sets and to characterise their uncertainty. Optimisation strategies using Newton-type methods are presented, along with practical implementations of Metropolis and Gibbs samplers to elicit the full posterior distribution of model parameters. Diagnostic tests guiding the selection of the distributions of latent variables are discussed. Case studies are presented illustrating the BATEA approach and the insights it generates about predictive uncertainty in CRR models operating under conditions of data and model uncertainty.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

New hybrid method to simulate discharges on ungaged catchments by combination of hydrological and hydraulic models

J. Lerat[a], V. Andréassian, C. Perrin

[a] *Cemagref, Hydrosystems and Bioprocesses Research Unit, Antony, France*

E-mail: julien.lerat@cemagref.fr

Keywords: Hydraulic model, hydrological modelling, coupling, lateral inflows

The estimation of discharges on ungauged basins is a central issue in hydrology and resulted in various modelling developments during the PUB decade. In this field, the potential of distributed and semi-distributed models is high considering their capacity to extract results on interior points of a catchment based on the calibration on downstream discharges. Unfortunately, recent attempts [Smith, 2004, Reed et al., 2004] proved disappointing when such models are compared to lumped ones on objective basis, revealing a paradoxical picture: spatial complexity integrated in distributed models may increase errors that lumped models can limit thanks to a compact structure.

As a continuation of the work of Boyle et al. [2001] and Ajami et al [2004], we propose here to explore the question of interior simulations in a particular context: in many basins, a gauging station can be found upstream of the interior location reducing the problem of ungauged basin to ungauged river reach. Interior discharges are then a combination of propagated upstream discharges and intermediate lateral inflows. This question is of particular importance for river hydraulic modellers that frequently struggle to determine lateral inflows [Vidal et al., 2007] to their model.

Our study applies a semi-distributed model (SD) using GR4J [Perrin et al, 2003] as a rainfall-runoff module and a linearised diffusing wave routing scheme [Moussa, 1996] on 49 French river reaches at the hourly time{step. Lumped models and pure propagation models are also applied as benchmarks. The improvements of interior simulations from benchmarks to SD models are then related to physical characteristics and recommendations are given to estimate a priori interest for complex semi-distributed schemes.

REFERENCES

Ajami, N., H. Gupta, T. Wagener, S. Sorooshian, (2004). Calibration of a semi-distributed hydrologic model for stream flow estimation along a river system. J. Hydrol., 298(1-4):112-135.

Boyle, D. P., H.V. Gupta, S. Sorooshian, V. Koren, Z. Zhang, M. Smith (2001). Toward improved stream flow forecasts: Value of semidistributed modeling. Water Resour. Res., 37(11):2749-2759.

Moussa, R. (1996). Analytical hayami solution for the diffusive wave flood routing problem with lateral in flow. Hydrol. Processes, 10(9):1209-1227.

Perrin, C., C. Michel, V. Andréassian, V. (2003). Improvement of a parsimonious model for stream flow simulation. J. Hydrol., 279(1-4):275-289.

Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D.-J. Seo (2004). Overall distributed model intercomparison project results. J. Hydrol., 298(1-4):27-60.

Smith, M. B., K. P. Georgakakos, X. Liang (2004). The distributed model intercomparison project (dmip). J. Hydrol., 298(1-4):1-3.
Vidal, J. P., S. Moisan, J.B. Faure and D. Dartus (2007). River model calibration, from guidelines to operational support tools. Environmental modelling and software, 22.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Uncertainty and evolution of rating curves: a key issue for the reliability of rainfall-runoff models?

T. Mathevet[a], L. Oudin[b], V. Andreassian[c], C. Perrin[c]

[a] *EDF-DTG, 21 avenue de l'Europe, BP 41, 38040 Grenoble cedex 09, France*

[b] *Université Pierre-et-Marie Curie, UMR Sisyphe, 4 place Jussieu, 75252 Paris cedex 05, France*

[c] *Cemagref, Hydrosystems and Bioprocesses Research Unit, Antony, France*

E-mail: thibault.mathevet@edf.fr

Keywords: Rating curve, rainfall-runoff modelling, uncertainty, extrapolation

The inputs of rainfall-runoff (RR) models (precipitation, temperature) and the data used for the calibration of their parameters (typically streamflow data) are known to be uncertain, due to measurement and spatial extrapolation errors. In the literature, the impact of input variables uncertainty [Oudin et al., 2006] and model structure [Perrin et al., 2001] on RR model simulations have been largely studied. However, there are very few studies dealing with the impact of the uncertainty of streamflow data used for calibration on the outputs of RR models. Those impacts can be potentially considerable since streamflows are used to calibrate the RR models. The problems related to the construction and extrapolation of rating curves are especially crucial and should deserve more attention.

In this paper, we propose to study the impact of streamflow uncertainty on RR model simulations and parameters. To explore this issue, we use three methods to generate runoff uncertainty:

1. a method based on the use of historical stream gauging and evolution of gauging site;
2. a method based on the use of a 'wrong' rating curves, which is the case when gauging site are poorly maintained;
3. a method based on the use of different extrapolation laws.

In each case, we then calibrate the RR model using these corrupted flow data and assess the reliability and robustness of the RR model. The impact of runoff uncertainty on model performance and parameters is demonstrated using a sample of catchments for which the evolution of the gauging measurements and rating curve are well-documented.

REFERENCES

Perrin C., C. Michel, V. Andreassian (2001), Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. J. Hydrol, 242, 275-301, doi:10.1016/S0022-1694(00)00393-0

Oudin, L., C. Perrin, T. Mathevet, V. Andréassian, and C. Michel (2006), Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models, J. Hydrol., 320, 62-83, doi:10.1016/j.jhydrol.2005.1007.1016.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

On the role of pre and post-processing in environmental data mining

Karina Gibert[1,2], Joaquín Izquierdo[3], Geoff Holmes[4],Ioannis Athanasiadis[5], Joaquim Comas[6], Miquel Sànchez-Marrè[2]

[1] *Department of Statistics and Operation Research, Technical University of Catalonia, Barcelona, Catalonia*
[2] *Knowledge Engineering and Machine Learning Group, Technical University of Catalonia, Barcelona, Catalonia*
[3] *Centro Multidisciplinar de Modelación de Fluidos, Universidad Politécnica de Valencia, Spain.*
[4] *Department of Computer Science, University of Waikato, New Zealand*
[5] *Istituto Dalle Molle di Studi sull'Intelligenza Artificiale, Lugano, Switzerland*
[6] *Laboratory of Chemical and Environmental Engineering (LEQUIA), University of Girona*

Abstract: The quality of discovered knowledge is highly depending on data quality. Unfortunately real data use to contain noise, uncertainty, errors, redundancies or even irrelevant information. The more complex is the reality to be analyzed, the higher the risk of getting low quality data. Knowledge Discovery from Databases (KDD) offers a global framework to prepare data in the right form to perform correct analyses. On the other hand, the quality of decisions taken upon KDD results, depend not only on the quality of the results themselves, but on the capacity of the system to communicate those results in an understandable form. Environmental systems are particularly complex and environmental users particularly require clarity in their results. In this paper some details about how this can be achieved are provided. The role of the pre and post processing in the whole process of Knowledge Discovery in environmental systems is discussed.

Keywords: preprocessing; postprocessing; data mining; knowledge discovery of data; statistics; artificial intelligence; environmental systems.

1. INTRODUCTION

Environmental systems (ES) typically contain many interrelated components and processes, which may be biological, physical, geological, climatic, chemical, or social. Whenever we attempt to analyze ES and associated problems, we are immediately confronted with complexity stemming from various sources. However, there is a great need for data analysis, modelling of ES and development of decision support systems in order to improve the understanding of ES behaviour and the management of these complex systems (specially under abnormal situations). As stated in [Gibert et al.2008], the special features of environmental processes demand a new paradigm to improve the analysis and consequently its management.

Knowledge Discovery of Data (KDD) appeared in 1989 referring to high level applications which include particular methods of *Data Mining (DM, see figure 1)*, oriented to extract useful and understandable knowledge from data. KDD processes and the application of DM techniques are specifically appealing for environmental data, since activities permitting

extraction of maximum useful information from data bases are *per se* very important although they use to be preparatory for an environmental software system development. Also the KDD approach facilitates the integration of different knowledge sources and fields of expertise and the involvement of end-user (domain expert) criteria and stakeholders' points of view in algorithm design and result interpretation. Finally, it facilitates the sharing and rapid re-use of data and extracted technical knowledge and experiences among domain experts.

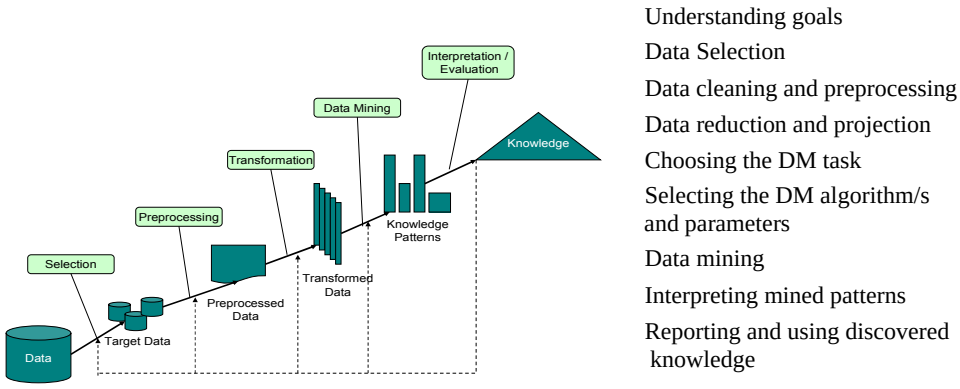


Figure 1 Outline of the Knowledge Discovery from Data process

Fayyad's proposal marked the beginning of a new paradigm in KDD research, considering prior and posterior analysis as important as the application of DM techniques itself:

> *"Most previous work on KDD has focused on [...] DM step. However, the other steps are of considerable importance for the successful application of KDD in practice".*

In fact, prior and posterior analysis requires great effort when dealing with real applications. Prior analysis is critical, mainly owing to two reasons:

- Real data sets tend to be imperfect, contains errors, outliers, missing data, extra noise and tools either for detecting or correcting it are required.

- Application of a certain data mining technique may require specific conditions for the data set (only binary variables, centered data, normality, only qualitative variables, etc). In this case, tools for verifying that those conditions hold as well as to transforming data in the appropriate way in order that they hold them, are required.

In environmental data, where errors of measuring (from automatic monitoring), uncertainty, imprecision, multi-scalarity, heterogeneity, non-linearities, non stacionariety, non-normality, are frequent, together with redundant variables, or irrelevant, or even contradictory, systematic and objective exploration as well as visualization and transformation of data is particularly critical for:

- better understanding the data set
- detecting imperfections in the data and managing them in the proper way
- correctly preparing data for the selected DM technique/s, if required assumptions do not hold

In (Witthen et al. 1993) a review of selected methods of machine learning with an emphasis on practical applications is presented, together with suggestions on how they might be used to address some important problems in the primary production industries, particularly agriculture.

Also, particular efforts in post processing the results directly provided by a DM technique are important in this context, in order to make these results directly understandable by an environmental scientist, who has to make real decisions upon them, which will surely have a real impact to the evironmental system behaviour. In fact, it can be said that the quality of the decisions will depend, not only on the quality of the data mining results themselves, but also on the capacity of the system to communicate the relevant results to the decision-maker as understandably as possible. The software tools used for applying DM techniques to a data set usually produce long listings plenty of results that may be useful in whatever particular application is performed. However, given a particular real case, not all this information is useful. Thus, it is important to:

- Identify the relevant information from the software outputs, depending on the aims of every particular analysis.
- Find the best way to present the selected results to the user.

It can be said that pre and post processing are one of the most important and critical parts of the whole KDD process. On the one hand, because *data cleaning* [Moore et alt, 1993], transformation, selection of DM techniques and optimization of parameters (if required) are often time consuming and difficult, mainly because the approaches taken should be tailored to each specific application, and human interaction is required; on the other hand, because the correctness of the DM itself is critically depending on the quality of the data and wrong or poor preprocessing may lead to incorrect results. Data miners should become conscious of the importance of performing very careful and rigorous preprocessing, and allocate sufficient time to this activity. In fact, once those tasks have been accomplished, the application of DM algorithms becomes trivial and can be automated, requiring only a small proportion of the time devoted to the whole KDD process. Regarding the post processing, the results provided by the software implementing the DM techniques, select of relevant information from the automatic outputs produced by the different software applications, and choose the proper way of transforming or synthesizing it to make directly understandable to the user is also critical to provide good decision support. This task is also difficult to standardize and time consuming and has to be designed ad-hoc for every particular application, requiring much human guidance. In real applications, the time devoted to both pre and post-processing is rarely below 70% of the time for the whole KDD process.

In this paper issues related to pre and post processing in environmental applications will be addressed, together with the most popular solutions for the different cases. This paper does not pretend to be exhaustive, but to provide tools to environmental scientists for addressing the most common problems arisen in pre and post processing, as no clear methodology for tackling these two steps of KDD process has no been established yet.

2. PREPROCESSING IN ENVIRONMENTAL SYSTEMS

As previously stated, environmental data often includes measurement errors (from automatic monitoring, sensors, etc), uncertainty, imprecision, multi-scalarity, heterogeneity, non-linearities, non stacionariety, non-normality, and tools are required to:

- better understand the data set
- detect imperfections in data sets and manage them in the proper way
- correctly prepare data for the selected DM technique/s, if required assumptions do not hold

This section is devoted to providing elements for pre-processing environmental data for KDD. Pre-processing ranges from the simplest descriptive techniques to the more sophisticated data analysis methods, depending on the nature of data and the goals of the

analysis itself. Authors think that most of the operations performed in a pre-processing step can be reduced to two main families of techniques:

- Detection techniques: Those oriented to detect imperfections in data sets or to verify the accomplishment of required assumptions for a particular analysis:
    - Outlier detection
    - Missing data detection
    - Influent observations detection
    - Normality assessment
    - Linearity assessment
    - Independence assessment
- Transforming techniques: Those oriented to perform transformations in the data set in order to correct the imperfections detected before, or to achieve the technical conditions to apply a certain analysis technique.
    - Outlier treatment
    - Missing data imputation
    - Dimensionality reduction techniques or data projection techniques
    - Creation of new transformed variables
        - Standardization
        - Aggregation
        - Transformation (logarithmic, quadratic)
        - Discretization
        - Recodification
    - Filtering
    - Resampling

Classically, statistics provided a wide set of possibilities to preprocess data. The global process of transforming a raw data set to a correct one ready for analysis is called *data cleaning* [Moore et alt, 1993]. Besides the classical statistical techniques for *data cleaning*, inductive techniques are an alternative for several activities of the environmental scientist, when analytical/traditional methods fail, are too slow, or simply do not exist. Finally, visualization techniques also play an important role in the correct preprocessing of data. In the following sections different situations and possibilities are addressed.

2.1. Visualization

Visualization is a powerful strategy for leveraging the visual orientation of sighted human beings. Sighted humans are extraordinarily good at recognizing visual patterns, trends and anomalies; these skills are valuable at all stages of the KD Miller (in press, to appear in 2007). For example, the presence of outliers, missing values, or errors are typical pre- and post-processing KDD tasks where visualization techniques can be valuable.

Graphs commonly used for classical exploratory visualization, like boxplots, histograms, time series plots or two-dimensional scatter plots, perform poorly considering the great number of variables involved in environmental datasets, along with their complex interrelations, and spatial-temporal references. Thus, more sophisticated visualization methods are required, as for example:

- Distributional plots,
- Three, four, and five dimensional plots (color and symbols may be used to represent the higher dimensions),
- Dimension scaling, for example log scales,
- Rotatable frames,
- Animation with time and interactive graphs,
- Geo-referenced visualizations and maps.

Most DM packages, as Weka, include visualization tools, while more advanced features are provided with wide-spread tools such as Matlab or a dedicated data language such as IDL

or the CommonGIS tool (Andrienko and Andrienko 2004). Reader is also pointed to dedicated visualization tools such as XGobi (Swayne et al. 1998). Visual representations are extremely effective, and may convey knowledge far better than numerical or analytical forms. They should be always considered in environmental KDD.

2.2. Outlier Detection

Outliers are objects with very extreme values in one or more variables (Barnett and Lewis 1978). Graphical techniques were once the most common method for identifying them, but increases in database sizes and dimensions have led to a variety of automated techniques. The use of standard deviations is possible when and only when considering a single variable that has a symmetric distribution, but outliers may also take the form of unusual combinations of two or more variables. The data point should be analyzed as a whole to understand the nature of the outlier and multivariate approach is required.

The treatment will depend on the nature of the outlier (error, member of another population, intrinsic extreme value, etc). The influence of outliers can dramatically affect the results or certain methods, a concern which should feature the choice of tools used throughout the rest of the process. See Moore and McCabe (1993) for an interesting discussion on the dangers of simply eliminating rows with outliers:

> *"In 1985 British scientists reported a hole in the ozone layer of the Earth's atmosphere over the South Pole. [...] The British report was at first disregarded, since it was based on ground instruments looking up. More comprehensive observations from satellite instruments looking down had shown nothing unusual. Then, examination of the satellite data revealed that the South Pole ozone readings were so low that the computer software [...] had automatically suppressed these values as erroneous outliers! Readings dating back to 1979 were reanalyzed and showed a large and growing hole in the ozone layer [...] suppressing an outlier without investigating it can keep valuable out of sight."* (Moore and McCabe 1993).

2.3. Missing Value Treatment

Sometimes, a number of cells are missing from the data matrix. These cells may be marked as a *, ?, NaN (Not a Number), blank space or other special character or special numeric code such as 99999. The latter can produce grave mistakes in calculations if not properly treated. It is also important to distinguish between random and non-random missing values (Allison 2002; Little and Rubin 1987). Non-random missing values are produced by identifiable causes that will determine the proper treatment, also influenced by the goals of the task. Imputation (see Rubin 1987) is a complex process for converting missing data into useful data using estimation techniques. It is important to avoid false assumptions when considering imputation methods, which may have a significant effect on the results extracted. All the methods have pros and cons, and the choice must be made with care. In particular, removing rows with missing cells from a dataset may cause serious problems if the missing values are not randomly distributed. It is of utmost importance to report any elimination performed.

However, many data-mining algorithms can treat missing values. Such algorithms should be preferred, in cases of scarce datasets.

A method proposed in Gibert et al.(2008b) is to perform a previous clustering with complete variables and to impute missing values locally to the discovered clusters. Such an approach may even help to identify errors in the data set (if some observation places in a wrong cluster and the reasons for that are investigated). Another method proposed in Athanasiadis and Mitkas (2007) was to use qualitative indicators for substituting missing or erroneous air quality measurements. In operational systems, instead of making numerical estimates of a missing records, data-mining techniques may be used for making qualitative

estimates (i.e by assigning quality labels instead of predicting values). Similar work has been done for substituting missing or erroneous values with fuzzy sets, probability distributions or confidence intervals.

2.4. Pre-processing imbalanced data sets

Imbalanced sets appear in many real applications areas, where the number of instances of one class (usually the normal class) is much bigger than the number of instances of the other classes (usually abnormal classes), which frequently are the most important ones, since they are the object of interest in the specific problem. Some applications in which imbalanced data sets typically appear are the identification of anomalies in certain systems designed to work under steady conditions, such as Water Distribution Systems [Izquierdo et al., 2007], the detection of oil spills from satellite images [Kubat et al., 1998], the identification of power distribution fault causes [Xu et al., 2007] and the prediction of pre-term births [Grzymala-Busse et al., 2003]). This issue is growing in importance since it appears more and more in most real domains of classification, especially in systems where normal data are abundant while abnormal ones are scarce.

Most classical machine learning algorithms generally perform poorly on imbalanced data-sets because they are designed to minimize the global error rate [Japkowicz and Stephen, 2002] and in this manner they are biased toward the majority class, that is to say, they tend to classify almost all instances within the majority class, and thus poorly classify the minority class examples.

Two kinds of solutions can be used to cope with imbalanced sets: pre-processing solutions, trying to balance data by over-sampling the minority classes, under-sampling the majority one or a combination of both (see [Batista et al., 2004; Chawla et al., 2002; Guo and Viktor, 2004]), and solutions at the algorithmic level, modifying the cost per class [Provost and Fawcett, 2001], adjusting the probability estimation by establishing a bias towards the minority classes [Weiss and Provost, 2003], etc.

In [Fernández et al., 2008] different pre-processing mechanisms are used in conjunction with a Fuzzy Rule Classification System to deal with imbalanced data sets.

As an example, in [Izquierdo et al., 2007] a complex hybrid model, which uses a calibrated classical model of a Water Supply System and a neuro-fuzzy technique, is used to obtain diagnosis of leaks and other anomalies in the system. To train the neural network both normal and abnormal data are needed. While normal data are abundant, abnormal ones are scarce since only a reduced number or records from reported anomalies are usually available. The calibrated classical model of the hydraulic network is used to simulate abnormal data, thus producing a data set with enough fuzzy examples to correctly train the neural network.

2.5. Uncertainty in environmental data

Most of the approaches used in KDD assume precise data. They assume that we deal with exact measurements. But in most, probably in all real-world scenarios, always imprecise measurement is usually obtained. There is always a degree of uncertainty. Even being able to measure a magnitude, its exact value will be never known. It can only be known that the measurement is somewhere in a certain range, bounded by the precision of the measurement instrument or measurement procedure itself.

The multiplicity of factors involved in environmental processes provokes uncertainty in data. A number of issues contribute to this complexity. Among others (i) a huge number of sensors are involved, (ii) knowledge related to the spatial structure of these sensors is not well known, (iii) different types of variables can affect the same sensor, (iv) one specific

variable can exhibit simultaneously different states. In addition, subjectivity is often influential and fuzziness becomes important.

Tracking and reporting of uncertainties related to measurement and other sources of noise is an area that is sometimes not treated rigorously, despite the implications. Therefore, the minimum theoretically achievable error of any model built on the data cannot be less than the error contained in the original data. Models with reported fit greater than this are overfitted and their performance measures do not reflect true predictive capacity. In general, as much data as possible should be available, there is less uncertainty, or at least that uncertainty can be better quantified.

To be exact, all recorded data involves uncertainty in some neighborhood, and the real measured values are really somewhere inside a certain interval, with a width depending on the accuracy of the measurement procedure. It has to be noted that we are not speaking about a probability, but about some kind of likelihood that certain crisp value is being obtained [Zadeh, 1995]. Thus, the use of single values (means, for example, or maximum and minimum values), although it implies a quick and sometimes convenient way of calculation or completion of registers with missing values, it is often simplistic and it should be corrected with appropriate safety values. The statistical approach of subjectively attributing probabilities to the inaccuracies many times gives the problem an artificial character, involving a high degree of randomness and runs the clear risk of inventing information about unknown distributions.

An appropriate conceptual tool for this type of data is the theory of fuzzy sets [Zadeh, 1965]. Many environmental research works make use of fuzzy logic to model uncertainty (see [Bazartseren et al., 2003; Juang, 2003; Oh and Pedrycz, 2004; Kuncheva et al., 2000; Faye et al., 2003; Izquierdo et al., 2006], among others).

2.6. Transformation and Creation of Variables

Sometimes transformation of variables may assist analysis. For example, normality may be forced when using ANOVA or, for ease of interpretation, variables with a large number of categorical labels can be grouped according to expert knowledge. Under some circumstances, discretization of continuous variables is appropriate (eg *Age* into *Child* under 18 years, *Adult* between 18 and 65 years, *Elderly* over 65 years). In fact, in real applications it is quite common to globally discretize any numeric attributes before applying learning algorithms to datasets, since a number of them cannot handle numeric attributes directly. In these cases prior discretization is essential. Even if it can, prior discretization often accelerates induction, and may produce simpler and more accurate classifiers. As it is generally done, global discretization denies the learning algorithm any chance of taking advantage of the ordering information implicit in numeric attributes. However, a simple transformation of discretized data preserves this information in a form that learners can use. In (Frank et alt, 99) it is shown that, compared to using the discretized data directly, this transformation significantly increases the accuracy of decision trees built by C4.5, decision lists built by PART, and decision tables built using the wrapper method, on several benchmark datasets. Moreover, it can significantly reduce the size of the resulting classifiers. This simple technique makes global discretization an even more useful tool for data preprocessing.

Noise is often a critical issue, and especially with environmental data some bias may exist that can be removed with a filter. Transformations should always be justified and documented, and the biases that may be introduced noted (Gibert and Sonicki 1999). Interpretability of transformed variables should be kept.

Creation of additional variables is also used in KDD. Here, expert knowledge is usually the guide. Exploratory variable creation without such assistance is almost always prohibitively time consuming, and as noted, may obfuscate physical interpretation and exacerbate noise. Efficient techniques for data reduction, however, do exist and are well used.

However, avoidance of unnecessary transformations is recommended, especially if the transformation decreases interpretability (for example Y = log(streamflow), although Y is normal). If transformations are definitely required, some bias may be introduced into the results; thus, it is convenient to minimize arbitrariness of the transformation as much as possible (in recoding *Age*, Adult may be defined from 18 to 65 or from 15 to 70), and this implies that the goals of the analysis must also be taken into account. For arithmetic transformations, imputation of missing data before the transformation is thought to be better. Note that where data is numerical and the scales changes between variables, normalization may be necessary.

Finally, a good practice is to select a data mining technique that fits well on the target phenomenon and the kind of available data, instead of adopting the approach of strongly transform the data till it fits on the technical assumptions of the preferred data mining technique, in spite of loosing interpretability, or introducing arbitrariness. This means that if data is non-normal, it is better to see, first of all, if there exist a data mining technique suitable for this purpose which does not requires the normality to be applied; or this means that an alternative to ID3 should be used if data is numeric, instead of forcing discretization.

2.7. Data Reduction and Projection

When the number of variables is too high to deal with in a reasonable way, which is not unusual in data mining context, a data reduction method can be applied. Either Data Projection or Feature Selection methods are suitable possibilities in these cases.

2.7.1. *Data projection*

This may be accomplished by eliminating some variables wholesale, or projecting the feature space of the original problem into a reduced fictitious space, with fewer dimensions. Principal Components Analysis (PCA) (Dillon and Goldstein 1984) is one of the best known techniques used for the latter purpose. Each principal component is a linear combination of the original variables, and the aim is to work with a reduced set of these, such that the loss of information is not relevant. Thus, Principal Component Analysis is suitable for synthesizing an original set of numerical variables into a small number of fictitious variables conserving as much information as possible from original dataset. Equivalent techniques are available for qualitative data, like multiple correspondence analysis (Lebart et al. 1984; Dillon and Goldstein 1984). However, it has to be taken into account that in most cases, interpretation of the new variables (or factors) may not be clear, and if this is the case, there will be implications for understandability of the final results.

2.7.2. Feature Selection and Feature Weighting

Datasets may contain irrelevant or redundant variables (Gibert et al. 2008). As previously stated, the quality of discovered knowledge is usually dependant on the quality of the data that they operate on. Also, the success of some learning schemes, in their attempts to construct models of data, hinges on the reliable identification of a small set of highly predictive attributes. The inclusion of irrelevant, redundant and noisy attributes in the model building process phase can result in poor predictive performance and increased computation.

Feature subset selectors are algorithms that attempt to identify and remove as much irrelevant and redundant information as possible prior to learning or knowledge discovery. Feature subset selection can result in enhanced performance, a reduced hypothesis search space, and, in some cases, reduced storage requirement. Automated techniques for identifying and removing unhelpful or redundant variables usually take one of two forms: direct examination of the relevance of candidate variables, or searching the best combination of attributes in terms of model performance and feedback. The former are called *filters* and the latter *wrappers* (see Hall 1999 for details). For a survey of common

feature selection techniques, see Molina et al. (2002). Usually, analyzing the feature subset selection provides better results than analyzing the complete set of variables (Andrew 1998).

However, attribute selection generally involves a combination of search and attribute utility estimation plus evaluation with respect to specific learning schemes. This leads to a large number of possible permutations where very few benchmark studies have been conducted. In (Andrew et al. 2003) a benchmark comparison of several attribute selection methods for supervised classification is presented. All the methods produce an attribute ranking, a useful devise for isolating the individual merit of an attribute. Attribute selection is achieved by cross-validating the attribute rankings with respect to a classification learner to find the best attributes. Results are reported for a selection of standard data sets and two diverse learning schemes C4.5 and naive Bayes.

Other techniques are based on *feature weighting* (see for example Aha 1998 and Núñez et al. 2003), which is a more general and flexible approach than feature selection. The aim is to assign a degree of relevance (a weight) to each attribute. Similarities (or dissimilarities) become emphasized according to the relevance of the attribute, and irrelevant attributes will not influence the results, so quality of inductive learning improves.

Feature weight assignment is frequently used to denote the relevance of attributes in similarity computations. When some attributes are irrelevant for the prediction task, the appropriate weight learning could improve the data mining process [Aha, 1998]. Empiric works [Wettschereck et al., 1997] and theoretical ones [Langley and Iba, 1993], suggest that the learning complexity is exponential regarding the number of irrelevant attributes. Therefore, the failures in the data mining process could be related to a similarity model, and in particular, with an incorrect weight assignment methodology.

In recent years, a great deal of research research works has been done in feature weight assignment. The main goal is to assign high weights to attributes that are identified as relevant, and at the same time, to assign low weights to those that are irrelevant. Most of the methods of weight assignment use a global scheme, that is, they associate a weight to the whole space of the attribute. In [Wettschereck et al., 1997] a conceptual framework for the weight assignment methods classification is presented, considering bias from the performance algorithm (wrapper) [Kohavi and John, 1998] or not (filter), previous transformations of data before analyzing relevance, if the relevance of an attribute is invariant or not all along the domain, or if they require domain specific knowledge. A comprehensive review of feature weighting algorithms can be found in [Núñez, 2004].

Although the weight assignment improves the accuracy in classification and retrieval tasks, feature selection is vital to reduce the dimensionality in learning tasks, completely eliminating irrelevant attributes (Martorell et al.2007). In general, feature weighting is more appropriate for tasks where features vary in their relevance, but such methods search larger spaces of weight assignments. On the other hand, feature selection algorithms perform better when the features used to describe instances are either highly correlated with the class label or completely irrelevant.

2.8. Hybrid approach

Frequently, data-driven (either stochastic or connectionist, among others) models are accused of lack of understandability (black boxes) because of the substantial drawback they exhibit: they are synthesized ONLY on the available data, with no detailed information of the underlying process. As a consequence, in prediction and extrapolation tasks and especially when data are noisy or sparse, they sometimes are inadequate and inaccurate. It is thus reasonable to argue that DM techniques should be used not only to replace

knowledge-based models but also to complement them and produce so-called hybrid models. Integration of both deterministic (numerical) and stochastic (data-driven), or logic (qualitative) models presumably should provide more accurate predictions and cope better with uncertainty, non-normality, non-linearity, or even with the original nature of data. It appears clearly more sensible to use the already available deterministic information given by the very well-known theory-driven models than stubbornly starting from the scratch, throwing away all knowledge, trying to use a data-driven model alone. Also, it seems better to take advantage of the precision of a numerical variable than discretizing all numerical ones to use a machine learning method for qualitative variables. This approach should also appear more appealing to scientists that have been working for years with theory-driven models who will be reluctant to sink all their *precious* knowledge into oblivion. From a conceptual point of view, this approach takes into account that (environmental) complexity comes not only from the system but also from outer perturbations producing stochastic influences (noise). A number of papers using hybrid modeling in different areas, in particular environmental modeling, can be quoted (Espert et al. (1999), Vojinovic et al. (2003), Krasnopolsky et al. (2006), Chercassky et al. (2006), Izquierdo et al. (2006), Izquierdo et al. (2008)). It can be foreseen that computational intelligence (machine learning) will be used not only for building data-driven models, but also for building optimal adaptive model structures of such hybrid models. Combining Artificial Intelligence techniques with Statistical ones can also improve the quality of discovered knowledge (Gibert et al.2005a, Pérez-Bonilla et al.2007, Gibert et al.2007).

3. THE ROLE OF POST PROCESSING

Apart from the important role of preprocessing, together with the correct selection of the data mining technique which will really answer the target questions, there is an important job to be done between getting the results of the data mining techniques and using them to support decision-making: to *understand* the results.

Indeed, the quality of decisions taken upon KDD results, depend not only on the quality of the results itself, but on the capacity of the system to communicate those results in an understandable form to the decision maker.

The software tools used for applying the DM techniques to a data set use to produce long listings plenty of results that may be useful in any particular application. The closer to statistical packages the software, the longer and more complex the output with more numerical information displayed. However, given a particular real case, not all this information is useful. Moreover, the major part is irrelevant. Thus, it is important to:

- Identify the relevant information from the software outputs, depending on the aims of every particular analysis. For example, from a regression analysis, it may be irrelevant to the environmental expert to know the exact value of *h_i* indexes, but from this information, the set of influential observations that have to be carefully analyzed should be reported.
- Find the best way to present the selected results to the user in such a way that it becomes directly understandable, given that the final user does not know the technical details of the Data Mining method used. So, probably, from the results of a logistic regression, it is more interesting to provide the interpretation of the estimated coefficients rather than the logistic equation itself.

As an example, consider an application where clustering is applied as the most suitable data mining technique, regarding the goals of the analysis. Most of the software implementing clustering algorithms provides information about the number of clusters discovered, and the set of objects belonging to every one of the clusters. Upon these results, the experts need to understand the underlying clustering criteria as well as the meaning of the classes themselves. In data mining contexts where the number of classes increases, and the number

of variables is high, tools that help the user to postprocess the clustering results till the conceptualization of the classes are very useful. In (Gibert et al.2005b) the *Class Panel Graph (CPG)* is presented as an integrated graphical tool that provides a perspective of the whole data set regarding the previously discovered classes, in such a way that identification of variables with particular behaviours in every class is easy. Figure 2 shows the aspect of a CPG which can display either histograms conditioned to the classes, or boxplots (and bar charts for qualitative variables).

The CPG supports the interpretation of the results as well as the process of conceptualizing the classes. The authors are not aware that commercial software offer facilities to show in a single integrated graph the behaviour of as many variables as possible. Also, in Pérez-Bonilla (2007) a methodology that automatically induces concepts from classes is presented (CCCS) in such a way that the interpretation of classes performed by the expert becomes easier on the basis of some preliminary concepts suggested by the system. Either the CPG as the CCCS methodology are implemented in the KLASS software, which will be presented below. Also, in a clustering context, it may be useful to display the prototypes of the classes in a 2-D or 3-D scatterplot of numerical variables (Gibert et al.2006), as implemented in the GESCONDA software, also described below.

As another case, let us consider here the classification tasks. Among the algorithms solving those taks, those that inductively construct decision trees are found. Other approaches using neural networks, evolutionary algorithms or Bayesian networks are also applied to solve the classification problem. Most of these algorithms work as black box approaches. As a result, it is often difficult to understand the decision model.

Visualization techniques can help to overcome these problems. One of them, for example used by SGIs MineSet sytem (Brunk et al., 1997), shows a scheme of the decision tree and allows the user to select among important parameters of the model. The user interactively can select different attributes and understand the model. In (Ankerst et al., 2000) a more sophisticated approach is used: each attribute value is shown by a colored pixel and all of them are arranged in bars; then the pixels corresponding to each attribute bar are sorted separately and the attribute exhibiting the purest value distribution is chosen to split the decision tree; the process is iteratively repeated until all leaves correspond to pure classes. In addition, information like number of training patterns corresponding to one node, purity of partitions, etc. can complete the picture.

In general, many tools are available to post process the results of many different data mining techniques, such as decision trees, neural nets, statistical modelling, etc. This paper do not pretends to be exhautive, but to provide an overview of the importance of inserting this kind of tools inside the methodology to make knowledge discovery more fruitful.

4.1. Weka

The Weka workbench (Witten and Frank 2005) contains a collection of visualization tools and algorithms for data analysis and predictive modelling, together with graphical user interfaces for easy access to this functionality. A command line interface is also included, for larger scale processing. It was originally designed as a tool for analyzing data from agricultural domains but is now used in many different application areas, largely for educational purposes and research. The main strengths of Weka are that it is (a) freely available under the GNU General Public License, (b) very portable because it is fully implemented in the Java programming language and thus runs on almost any computing platform, (c) contains a comprehensive collection of data preprocessing and modelling techniques, and (d) is easy to use by a novice due to the graphical user interfaces it contains.

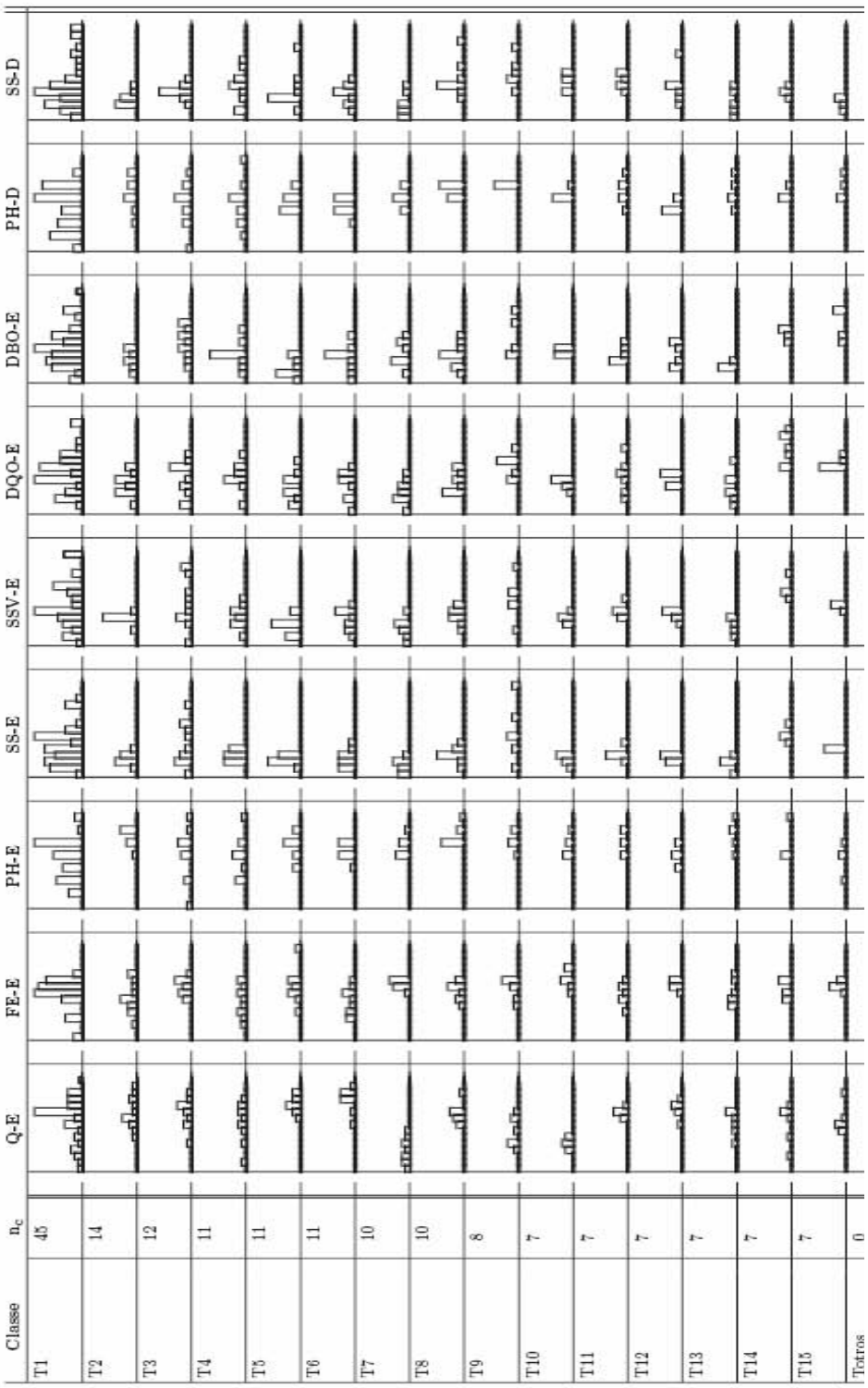


Figure 2: Example of CPG where it can be seen that group T7 contains elements with lower values on variable Q-E (inflow of a WWTP), group T15 contains elements with greater values of DQO-E (chemical organic matter).

Pre-processing in Weka is mainly achieved with filters. The idea is to pass data through the filter in order to affect some form of transformation. Two general types of filter are provided, supervised and unsupervised. Supervised filters use class information to affect the transformation while unsupervised filters are class-blind. Beyond that filters work on data either at the instance (ie example level) or the attribute level. Thus there are four types of filter:

Supervised attribute: providing operations for selecting the attributes more correlated with a class (Attribute Selection) or some coercions (Discretize, NominalToBinary) or reordering (ClassOrder). Also PLSfilter is a specialised filter particularly useful in spectral applications where attribute co-linearity is expected.

Supervised instance: These filters mainly support experimentation or down-sampling of a problem (Resample, SpreadSubsample or StratifiesRemoveFolds). They work on the entire set of instances in a dataset.

Unsupervised attribute: The bulk of the available filters in Weka are of this type, including operations to add columns to the data set, either with a new attribute (Add, Copy, ClusterMembership, FirstOrder), an identifier (AddID) or a new partition (AddCluster) or combining existing attributes (AddExpression), or transforming them (Normalize, Standardize, Discretize, PKDiscretize, MergeTwoValues, NumericToBinary, NumericToNominal, NominalToBinary, NumericTransform, StringtoNominal, StringToWordVector, Center, Wavelet, ChangeDateFormat, SwapValues, MathExpression, Obfuscate), to change some of the rows of a given column (NumericCleaner –for extreme values, AddNoise, AddValues, ReplaceMissingValues) or to mark some rows of a given column (Interquartile Range). Also, operations to eliminate some columns of the data set (Remove, RemoveType, RemoveUseless) or operations for creating derivative datasets from the original one (KernelFilter, MakeIndicator, PartitioneMultifilter, RandomProjection, TimeSeriesDelta, TimeSeriesTranslate,)

Unsupervised instance: Including filters to Normalize variables, transform to a sparse format NonSparseToSparse, produce random subsamples with or without replacement Resample, randomly reorder the set of instances (Randomize), or to remove subsets of instances under different criteria (RemoveFrequentValues of a nominal attribute, RemovePercentage, RemoveRange, RemoveWithValues, RemoveFolds)

4.2. Preprocessing with SPSS

Pre-processing techniques in SPSS range from plain detection and correction of data errors made at the introduction stage, to sophisticated transformations, including re-codification performed on one or several variables or the creation of new variables from other already existing ones. SPSS provides the nice possibility of apply transformations and re-codifications either to the whole set of data or to a predefined subset (by using the **select cases** option) Also, sometimes it is necessary to change the order, to merge different files or to select specific cases for analysis. Another need, which is frequent in the case of working with real data bases, comes from the fact that files are not suitably organized. SPSS provides tools for all these tasks.

The menu **Transform**, in the SPSS's main menu bar, includes a number of options allowing a number of transformations:

- **Computing variables** allows calculating values for a variable based on numeric transformations of other variables.
- **Count Occurrences** creates a variable that counts the occurrences of the same value(s) in a list of variables.

Recoding values allows data values modifications by recoding them. It is particularly useful for collapsing or combining categories. Values can be recoded into existing variables or, alternatively, new variables can be created to recode values of existing variables.

Rank cases allows creating new variables containing ranks, normal scores and percentile values for numeric variables. The **Automatic Recode** option allows converting string and numeric values into consecutive integer numbers.

Missing values in functions make it possible to treat missing values in different ways.

Data files are not always organized in the ideal form for some specific needs. A wide range of file handling and transformation capabilities is available. Among them:

Sort cases that sorts rows based on the value or one or more variables.

Transpose cases and variables creates a new data file with rows into columns and vice versa.

Aggregate data aggregates groups of cases in the active dataset into single cases and creates a new file or new variables in the active data set.

Select subsets of cases allows the analysis to be restricted to a subset of cases or perform simultaneous analyses on different subsets.

Weight data weights cases for analyses based on the value of a weight variable. It is particularly useful to build contingency matrices.

Restructure data can create a single case from multiple cases or create multiple cases from a single one.

Some other additional data preparation features can be found in the **Data Preparation** (*Data Validation* in older versions) module. It allows creating reports, charts and analyses without additional preliminary work. Among other tasks, one can determine how certain values should be treated by assigning variable properties that describe the data, identify anomalies such as missing values, outliers, duplicate information, or create new variables with a lower number of categories to represent ranges of values from variables with larger number of possible values.

4.3. Preprocessing and postprocessing in GESCONDA

GESCONDA (Gibert et al.2006) is the name given to an Intelligent Data Analysis System developed with the aim of facilitating Knowledge Discovery (KD) and especially oriented to environmental databases. On the basis of previous experiences, it was designed to include extensive preprocessing tools: data cleaning, missing data management, outlier analysis and treatment, statistical univariate analysis, statistical bivariate analysis, visualization tools, attribute or variable transformation facility, including discretizations, recodifications and creations of new variables, feature weighting for supervised and unsupervised data sets (Gibert et al.2006b). Portability of the software between platforms is provided by a common Java platform.

GESCONDA contains a postprocessing module including tools for validation of the results of different data mining techniques, such as results from clustering (scatter plot of pairs of numerical variables, marked by class and superimposing the position of class prototypes, rates of missclassification if a reference partition is available), rule induction (eliminating low precision rules, evaluate an inducted rules base over a test set and estimate the quality of the rule base), etc.

integration of different knowledge patterns for a predictive task, or planning, or system supervision, together with AI and statistics mixed techniques, consideration of knowledge use by end-users.

4.4. Pre and post processing in KLASS

KLASS (Gibert et al.2005b) is a software package originally conceived for Knowledge Discovery (KD) in real domains with complex structure (Gibert et al.1999). It provides a mixture of statistical and artificial intelligence tools to support KD, including basic statistics and providing an integrated system to support the whole process of KDD including pre and post processing, provided that the main data mining technique to be used is related with clustering or rule induction.

Regarding preprocessing, KLASS offers functionalities for basic statistics (simple or by groups), histograms, boxplots (side-by-side), (letter)plots, cross-tables. The performance of the system is quite high, since the user has control over many parameters of the graphics (like the number of classes of a histogram, or the limits of the axis of a plot), providing a very flexible tool. It also offers a complete module of data management, including missing data treatment, creation of transformed variables either using mathematical expressions or via recodification or discretization (here the Boxplot based discretization is provided, which discretizes the numerical variable in such a way that the resulting qualitative variable maximizes association degree with a previously discovered class variable, Pérez-Bonilla et al.2007). Construction of a prior expert knowledge base (which can be non-complete) is also available and it can be used to bias a posterior clustering process, by means of the Clustering Based on Rules option (Gibert et al.1999), in such a way that the final classes hold the semantic constraints expressed by the rules.

Regarding the postprocessing, Klass offers some interesting tools to support the interpretation of a clustering results, apart from the classical representation of the dendrogram; It also provides the Class Panel Graph (Gibert et al.2005), which is a very interesting possibility in clustering contexts to understand better the meaning of the classes. It also implements the CCCS methodology (Pérez-Bonilla 2007) for assigning concepts to every class, improving even more the support to the understanding of the results. There is also a function for visualizing knowledge bases, containing probabilities or not, and selecting the rules with degrees of certainty over a certain threshold.

One of the particularities of the system is that it is designed in such a way that the outputs, either graphical or numerical or textual, are produced in LaTeX font files, which are directly processed by the kernel of KLASS and automatically sent to the LaTeX viewer and displayed on the screen. From the final user point of view, this makes no difference with other systems, since graphical representations are directly displayed on the screen as well as other results. However, as reporting the results of the KDD process is always involved with the elaboration of technical papers, KLASS also includes a reporting facility in such a way that the user can specify a set of steps to be performed sequentially and a single big LaTeX document including all the results is produced. The user only needs to edit this document and add personal comments on it to get a complete report of the analysis. KLASS provides either standard or personalized reports. It is a flexible possibility since it is possible to automatically transform every result of the single steps into PostScript or PDF documents, which can be managed as usual, for example, pasting it into a Word document.

If the document to be produced is long, with a complex structure and contains hard mathematical notation, LaTeX offers nice advantages and the LaTeX results provided by KLASS are really useful. In this case, LaTeX is a widely used text processor, owing to the excellent support it provides to the generation of high quality mathematical formulae and scientific notation. However, including graphical representations from commercial statistical packages in a LaTeX document requires the use of special LaTeX packages to deal with graphical formats and makes a little bit more complicate the elaboration of the document, which requires transformation to PostScript or PDF to be completely visualized. Since the results of KLASS are produced in native LaTex code, inclusion of those graphics in the final report becomes trivial.

On the other hand, making the native LaTeX code accessible to the user permits the user to adjust labels or size of the titles of graphical representations. In this way, the quality of the image is maintained to its final use.

ACKNOWLEDGEMENTS

The project TIN2004-01368 has partially financed the development of GESCONDA.

REFERENCES

Aha, D. 1998. Feature weighting for lazy learning algorithms. In: Liu, H., Motoda, H. (Eds.), *Feature Extraction, Construction and Selection: A Data Mining Perspective*. Kluwer.

Allison, P. 2002. *Missing Data*. Sage, Thousand Oaks, CA, USA.

Almasri, M. and Kaluarachchi, J. 2005. Modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrogen loading and recharge data. *Environmental Modelling and Software*, 20(7): 851-871.

Mark Andrew Hall and Geoffrey Holmes. Benchmarking attribute selection techniques for discrete class data mining. IEEE Transactions on Knowledge and Data Engineering, 15(3):1437-1447, November/December 2003.

Mark Andrew Hall and Lloyd Smith (1998): Practical feature subset selection for machine learning. In Proc 21st Australian Computer Science Conference, pages 181-191, Perth, Australia. Springer.

Andrienko, G. and Andrienko, A. 2004. Research on visual analysis of spatio-temporal data at fraunhofer ais: an overview of history and functionality of commonGIS. In: *Proceedings of the Knowledge-Based Services for the Public Services Symposium, Workshop III: Knowledge Discovery for Environmental Managment*. KDnet, pp. 26-31.

Ankerst, M., Ester, M., and Kriegel, H. (2000). Towards an effective cooperation of the computer and the user for classification. In *SIGKDD Int. Conf. on Knowledge Discovery & Data Mining (KDD 2000)*, Boston, MA, 179-188.

Athanasiadis, I., Kaburlasos, V., Mitkas, P. and Petridis, V. 2003. Applying machine learning techniques on air quality for real-time descision support. In: *Information technologies in Environmental Engineering*.

Athanasiadis, I., Karatzas, K. and Mitkas, P. 2005. Contemporary air quality forecasting methods: A comparative analysis between statistical methods and classification algorithms. In: *Proceedings of the 5th International Conference on Urban Air Quality*.

Athanasiadis, I. and Mitkas, P. 2004. Supporting the decision-making process in environmental monitoring systems with knowledge discovery techniques. In: *Proceedings of the Knowledge-Based Services for the Public Services Symposium, Workshop III: Knowledge Discovery for Environmental Managment*. KDnet, pp. 1-12.

Athanasiadis, I.N. & Mitkas, P.A. (2007). Knowledge discovery for operational decision support in air quality management. Journal of Environmental Informatics. 9:2, 100-107

Babovic, V. 2005. Data mining in hydrology. *Hydrological Processes*, 19:1511-1515.

Barnett, V. and Lewis, T. 1978. *Outliers in Statistical Data*. Wiley.

Batista, G., Prati, R. and Monard, M. A study of the behaviour of several methods for balancing machine learning training data, *SIGKDD Explorations* 6 (1), 20-29, 2004.

Bazartseren, B. Hildebrandt, G. and Holz, K. P. Short-term water level prediction using neural networks and neuro-fuzzy approach. *Neurocomputing*, 55, 439-450, 2003.

Belanche, L., Valdés, J., Comas, J., Rodríguez-Roda, I. and Poch, M. 2001. Towards a model of input-output behaviour of wastewater treatment plants using soft computing techniques. *Environmental Modelling and Software*, 5(14): 409-419.

Brunk, C., Kelly, J., and Kohavi, R. (1997). Mineset: An integrated system for data mining. In *SIGKDD Int. Conf. on Knowledge Discovery & Data Mining (KDD 1997)*, Newport Beach, CA, 135-138.

Cadez, I. and Smyth, P. 1999. Modelling of inhomogeneous markov random fields with applications to cloud screening. Tech. Rep. UCI-ICS 98-21.

Camargo, S., Robertson, A., Gaffney, S. and Smyth, P. 2004. Cluster analysis of western north pacific tropical cyclone tracks. In: *Proceedings of the 26th Conference on Hurricanes and tropical Meteorology*. pp. 250-251.

Cendrowska, J. 1998. Prism: An algorithm for inducing modular rules. *International Journal of Man-Machine Studies*, 27(4): 349-370.

Chawla, N., Bowyer, K., Hall, L. and Kegelmeyer, W. Smote: Synthetic minority over-sampling technique, *Journal of Artificial Intelligent Research* 16, 321-357, 2002.

Chercassky, V., Krasnopolsky, V., Solomatine, D. P., Valdes, J. 2006. Computational intelligence in earth sciences and environmental applications: Issues and challenges. Neural Networks 19, 113-121.

Cohen, W. 1995. Fast effective rule induction. In: *Proceedings of the Twelfth International Conference on Machine Learning*, Prieditis, A., Russell, S. (Eds.). Morgan Kaufmann, pp. 115-123.

Comas, J., Dzeroski, S., Gibert, K., Rodríguez-Roda, I and Sànchez-Marrè, M. 2001. Knowledge discovery by means of inductive methods in wastewater treatment plant data. *AI Communications*, 14(1): 45-62.

Comas, J., Llorens, E., Martí, E., Puig, M.A., Riera, J.L., Sabater, F. and Poch, M. 2003. Knowledge Acquisition in the STREAMES Project: The Key Process in the Environmental Decision Support System Development. *AI Communications*, 16(4): 253-265.

Cortés, U., Rodríguez-Roda, I., Sànchez-Marrè, M., Comas, J., Cortés, C. and Poch, M. 2002. DAI-DEPUR: An environmental decision support system for supervision of municipal waste water treatment plants. In: *Proceedings of the 15th European Conference on Artificial Intelligence (ECAI'2002)*. pp. 603-607.

Dillon, W. and Goldstein, M. 1984. *Multivariate Analysis*. Wiley, USA.

Dixon, M., Gallop, J.R., Lambert, S.C. and Healy, J.V. 2007. Experience with data mining for the anaerobic wastewater treatment process. *Environmental Modelling and Software*, 22: 315-322.

Domingos, P. 1996. Unifying Instance-Based and Rule-Based Induction. *Machine Learning*, 24: 141−168.

Draper, N. and Smith, H. 1998. *Applied Regression Analysis*. Wiley.

Dubes, R. and Jain, A. 1988. *Algorithms for Clustering Data*. Prentice Hall.

Dzeroski, S. and Drumm, D. 2003. Using regression trees to identify the habitat preference of the sea cucumber (holothuria leucospilota) on rarotonga, cook islands. *Ecological Modelling*, 170(2-3): 219-226.

Dzeroski, S., Grbovic, J., Walley, W. and Kompare, B. 1997. Using machine learning techniques in the construction of models. ii. data analysis with rule induction. *Ecological Modelling*, 95(1): 95-111.

Ekasingh, B., Ngamsomsuke, K., Letcher, R. and Spate, J. 2003. A data mining approach to simulating land use decisions: Modelling farmer's crop choice from farm level data for integrated water resource management. In: *Advances in Hydrology: Proceedings of the International Conference on Water and Environment*, Singh, V., Yadava, R. (Eds.), pp. 175-188.

Ekasingh, B., Ngamsomsuke, K., Letcher, R. and Spate, J. 2005. A data mining approach to simulating land use decisions: Modelling farmer's crop choice from farm level data for integrated water resource management. *Journal of Environmental Management*.

Ellis, F. 1996. The application of machine learning techniques to erosion modelling. In: *Proceedings of the Third International Conference on Integrating GIS and Environmental modelling*. National Center for Geographic Information and Analysis.

Espert, V., López, P. A. and Izquierdo, J., 1999. Fundamentals of a water quality model solution for dissolved oxygen in one-dimensional receiving system. Numerical Modelling of Hydrodynamic Systems. Proc. of Intnl. Workshop, 444-445.

Faye, R. M., Sawadogo, S., Lishou, C. and Mora-Camino, F. Long-term fuzzy management of water resource systems. *Applied Mathematics and Computation* 37: 459-475, 2003.

Fayyad, U. and Piatetsky-Shapiro, G., P, S. 1996a. Advances in knowledge discovery and data mining. In: *Data Mining to Knowledge Discovery: an Overview*. American Association for Artificial Intelligence, pp. 1-34.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. 1996b. From data mining to knowledge discovery in databases (a survey). *AI Magazine*, 3(17): 37-54.

Fernández, A., García, S., del Jesus, M.J. and Herrera, F. A Study of the Behaviour of Linguistic Fuzzy Rule Based Classification Systems in the Framework of Imbalanced Data-Sets. *Fuzzy Sets and Systems*, doi:10.1016/j.fss.2007.12.023, 2007

Frank E, Witten I A (1999): Making better use of global discretization. In Proc 16th International Conference on Machine Learning, Bled, Slovenia, pages 115-123. Morgan Kaufmann.

Gatts, C., Ovalle, A. and Silva, C. 2005. Neural pattern recognition and multivariate data: Water typology of the Paraiba do Sul River, Brazil. *Environmental Modelling and Software*, 20(7): 883-889.

Gibbs, M., Morgan, N., Maier, H., Dandy, GC, H. M. and Nixon, J. 2003. Use of artificial neural networks for modelling chlorine residuals in water distribution systems. In: *MODSIM 2003: Proceedings of the 2003 International Congress on Modelling and Simulation*. pp. 789-794.

Gibert, K., Spate, J., Sànchez-Marrè, M., Comas, J., Athanasiadis, I. 2008: Data Mining for Environmental Systems, In State of the art and Futures in Environmental Modelling and Software. IDEA Series (Jackeman, A. J., Rizzoli, A., Voinov, A. and Chen, S. (eds), (in press). Elsevier, Amsterdam, The Netherlands.

Gibert, K., Sanchez-Marre, M. and Rodriguez-Roda, I. 2006. GESCONDA: an intelligent data analysis system for knowledge discovery and management in environmental databases. *Environmental Modelling and Software*, 21:115-120.

Gibert, K., Annicchiarico, R., Cortés, U. and Caltagirone, C. 2005a. *Knowledge Discovery on Functional Disabilities: Clustering Based on Rules Versus Other Approaches*. IOS Press.

Gibert, K., Flores, X., Rodríguez-Roda, I. and Sànchez-Marrè, M. 2004. Knowledge discovery in environmental data bases using GESCONDA. In: *Proceedings of IEMSS 2004: International Environmental Modelling and Software Society Conference*, Osnabruck, Germany.

Gibert, K., Nonell, R., Velarde, JM, Colillas, MM 2005b. Knowledge discovery with clustering: Impact of metrics and reporting phase by using klass. *Neural Network World*, 319-326.

Gibert, K, Rodríguez-Silva, G (2007) Knowledge Discovery in a Wastewater Treatment Plant with Clustering based on Rules by States. In Artificial Intelligence Research and Development. Frontiers in Artificial Intelligence and Applications, v 163, pp 359—367. IOS Press.

Gibert, K, Sànchez-Marrè, M, Comas, J. (2006b) The impact of feature weighting in environmental class discovery using GESCONDA, In BESAI proceedings, ECAI 2006.

Gibert, K., Sànchez Marrè, M. and Flores, X. 2005c. Cluster discovery in environmental databases using GESCONDA: The added value of comparisons. *AI Communications*, 4(18): 319-331.

Gibert, K. and Sonicki, Z. 1999. Clustering based on rules and medical research. *Journal on Applied Stochastic Models in Business and Industry*, formerly JASMDA 15(4): 319-324.

Gibert K, Saxena, K. Morris, J., et al. 2008: The advantages of using clustering for missing inputation, World Health Organization (in press).

Grzymala-Busse, J.W., Goodwin, L.K. and Zhang, X. Increasing sensitivity of preterm birth by changing rule strengths, *Pattern Recognition Letters* 24 (6), 903-910, 2003.

Guo, H. and Viktor, H. L. Learning from imbalanced data sets with boosting and data generation: The databoosting approach, *SIGKDD Explorations* 6 (1), 30-39, 2004.

Guo, Q., Kelly, M. and Graham, C. 2005. Support vector machines for predicting distribution of sudden oak death in california. *Ecological Modelling*, 182(1): 75-90.

Hall, M. 1999. *Feature selection for discrete and numeric class machine learning*. Tech. rep., Department of Computer Science, University of Waikato, working Paper 99/4. URL http://www.cs.waikato.ac.nz/~ml/publications1999.html

Han, J. and Kamber, M. 2001. *Data Mining: Concepts and Techniques*. Morgan Kaufmann.

Hastie, T., Tibshirani, R. and Friedman, J. 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag.

Holmes, G., Cunningham, S., Dela Rue, B. and Bollen, A. 1998. Predicting apple bruising using machine learning. In: Proceedings of the Model-IT Conference, *Acta Horticulturae*, 476: 289-296.

Izquierdo, J., Pérez, R., López, P. A., Iglesias, P. L., 2006. Neural Identification of Fuzzy Anomalies in Pressurized Water Systems, Summit on Environmental Modelling and Software, 3rd Biennial meeting of the International Environmental Modelling and Software Society, Proceedings, Burlington (Vt), USA.

Izquierdo, J., López, P.A., Martínez, F.J. and Pérez, R., 2007. Fault detection in water supply systems using hybrid (theory and data-driven) modelling, *Mathematical and Computer Modelling*, 46, 3-4, 341-350.

Izquierdo, J., Montalvo, I., Pérez, R. and Herrera, M., 2008. Sensitivity analysis to assess the relative importance of pipes in water distribution networks, *Mathematical and Computer Modelling*, 48, 268–278.

Japkowicz, N. and Stephen, S. The class imbalance problem: A systematic study, *Intelligent Data Analysis* 6 (5), 429-450, 2002.

Juang, Ch. F. Temporal problems solved by dynamic fuzzy network based on genetic algorithm with variable-length chromosomes. *Fuzzy Sets and Systems* 142: 199-219, 2003.

Kolodner, J. 1993. *Case-Based Reasoning*. Morgan Kaufmann.

Kohavi and John, 1998 R. Kohavi, and G-H. John, The Wrapper Approach, in Feature Selection for Knowledge Discovery and Data Mining, H. Liu & H. Motoda (eds.), Kluwer Academic Publishers, pp33-50. 1998

Kralisch, S., Fink, M., Flügel, W.-A. and Beckstein, C. 2001. Using neural network techniques to optimize agricultural land management for minimisation of nitrogen loading. In: *MODSIM 2001: Proceedings of the 2001 International Congress on Modelling and Simulation*. pp. 203-208.

Krasnopolsky, V.; Fox-Rabinovitz, M. S., 2006. Complex hybrid models combining deterministic and machine learning components for numerical climate modeling and weather prediction. Neural Networks, 19, 122-134

Kubat, M., Holte, R. and Matwin, S. Machine learning for the detection of oil spills in satellite radar images, *Machine Learning* 30 (2-3), 195-215, 1998.

Kuncheva, L. I., Wrench, J., Jain, L. C. and Al-Zaidan, A. S. A fuzzy model of heavy metal loadings in Liverpool bay. Env. *Modeling and Software* 15: 161-167, 2000.

Langley, P., and Iba, W. (1993). Average-case analysis of a nearest neighbor algorithm. Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (pp. 889-894). Chambery, France.

Larose, D. 2004. *Discovering Knowledge in Data: An Introduction to Data Mining.* John Wiley.

Lebart, L., Morineau, A. and Warwick, K. 1984. *Multivariate Descriptive Statistical Analysis.* Wiley, New York, USA.

Little, R. and Rubin, D. 1987. *Statistical Analysis with Missing Data*. Wiley.

Martín, F.J. and Plaza, E. 2004. Ceaseless Case-based Reasoning. In: *Procc. of 7th European Conference on Case-Based Reasoning (ECCBR'2004)*, pp. 287-301, LNAI-3155, Madrid, Spain.

X. L. Martorell, R. Massanet, K. Gibert, M. Sánchez-Marrè, J. C. Martín-Sánchez, A. Martorell (2007) A Knowledge Discovery methodology for identifying vulnerability factors of mental disorder in an intellectually disabled population. In Artificial Intelligence Research and Development. Frontiers in Artificial Intelligence and Applications, v 163, pp 426—435. IOS Press

Mas, J., Puig, H., Palacio, J. and Sosa-Lopez, A. 2004. Modelling deforestation using GIS and artificial neural networks. *Environmental Modelling and Software*, 19(5): 461-471.

McKenzie, N. and Ryan, P. 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma* (89): 67-94.

Michalski, R. and Chilausky, R. 1980. Learning by being told and learning by examples: An experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis. *International Journal of Policy Analysis and Information Systems*, 4(2): 125-161.

Miller, H. J. (in press, to appear in 2007). Geographic data mining and knowledge discovery. In: *Handbook of Geographic Information Science*, Wilson, J. P., Fotheringham, A. S. (Eds.). Blackwell Publishing.

Molina, L., Belanche, L. and Nebot, A. 2002. Feature selection algorithms: A survey and experimental evaluation. In: *ICDM 2002: Proceedings of the IEEE International Conference on Data Mining*. pp. 306-313.

Moore, D. and McCabe, G. 1993. *Introduction to the practice of statistics*. WH Freeman, New York, second edition.

Mora-López, L. and Conejo, R. 1998. Qualitative reasoning model for the prediction of climatic data. In: *ECAI 1998: Proceedings of the 13th European Conference on Artificial Intelligence.* pp. 61-75.

Núñez, H. 2004. Feature Weighting in Plain Case-Based Reasoning. Ph.D. Thesis. Universitat Politècnica de Catalunya.

Núñez, H., Sànchez-Marrè, M., Cortés, U., Comas, J., Martinez, M., Rodríguez-Roda, I. and Poch, M. 2004. A comparative study on the use of similarity measures in case-based reasoning to improve the classification of environmental system situations. *Environmental Modelling and Software*, 19(9): 809-819.

Núñez, H., Sànchez-Marrè, M. and Cortés, U. 2003. Improving similarity assessment with entropy-based local weighting. In: *Lecture Notes in Artificial Intelligence, (LNAI-2689): Proceedings of the 5th International Conference on Case-Based Reasoning (ICCBR2003)*. Springer-Verlag, pp. 377-391.

Oh, S. K. and Pedrycz, W. Self-organizing polynomial neural networks based on polynomial and fuzzy polynomial neurons: analysis and design. *Fuzzy Sets and Systems* 142, 163-198, 2004.

Parr Rud, O. 2001. *Data Mining Cookbook- Modelling data for marketing, risk, and CRM*. Wiley.

Pérez-Bonilla A, Gibert, K (2007) Automatic generation of conceptual interpretation of clustering. In Progress in Pattern Recognition, Image analysis and Applications.Lecture Notes in Computer Science v 4756, pp 653-663. Springer.

Poch, M., Comas, J., Rodríguez-Roda, I., Sànchez-Marrè, M. and Cortés, U. 2004. Designing and building real environmental decision support systems. *Environmental Modelling and Software*, (19): 857-873.

Provost, F. and Fawcett, T. Robust classification for imprecise environments, *Machine Learning* 42 (3), 203-231, 2001.

Recknagel, F. 2001. Applications of machine learning to ecological modelling. *Ecological Modelling*, 146(1-3): 303-310.

Riaño, D. 1998. Learning rules within the framework of environmental sciences. In: *ECAI 1998: Proceedings of the 13th European Conference on Artificial Intelligence.* pp. 151-165.

Robertson, A., Kirshner, S. and Smyth, P. 2003. Hidden markov models for modelling daily rainfall occurrence over brazil. Tech. Rep. UCI-ICS 03-27. URL http://www.datalab.uci.edu/papers-by-date.html

Rodríguez-Roda, I., Comas, J., Colprim, J., Poch, M., Sànchez-Marrè, M., Cortés, U., Baeza, J. and Lafuente, J. 2002. A hybrid supervisory system to support wastewater treatment plant operation: Implementation and validation. *Water Science and Technology*, 45(4-5): 289-297.

Rodríguez-Roda, I., Comas, J., Poch, M., Sànchez-Marrè, M. and Cortés, U. 2001. Automatic knowledge acquisition from complex processes for the development of knowledge based systems. *Industrial and Engineering Chemistry Research*, 15(40), 3353-3360.

Rodríguez-Roda, I., Poch, Sànchez-Marrè, M., Cortés, U. and Lafuente, J. 1999. Consider a case-based system for control of complex processes. *Chemical Engineering Progress*, 6(95): 39-48.

Rubin, D. 1987. *Multiple Imputation for Nonresponse in Surveys.* Wiley.

Sanborn, S. and Bledsoe, B. 2005. Predicting streamflow regime metrics for ungauaged streams in colorado, washington, and oregon. *Journal of Hydrology* (*under review*).

Siebes, A. 1996. *Data mining: What it is and how it is done*. In: SEBD. p. 329.

Smith, C. and Spate, J. 2005. Gully erosion in the ben chifley catchment (*in preparation*).

Sànchez-Marrè, M., Cortés, U., Martínez, M., Comas, J. and Rodríguez-Roda, I. 2005. An Approach for Temporal Case-Based Reasoning: Episode-Based Reasoning. In: *Procc. of 6th International Conference on Case-Based Reasoning (ICCBR'2005)*. LNAI-3620, pp. 465 – 476.

Sànchez-Marrè, M., Gibert, K. and Rodríguez-Roda, I. 2004. GESCONDA: A Tool for Knowledge Discovery and Data Mining in Environmental Databases. Vol. 11 of Research on Computing Science. Centro de Investigación en Computación, Instituto Politécnico Nacional, México DF, México, pp. 348-364.

Sànchez-Marrè, M., Cortés, U., Béjar, J., de Gracia, J., Lafuente, J. and Poch, M. 1997. Concept formation in wastewater treatment plants by means of classification techniques: a compared study. *Applied Intelligence*, 7(2):147-165 .

Stadler, M., Ahlers, D., Bekker, R.M., Finke, J., Kunzmann, D. and Sonnenschein, M. 2006. Web-based tools for data analysis and quality assurance on a life-history trait database of plants of Northwest Europe. *Environmental Modelling and Software*, 21:1536-1543.

Spate, J. 2002. *Data in hydrology: Existing uses and new approaches*. Australian National University, Master thesis.

Spate, J. 2005. Modelling the relationship between streamflow and electrical conductivity in Hollin Creek, southeastern Australia. In: *Proceedings of the 6th International Symposium on Intelligent Data Analysis*, Fazel Famili, A., Kok, J., Peña, J. (Eds.). pp. 419-440.

Spate, J. 2006. *Machine learning as a tool for investigating environmental systems*. Australian National University, PhD Thesis.

Spate, J., Croke, B. and Jakeman, A. 2003. Data mining in hydrology. In: *MODSIM 2003: Proceedings of the 2003 International Congress on Modelling and Simulation*. pp. 422-427.

Spate, J. and Jakeman, A., 2006 Review of data mining techniques and their application to environmental problems. *Environmental Modelling and Software* (*under review*).

Su, F., Zhou, C., Lyne, V., Du, Y. and Shi, W. 2004. A data-mining approach to determine the spatio-temporal relationship between environmental factors and fish distribution. *Ecological Modelling*, 174(4): 421-431.

Swayne, D., Cook, D. and Buja, A. 1998. Xgobi: Interactive dynamic data visualization in the x window system. *Journal of Computational and Graphical Statistics*, 7(1).

Sweeney, A., Beebe, N. and Cooper, R. 2007. Analysis of environmental factors influencing the range of anopheline mosquitoes in northern Australia using a genetic algorithm and data mining methods. *Ecological Modelling*, 203(3-4): 375-386.

Ter Braak, C., Hoijtink, H., Akkermans, W. and Verdonschot, P. 2003. Bayesian model-based cluster analysis of predicting macrofaunal communities. *Ecological Modelling*, 160(3): 235-248.

Vellido, A., Martí, J., Comas, I., Rodríguez-Roda, I. and Sabater, F. 2007. Exploring the ecological status of human altered streams through generative topographic mapping. Environmental Modelling and Software, 22(7): 1053-1065.

Vojinovic, Z.; Kecman, V; Babovic, V., 2003. Hybrid approach for modelling wet weather response in wastewater systems. J. of Water Resources, Planning and Mgment., 129(6), 511-521

Voss, H., Wachowicz, M., Dzeroski, S. and Lanza, A. (Eds.). 2004. *Knowledge Discovery for Environmental Management*. Knowledge-Based Services for the Public Sector Conference. KDnet.

Ward, J. 1963. *Hierarchical Grouping to Optimize an Objective Function*.

Weiss, G. and Provost, F. 2001. *The effect of class distribution on classier learning: An empirical study*. Tech. rep., Department of Computer Science, Rutgers University, technical Report ML-TR-44.

Weiss, G. and Provost, F. Learning when training data are costly: The effect of class distribution on tree induction, *Journal of Artificial Intelligence Research* 19, 315-354, 2003.

URL http://www.research.rutgers.edu/\~{}gweiss/papers/ml-tr-44.pdf

Witten I. H., Cunningham S., Holmes G., McQueen R. J., and Smith L.A. (1993): Practical Machine Learning and its Potential Application to Problems in Agriculture. In

Proc New Zealand Computer Conference, volume 1, pages 308-325, Auckland, New Zealand, 1993.

Whitten, I. and Frank, E. 1991. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann Publishers.

Wettschereck, D., D. W. Aha, and T. Mohri. 1997. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. *Artificial Intelligence Review*, Special Issue on lazy learning Algorithms.

Wong, I.W., Bloom, R., McNicol, D.K., Fong, P., Russell, R. and Chen, X. 2007. Species at risk: data and knowledge management within the WILDSPACETM Decision Support System. *Environmental Modelling and Software*, 22: 423-430.

Wnek, J. and Michalski, R. 1991. Hypothesis-driven constructive induction in aQ17: A method and experiments. In: *Proceedings of the IJCAI-91 Workshop on Evaluating and Changing Representation in Machine Learning*. pp. 13-22.

Xu, L., Chow, M. and Taylor, L. Power distribution fault cause identification with imbalanced data using the data mining-based fuzzy classification e-algorithm, *IEEE Transactions on Power Systems* 22 (1), 164-171, 2007.

Yeates, S. and Thomson, K. 1996. Applications of machine learning on two agricultural datasets. In: *Proceedings of the New Zealand Conference of Postgraduate Students in Engineering and Technology*. pp. 495-496.

Zadeh, L. Fuzzy sets. *Information and Control*, 8, 338-353, 1965.

Zadeh, L. Probability theory and fuzzy logic are complementary rather than competitive. *Technometrics*, 37(3), 271-276, 1995.

Zhu, X. and Simpson, A. 1996. Expert system for water treatment plant operation. *Journal of Environmental Engineering*, 822–-829.

Zoppou, C., Neilsen, O. and Zhang, L. 2002. Regionalization of daily stream flow in Australia using wavelets and k-means. Tech. rep., Australian National University, (http://wwwmaths.anu.edu.au/research.reports/mrr/mrr02.003/abs.html), accessed 15/10/2002.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Data mining using k-means clustering and classification and regression trees (CART) as post-processing methods: identifying management and environmental factors for explaining sugarcane yield in Northern Argentina (1971-2005)

D.O. Ferraro

Instituto de Investigaciones Fisiológicas y Ecológicas Vinculadas a la Agricultura (IFEVA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) / Facultad de Agronomía, Universidad de Buenos Aires, Av. San Martín 4453, C1417DSE Buenos Aires, Argentina.

***Keywords*:** modelling; crop yield; agriculture; sugarcane; Argentina.

Final crop yield in an agroecosystem is determined by agronomic management and environmental factors interactions. Gains in understanding the magnitude and nature of these interactions are keys for the design of efficient and productive systems. Traditionally, this knowledge is acquired for a particular area indirectly, by means of simulation growth models (Lisson, et al., 2005). These models deliberately reduce the system complexity for identifying the key aspects as well as to reduce the dataset required for parameterization. Alternatively, direct information of crop production can be stored in databases, which document what has actually happened in the farming systems, capturing large scale information on a wide range of variables that may potentially influence crop yield. However, the analysis of these large databases requires statistical methods capable of dealing with multivariate and nonlinear data structures.

Mathematical tools based on data mining provide an adequate framework for extracting useful information from large databases as well as they can also lead to the knowledge discovery for understanding environmental patterns and process. Recently, the use of learning algorithms of data mining has increased for analyzing small scale information (e.g hyper spectral remote sensing data) in order to provide crop management information for use in precision farming (Waheed, et al., 2006). However, the analyses of crop yield variability related to soil properties, agronomic practices or farmers' resource allocation decisions (e.g. agronomic factors) is usually carried out through traditional techniques such as linear regression and correlation.

This work proposes the use of two data mining techniques, k-means clustering (Jain and Dubes, 1988) and classification and regression trees (CART) (Breiman, et al., 1984) in order to identifying management and environmental factors for explaining sugarcane yield in Northern Argentina from a large database over the period of 1971 through 2005. The database contains information (i.e. field attributes) on the variety, crop age, area, farm of origin, and rainfall amount for every individual field of sugarcane harvested in each year. Sugarcane yield variable recorded was cane yield (TCH, t cane ha−1).

Firstly, TCH is analyzed through a k-means cluster analysis to group the sugarcane fields increasing cluster internal homogeneity and external or between-group heterogeneity. Then,

a classification tree partitions the space of all possible field attributes, starting with all field attributes (at the root of the tree) and successively splitting that space in subsets in which each subset is more likely to be assigned to one of the k-means clusters than the subset from which it is split. Also, the classification procedure is capable of provide information about the importance score (i.e. importance ranking) of the independent variables for explaining crop yield, which are ranked in descending order of their contribution to tree construction. The variable with the highest sum of improvements is scored 100, and all other variables have lower scores ranging downwards towards zero (Steinberg and Colla, 1995).

The methodology presented here is able to deal with large datasets in a robust way for detecting data patterns without the usual assumptions of data balance or distribution. Also, although CART are not inference tools, the information from this kind of studies could be useful to elaborate on testable hypothesis; pinpoint future mechanistic studies and designing more comprehensive management decisions for farming systems.

REFERENCES

Breiman, L., Friedman, R., Olshen, R., and Stone, C., *Classification and regression trees*, Pacific Grove, California, 1984.

Jain, A. K., and Dubes, R. C., *Algorithms for Clustering Data*, Prentice-Hall, Englewood Hills, NJ, 1988.

Lisson, S. N., Inman-Bamber, N. G., Robertson, M. J., and Keating, B. A., The historical and future contribution of crop physiology and modelling research to sugarcane production systems, *Field Crops Research, 92(2-3)*, 321-335, 2005.

Steinberg, D., and Colla, P., *CART: Tree-Structured Non-Parametric Data Analysis*, San Diego, 1995.

Waheed, T., Bonnell, R. B., Prasher, S. O., and Paulet, E., Measuring performance in precision agriculture: CART--A decision tree approach, *Agricultural Water Management, 84(1-2)*, 173-185, 2006.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Experiments in Auto-Calibration of Large Scale Environmental Models using a Data Mining Approach

David A. Swayne[1], Markiyan Sloboda[1], Simon Hood[1], Vimal Sharma[1], William Booty[2], David Lam[2]
[1]Computing Research Laboratory for the Environment
Department of Computing and Information Science, University of Guelph,
Guelph, Ontario, Canada
[2]National Water Research Institute, Burlington Ontario, Canada
{dswayne, vsharma, msloboda}@uoguelph.ca, shood01@guelphhumber.ca {Bill.Booty, David.Lam}@ec.gc.ca

Abstract: Many artificial intelligence paradigms are good candidates for the very difficult problem of model autocalibration. Quasi-Newton methods, optimization techniques and genetic algorithms are typical candidates for current algorithm research for this problem. A less well-travelled road to discovery consists of techniques for knowledge acquisition and data mining for autocalibration. This problem is largely intractable except by "expert" guessing, and the results are usually unsatisfactory. Our approach has been, from the environmental side, to use well-known and understood models, rather than develop our own. Many of these models involve large amounts of computer time for a single instance. We have partially overcome the computer time issue through the use of virtual files and running the codes on a clustered supercomputer architecture. These two preliminary steps have left us with a very typical knowledge discovery problem: several thousand model runs from which we hope to abstract the following: "good" solutions, knowledge of the trends toward these good solutions, and possibly a transformation of the model into (for instance) a Bayesian network approximation of the original model. This would allow us to "throw away" the computational model, instead selecting an output set which is a good approximation to what the model would produce under ideal circumstances. This research was not well-timed for a paper presentation at iEMSs in Barcelona, but our project has produced some preliminary successes, from which we hope to provoke discussion at this workshop.

Keywords:autocalibration, parallel computing, data mining, clustering

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Swarm Behaviour for Visual Data Mining

I. Montalvo, J. Izquierdo, R. Pérez and M. Herrera

Centro Multidisciplinar de Modelación de Fluidos, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain
(imontalvo, jizquier, rperez, mahefe@gmmf.upv.es)

Keywords: Particle Swarm Optimization, Visualization, Data Mining

The increasing complexity of many data analysis procedures makes it really difficult for the user to extract useful information out of the results given by the various used techniques. Information visualization and visual data mining can help to deal with the flood of information and to interpret those results. Visual data mining on heterogeneous, imprecise and incomplete information systems need different representations according to human perception (Berthold and Hand 2003). The advantage of visual data exploration is that the user is directly involved in the data mining process. For example, it should allow the user to navigate inside the data, to interact with objects, etc., to be able to create living experience and to see the whole world or concentrate on specific details (Valdés 2004). In this contribution we explore a visualization approach based on techniques borrowed from swarm behaviour, particularly Particle Swarm Optimization (Montalvo et al. 2008). It is able to visualize multidimensional data, uses both standard 2D/3D and dense pixel displays and is based on different interaction/distortion techniques, such as projection, filtering and zooming. The basic idea is to present abstract multidimensional data in some visual form, allowing the user to get concepts visually, draw conclusions, and directly interact with the data.

ACKNOWLEDGEMENTS

This work has been performed under the support of the projects Investigación Interdisciplinar nº 5706 (UPV) and DPI2004-04430 of the Dirección General de Investigación del Ministerio de Educación y Ciencia (Spain), including the support for I+D+i projects from the Consellería de Empresa, Universidad y Ciencia of the Generalitat Valenciana, and FEDER funds. Thanks also to the support of the Grant MAEC-AECI 0000202066, which was awarded to one of the authors by the Ministerio de Asuntos Exteriores y Cooperación of Spain.

REFERENCES

Berthold, M., Hand, D.J., Intelligent Data Analysis: An Introduction. Second Edition. Springler-Verlag Berlin Heidelberg, 3-540-43060-1, 2003.

Montalvo, I., Izquierdo, J. Pérez, R., Fuertes, V.S., Simulation of people evacuation processes through Particle Swarm Optimization. Seventh Interdisciplinary Conference of the International Society of the Arts, Mathematics and Architecture, Universidad Politécnica de Valencia, 2008.

Valdés, J., Building Virtual Reality Spaces for Visual Data Mining with Hybrid Evolutionary-Classical Optimization: Application to Microarray Gene Expression Data. Proceedings of the IASTED International Joint Conference on Artificial Intelligence and Soft Computing, Marbella, España, 2004.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Data Mining vs. Mathematical Modelling: Nonlinear Dynamics and Chaos Theory

Morales-Acoltzi T[a,c,e], Medel-Rojas A.[c], Morales-Aceves F. J[b], Cortés-Pérez E[f] and Morales-Calva X[d].

[a]Departamento de Meteorología Tropical, Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Circuito Exterior, Ciudad Universitaria Mexico D.F. 04510 (acoltzi@atmosfera.unam.mx)
[b]Departamento de Ciencias de la Computación, Universidad Iberoamericana, Campus Puebla, Mexico (aceves83@gmail.com)
[c]Laboratorio de Investigación en Tecnologías Inteligentes, Instituto Tecnológico de Apizaco, Tlaxcala, Mexico (alfonso@toro.itapizaco.edu.mx)
[d]Centro de Investigación Interdisciplinaria y Desarrollo Regional, Universidad Autónoma de Tlaxcala, Mexico ()
[e]Ciencias Ambientales, Facultad de Agro biología, Universidad Autónoma de Tlaxcala, Mexico
[f]Departamento de Ingeniería en Computación, Universidad del Istmo, Ixtepec, CU s/n, Oaxaca, Mexico (cope_144@hotmail.com)

Keywords: Nonlinear Dynamics; Chaos; Attractor; Modelling; Fractal Dimension.

Temporal behavior of natural phenomena has been difficult to characterize and quantify [Boxian et al. 1994; Broomhead et al. 1986]. Complexity of a natural phenomenon does not depend on the number of causes that govern it but essentially on the number of their interconnections, on the magnitude of such linkages and on the feed-back processes [Drazing, 1992]. An extraordinary advance of the environmental sciences will take place in the next years as a result the new technologies used in data mining [Sivakumar, 2000; Sivakumar, 2004; Turcotte, 2003]. Specifically, nonlinear analysis it is known that the long-term behavior of the motion and states of the atmosphere can be described by the global attractor. Namely, starting with a given initial value, the solution will tend to the attractor as time goes to infinity. An attractor is a set to which a dynamical system evolves after a long enough time. That is, points that get close enough to the attractor remain close even if slightly disturbed. The fractal dimension of the atractor in the phase space provides very useful information about the nature of the processes that generated the sequence of values measured in time [Grassberger P. and I. Procaccia, 1983]. In a deterministic system, the present values of the measured properties are related to their previous values. The dimension of the attractor tells us the number of independent variables in this relationship [Eckman, J.P and D. Ruelle, 1992]. "The number of independent variables is the smallest integer that is greater than or equal to the fractal dimension of the attractor". For example, the fractal dimension of the attractor for the Lorenz system is equal to 2.03. Thus a set of 3 equations with 3 independent variables can generate the sequences of values in time of the Lorenz system.
We managed to develop a system combining the **MATLAB** language and the **TISEAN** routines in order to make comparisons between our system and evaluate the results obtained using the **VRA software** and also to a great extent eliminate the "black box" pertaining to this. However there is an advantage to presenting the environment, totally in graph form.

According to these results, we might conclude that the values calculated for the Correlation Dimension are correct and thus its interpretation; this will generally be achieved in the case of observed time series. The key was to use a theoretical time series in our analysis, which is the numerical solution for Lorenz's system, for the variable "z".

BIBLIOGRAFÍA

Boxian W. and Lye M.L. Identification of Temporal scaling behavior of flood: a study of fractas. Fractals, Vol.2 No 2,283-286, 1994

Broomhead D.S. and Gregory P. King. Extracting Qualitative Dynamics from Experimental Data, Physica D,20,217-236,1986.

Drazing Nonlinear Systems. Cambridge, 317 pp, 1992

Eckman, J.P and D. Ruelle. Fundamental limitations for estimating dimensions and Lyapunov Exponents. Physica D,56,185-187, 1992.

Grassberger P. and I. Procaccia, "Characterization of Strange Atractors", Phys. Rev. Lett. 50; 346:349, 1983.

Sivakumar B. "Chaos Theory in hydrology: important issues and interpretations" Journal of hydrology, Vol 227,Issues 1-4,1-20, 2000.

Sivakumar B. "Chaos Theory in Geophysics Past, present and Future "Chaos", Solitons & Fractals, Vol.19, Issue 2, 441-462, 2004.

Turcotte, D.L. "The Relationship of fractals in Geophysics to "New Science". Chaos, Solutions & Fractals, Vol. 27, Issue 1,. 210-217, 2003

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Pre and Postprocessing in KLASS

Karina Gibert[1,2], R. Nonell[1]

[1] Department of Statistics and Operation Research, Technical University of Catalonia, Barcelona
[2] Knowledge Engineering and Machine Learning Group, Technical University of Catalonia, Barcelona, Catalonia

Keywords: Knowldege Discovery from Databases, Artificial Intelligence, Statistics, Prior Knowledge, Clustering based on Rules, PostProcessing, Interpretation Support Tools, Reporting,Data Mining.

KLASS (Gibert et alt 2005b) is a software originally conceived for Knowledge Discovery (KDD) in real domains with complex structure (Gibert et alt 1999). It provides some mixtures of statistical and artificial intelligence tools to support KDD, including basic statistics and providing an integrated system to support the whole process of KDD including pre and post processing, provided that the main data mining technique to be used is related with clustering or rule induction (Gibert et alt 2005c).

Regarding preprocessing, KLASS offers functionalities for basic statistics (simple or by groups), histograms, boxplots (side-by-side), (letter)plots, cross-tables. The performance of the system is quite high, since the user has control over many parameters of de graphics (like the number of classes of a histogram, or the limits of the axis of a plot), being a very flexible tool. It also offers a complete module of data managing, including missing data treatment, creation of transformed variables either using mathematical expressions or via recodification or discretization (here the Boxplot based discretization is provided, which discretizes the numerical variable in such a way that the resulting qualitative variable maximizes association degree with a previously discovered class variable, Pérez-Bonilla et alt 2007). Construction of a prior expert knowledge base (which can be non-complete) is also available and it can be used to biass a posterior clustering process, by means of the Clustering Based on Rules option (Gibert et alt 1999), in such a way that the final classes hold the semantic constraints expressed by the rules.

Regarding the postprocessing, Klass offers some interesting tools to support the interpretation of a clustering results, apart from the classical representation of the dendrogram; It also provides the Class Panel Graph (Gibert et alt 2005), which is a very interestig possibility in clustering contexts to understand better the meaning of the classes. It also implements the CCCS methodology (Pérez-Bonilla 2007) for assigning concepts to every class, improving even more the support to the understanding of the results. There is also a function for visualizing knowledge bases, either probabilized or not, and selecting the rules with degrees of certainty over a certain threshold.

One of the particularities of the system is that it is designed in such a way that the outputs, either graphical or numerical or textual,
are produced in LaTeX font files, which are directly processed by the kernel of KLASS and automatically sent to the LaTeX viewer and displayed on the screen. From the final user point of view, this makes no difference with other systems, since graphical representations are directly displayed on the screen as well as other results. However, as reporting the results of the KDD process is always involved with the ellaboration of technical papers, KLASS also includes a reporting facility in such a way that the user can specify a set of steps to be performed sequentially and a single big LaTeX document including all the results is produced. The user only needs to edit this document and add personal comments on it to get a complete report of the analysis. KLASS provides either standard or personalized reports. It is a flexible possibility since it is possible to automatically transform every result of the single steps into PostScript or PDF documents, which can be managed as usual, for example, pasting it into a Word document.

If the document to be produced is long and with complex structure and contains hard mathematical notation, LaTeX offers nice advantages and the LaTeX results provided by KLASS are really useful. In this cases, LaTeX is a widely used text processor, owing to the excellent support it provides to the generation of high quality mathematical formulae and scientific notation. However, including graphical representations from commercial statistical packages in a LaTeX document requires the use of special LaTeX packages to deal with graphical formats and makes a little bit more complicate the ellaboration of the document, which requires transformation to PostScript or PDF to be completely visualized. Since the results of KLASS are produced in native LaTex code, inclusion of those graphics in the final report becomes trivial.

On the other hand, making the native LaTeX code accessible to the user permits the user to adjust labels or size of the titles of graphical representations. In this way, the quality of the image is maintained to its final use.

Acknowledgments: The project TIN2004-01368 has partially financed the development of GESCONDA.

References

Gibert, K., Annicchiarico, R., Cortés, U. and Caltagirone, C. 2005a. *Knowledge Discovery on Functional Disabilities: Clustering Based on Rules Versus Other Approaches*. IOS Press.

Gibert, K., Nonell, R., Velarde, JM, Colillas, MM 2005b. Knowledge discovery with clustering: Impact of metrics and reporting phase by using klass. *Neural Network World,* 319-326.

Gibert, K., Nonell, R, 2005c. Descriptive statistics with KLASS. Supporting LaTeX documents ellaboration, In Procs. 3rd World Conferencs on Computational Statistics and Data Analysis (IASC 2005), pp 90. Limassol, Cyprus.

Gibert, K. and Sonicki, Z. 1999. Clustering based on rules and medical research. *Journal on Applied Stochastic Models in Business and Industry*, formerly JASMDA 15(4): 319-324.

Pérez-Bonilla A, Gibert, K (2007) Automatic generation of conceptual interpretation of clustering. In Progress in Pattern Recognition, Image analysis and Applications.Lecture Notes in Computer Science v 4756, pp 653-663. Springer.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

GESCONDA: from Environmental Data Mining to Environmental Decision Support

M.Sànchez-Marrè[a] **and K. Gibert** [a,b]

[a] *Knowledge Engineering and Machine Learning Group, Technical University of Catalonia Catalonia, Jordi Girona 1-3, 08034 Barcelona, Catalonia.*
(miquel@lsi.upc.edu, karina.gibert@upc.edu)
[b] *Department of Statistics and Operations Research, Technical University of Catalonia, Jordi Girona 1-3, 08034 Barcelona, Catalonia.*

Keywords: Knowledge Discovery from Databases; Data Mining; Environmental applications; integrated approach

GESCONDA is a tool for intelligent data analysis and implicit knowledge management of databases, with special focus on environmental databases. Differing from existing commercial systems, the more relevant aspects of this proposal are the incorporation of the statistical data filtering and pre-processing in the same software tool together with the intelligent data analysis techniques as well as the interaction of different data mining methods. Either statistical techniques or Artificial Intelligence techniques or even mixed techniques are combined and used to extract the knowledge contained within data.

The authors are not aware of the existence of a specific software for knowledge discovery and data mining of environmental databases, taking into account the special features of environmental domains, such as the temporal and dynamic aspects of data, including both statistical data mining and statistical modelling methods, or the problem of noisy data, and data filtering with no clear relevant or irrelevant features. In fact, these are major differences with other commercial or freeware software.

The issue of our work aims at designing and building an Intelligent Knowledge Data Discovery and Data Mining System, especially suitable for environmental data analysis. The software tool, which is called GESCONDA [Sànchez-Marrè et al., 2002], is built-up. Addition of new functionalities will take place in the near future.

On the basis of previous experiences described in Sànchez-Marrè et al. [1997, 1999], it was decided that GESCONDA would have a multi-layer architecture of 4 levels connecting the user with the environmental system or process. These 4 levels are briefly described below:

- Data Filtering: Provides statistical tools for data cleaning, including one-way or two-way analysis, even graphical representations, missing data or outlier analysis, as well as management and variable transformations.

- Recommendation and Meta-Knowledge Management: supporting the formal definition of problem goals, meta-knowledge of variables and examples, methods recommender, parameter setting, and domain knowledge elicitation.

- Knowledge Discovery: including several statistical and machine learning data mining algorithms like clustering, decision tree induction or case based resoning, among others, as well as some mixed techniques, some of them developed by the authors, as detailed in Gibert, [1998], Comas et al. [2001], Sánchez-Marrè et al. [1999], and Gibert [2004].

• Knowledge Management: making possible the integration of different knowledge patterns for a predictive task, or planning, or system supervision, as well as the validation of the knowledge patterns produced in the previous step. User interaction is important in this phase, and the system supports it.

GESCONDA will be useful to acquire relevant knowledge from environmental systems, on the basis of available databases. This knowledge will be used afterwards in the implementation of reliable EDSSs. The portability of the software is provided by a common Java platform. In next sections suggestions on the use of the software are provided.

GESCONDA is a standard Java application with a friendly graphical user interface (GUI).

Input data files can be analysed by GESCONDA. They follow the standard format of instance arranged in rows, and attributes in columns. Prior to the data file loading, the user should introduce the meta-information associated with each variable into the system. The variable type must be specified: quantitative or qualitative, and in this case, also the list of modalities. For ordered qualitative variables, the ordering of modalities must be also provided. Also, the weight of the variable can be modified. In addition, the variables can be declared as active for the analyses or not, depending on the user's requirements.

After data file loading, all changes can be saved into a GESCONDA database file format (GSP/GCDA file), in order to recover the work in future working sessions with the tool. Recovering the work is done through the opening of a previously created database (GSP/GCDA file). Once data are loaded, the first thing to do in order to extract knowledge patterns from data is the descriptive statistical analysis and the data filtering task. This operations let the user check whether there are errors, outliers, badly codified data, missing data, as well as summarise main data features, such as the minimum and maximum values, the mean, standard deviation, variance, and so on. In addition, tools for arranging the variables according to the user needs are provided, when variable transformations, such as linear transformation, variable re-coding or variable standardisation, are needed prior to analysis. Other facilities such as random variable generation following several distributions such as Bernouilli, Binomial, Gaussian, Exponential, Uniform, Discrete or specific probability value computation are also available.

According to the user's goal, which could be to discover some concepts hidden in the data (clustering or grouping), or to discover some discriminant knowledge (decision trees, classification rules) to predict to which kind of concept (class or cluster) each instance belongs, different data mining techniques can be used. Furthermore, they could be combined to make a more accurate data analysis. Different scenarios are possible.

A common scenario for our system could be when an unknown environmental database, with a huge amount of instances and/or features, is faced. In this case, one possibility is to start using a clustering technique to identify typical situations in the target environmental process. Several methods are available in GESCONDA, such as K-means [Dubes and Jain, 1988], Isodata [Ball and Hall, 1965], Nearest-Neighbour classifier [Cover and Hart, 1968; Duda and Hart, 1973], Marata and COBWEB/3 [McKusick and Thompson, 1990].

After experimenting with different techniques, and trying several parameter values of the methods, the tool provides the user with a sensitivity analysis regarding the applied methods, which can be used for finding the stable set of classes, as detailed in Gibert et al. [2004]. The obtained classes and prototypes can be visualized. The resulting class is recorded as a new attribute or variable.

Afterwards, an inductive decision tree technique can be used to discover a predictive knowledge model, such a decision tree, to find the best set of attributes predicting the class label for new instances of the environmental database. In GESCONDA, the user can select and test ID3 [Quinlan, 1986], CART [Breiman et al., 1984], C4.5 [Quinlan, 1993], with optional pruning techniques. Another complementary action, to predict the class label for a new instance, is to directly induce classification rules. Several methods exist in the machine learning field; RULES [Pham & Aksoy, 1995], PRISM [Cendrowska, 1987], CN2 [Clark & Niblett, 1989], and RISE [Domingos, 1996] are implemented in GESCONDA. The user can test several parameters and validate the obtained classification rules. Some validation

techniques (simple validation, cross-validation) are also available to test the quality of the induced models.

Both from the decision tree model or from the directly induced classification rules, a predictive knowledge model, implemented as a knowledge base, can be directly built with the final classification rules. This knowledge pattern can be used, for instance, to set-up an IEDSS for predictive tasks in an environmental domain.

For data without qualitative variables, a statistical modelling component is also available. In that case, quantitative models with several charts, graphs and model parameter estimation are found after a validation process, like multiple linear regression, ANOVA analysis [Lebart, 1990] or correlation models.

As an example of the use of GESCONDA, an IEDSS was built-up to supervise a wastewater treatment plant operation as described in Rodríguez-Roda et al. [2002].

Currently, GESCONDA is composed by several statistical data filtering analysis methods, such as one-way and two-way descriptive statistics, missing data analysis, clustering, relationship between variables, hybrid Artificial Intelligence and Statistical methods, as well as several machine learning techniques, coming from Artificial Intelligence, such as conceptual clustering methods, decision tree induction and classification rule induction.

The project purpose is to extend these intelligent system with some new agents and computational modules, such as case-based reasoning techniques, soft computing methods, support vector machines approaches, statistical models, dynamical analysis techniques, and hybrid methods integrating Artificial Intelligence approaches and Statistical ones.

The prototype is evolving from a simple data mining and knowledge discovery tool to a more complex intelligent environmental decision support tool, with a high emphasis on environmental features like:

- Huge amount of data
- Incomplete information: many missing values
- Many descriptive features: feature relevance problem
- Temporal / Spatial feature: Dynamic and Spatial data analysis
- Different Data format: Spatial data formats

ACKNOWLEDGEMENTS

The authors wish to thank the Spanish Ministry of Science and Technology for the partial financial support of the project TIN2004-01368.

REFERENCES

Ball G.H. and Hall D.J. ISODATA, a novel method of data analysis and pattern classification. Technical Report, Stanford Research Institute, 1965.

Bratko I., Dzeroski S., Kompare B. And Urbancic T. *Analysis of environmental data with machine learning methods*. Jozef Stefan Institute, 2000.

Breiman, L., Friedman, J.H., Olshen R.A. and Stone, C.J. *Classification and Regression Trees*, Wadsworth, Belmont, 1984.

Cendrowska,. J 1987. PRISM: an algorithm for inducing module rules. *Int'l Journal of Man-Machine Studies* 27(4):349-370.

Clark & Niblett, 1989. P. Clark & T. Niblett. The CN2 induction algorithm. *Machine Learning* 3:261-283.

Comas J., Dzeroski S. Gibert, K., Rodríguez-Roda I. and Sànchez-Marrè M. Knowledge Discovery by means of inductive methods in wastewater treatment plant data. *AI Communications* 14(1):45-62, January 2001.

Cover T.M. and Hart P.E. Nearest neighbour pattern classificat-ion. *IEEE Transactions on Information Theory*,13,21-27,1968

Domingos, P 1996 Unifying Instance-Based and Rule-Based Induction, *Machine Learning* 24(2):141-168.

Dubes R. and Kain A. Algorithms for Clustering Data. Prentice-Hall, Englewood Cliffs, USA, 1988.

Gibert K. AI and Statistics techniques for Knowledge Discovery and Data Mining. In *Tendencias de la Minería de Datos en España*. (Eds.) R. Giráldez, J. C. Riquelme, Jesús S. Aguilar-Ruiz. In press, 2004.

Gibert K. and Cortés U. Clustering based on rules and Knowl-edge Discovery in ill-structured domains. *Computación y Sis-temas* 1(4):213-227. CIC, Instituto Politécnico Nacional, 1998.

Gibert K., Cortés U, Aluja, T., Knowledge Discovery with Clustering Based on Rules. Interpreting Results,In *Principles of Data Mining and Knowledge Discovery*, J. M. Zytkow, M. Quafafou Eds., LNAI 510, p 83-92, Springer-Verlag, 1998.

Gibert K., Flores X. and Sànchez-Marrè M. Comparison of classificat-ions in environmental databases using GESCONDA. *In 4th ECAI Workshop on Binding Environmental Sciences and Artificial Inte-lligence (BESAI'2004)*, 13-1:13-10, at ECAI'2004. València, 2004.

Kanevski M., R. Parkin, A. Pozdnukhov, V. Timonin, M. Maignan, V. Demyanov and S. Canu. Environmental Data Mining & Modelling based on Machine Learning algorithms and Geostatistics, *Environmental Modelling & Software* 19(9): 845-856, 2004.

Lébart, L. Traitement statistique des données. DUNOD, Paris, 90

McKusick K. and Thompson K. COBWEB/3: A Portable Implementation, Technical Report FIA-90-6-18-2, NASA Ames Research Center, June 20, 1990.

Morabito F. C. *Environmental data interpretation: the next challenge for intelligent systems*. NATO Advanced Research Workshop on Systematic Organization of Information in Fuzzy Systems. 2001, Vila Real, Portugal.

Pham & Aksoy, 1995. RULES: a simple ruler extraction system. *Expert Systems with Applications* 8(1).

Quinlan J.R. *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Mateo, CA, 1993.

Quinlan J.R. Induction of decision trees, *Machine Learning* 1(1) (1986), 81- 106.

Rodríguez-Roda I., Comas J., Colprim J., Poch M., Sànchez-Marrè M., Cortés U., Baeza, J. and Lafuente J. A hybrid supervisory system to support wastewater treatment plant operation: implementation and validation. *Water Science & Technology*, 45(4-5), 289-297 (2002).

Sànchez-Marrè M., Gibert K., Rodríguez-Roda R., Bueno E., Mozo L., Clavell A., Martín M. and Rougé P. (2002). Development of an Intelligent Data Analysis System for Knowledge Management in Environmental Data Bases. In Rizzoli, A. E. and Jakeman, A. E. (eds.) Integrated Assessment and Decission Supoort. *Proceedings of the First biennial Meeting of the International Environmental Modelling and Software Society*, Vol. 3, pp:420-425. iEMSs 2002.

Sànchez-Marrè M., U. Cortés, I. R.-Roda, M. Poch. Sustainable Case Learning for continuous domains. *Environmental Modelling & Software* 14:349-357, 1999.

Sànchez-Marrè M., Béjar J., Cortés U., Gràcia J., Lafuente J. and Poch M. Concept formation in WWTP by means of classification techniques: a compared study. *Applied Intelligence*, 7(2), 147-166. 1997.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Qualitative models for Sustainability

Workshop 5 (Wednesday, July 9th, 2008) *at*
iEMSs – the International Environmental Modelling and Software Society
Barcelona, July 7-10, 2008

Organisers:

- Bert Bredeweg
  University of Amsterdam, Informatics Institute, Human Computer Studies lab
  The Netherlands
  E-mail: bredeweg@science.uva.nl
- Paulo Salles
  University of Brasilia, Institute of Biological Sciences
  Brazil
  E-mail: psalles@unb.br

Abstract
Qualitative Reasoning is a research area within Artificial Intelligence that focuses on means to articulate and communicate conceptual knowledge such as system structure, causality, the start and end of processes, the assumptions and conditions under which facts are true, qualitative distinct behaviours, etc. Humans continuously reason about the physical world that surrounds them. Understanding how systems behaved in the past and how they may behave in the future is a crucial aspect of human nature. Qualitative Reasoning investigates how this aspect of human intelligence can be automated on computers.

NaturNet-Redime (NNR) is a project co-funded by the European Commission within the 6th Framework Program (2002-2006). The general objective of the NNR project is the provision of a web-portal to improve awareness on Sustainable Development. The project develops and demonstrates prototype technology and educational programs towards implementing the European Union's Strategy for Sustainable Development, focussing on environmental management. In this workshop the partners from the NNR project will present their latest developments on the main themes within the approach taken: software, models and model library, curriculum, and evaluative case studies with stakeholders. In addition the workshop is open for contributions of researchers and practitioners outside the NNR project who are working with or developing materials along similar dimensions. The main goals of the workshop are to share results, identify problems and opportunities, and to support the development of a user-group on qualitative approaches to environmental management.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Garp3 - Workbench for Qualitative Modelling and Simulation

B. Bredweg, A. Bouwer, F. Linnebank and J. Liem

Laboratory Human Computer Studies Laboratory, University of Amsterdam
Kruislaan 419 (matrix 1), 1098VA Amsterdam, The Netherlands
{bredeweg,bouwer,linneban,jliem }@science.uva.nl

Keywords: Qualitative Reasoning, Workbench, Software, Diagrammatic representation, Modelling, Conceptual modelling, Education, Training, Decision-making.

Abstract

Garp3 is a domain independent, multi-platform, qualitative modelling and simulation environment. It allows modellers to articulate and refine their conceptual domain knowledge and analyse this knowledge through simulation. Garp3 has been successfully applied in Ecology and Sustainable Development (SD) and is freely available via (http://www.garp3.org).

Garp3 and the NaturNet-Redime Project

Ecologists in the NaturNet-Redime project (http://www.naturnet.org) are building qualitative models about sustainable development issues through several case studies. For this purpose the Garp3 workbench for building, simulating, and inspecting qualitative models was developed (Bredeweg e al, 2006). The main goals of the development was making qualitative reasoning technology usable for non-computer scientists by creating a uniform user interface, a diagrammatic visual language for representing model content, and graphical buttons to communicate the available user options and manipulations. Garp3 is implemented in SWI-Prolog (http://www.swi-prolog.org) and seamlessly integrates three previously developed software components: Garp2 for simulating models, Homer for building models, and VisiGarp for inspecting simulation results.

To further support the ecologists in their modelling efforts a structured approach to modelling was developed. This framework helps modellers refine their initial ideas, represented in concept maps, into detailed conceptualisations, such as structural models, causal models, and expected model behaviour. The diagrams in the final steps of the framework are close to the actual modelling primitives used to implement qualitative models. The Sketch environment in Garp3 supports modellers with tools to create the required diagrams. These Sketches not only help modellers in creating a model, but also serve as a more general description of the final model for other users of the model.

One of the main goals of the NaturNet-Redime project is to create a sustainable development curriculum that allows students to learn about specific issues through qualitative modelling and simulation. To advance this goal the case study models developed in project are being integrated into a single library of sustainability concepts. Students can run and adapt different scenarios, let Garp3 automatically gather the correct knowledge relevant to the simulation and predict the possible outcomes, and analyse the results. To support the integration of the different case study models multiple model support and copy/paste functionality have been added to Garp3. The copying functionality assures that models remain syntactically correct, avoids adding redundant knowledge, preserves existing knowledge, and merges conceptual knowledge in a semantically correct manner as much as possible. The copy functionality also allows modellers to reuse parts of

already existing models. Furthermore, a qualitative model repository was developed to allow modellers to share and search for potentially reusable models. The copying functionality is currently actively used in the project.

Modelling and Simulation

The model implementation process starts with creating model ingredient definitions, e.g. the objects (entities), relations (configurations), quantities, external influences on the system (agents) and the different assumptions made by the modeller. Secondly, modellers create different scenarios (initial situations of the system) (Figure 1) using the defined modelling primitives. Thirdly, the modeller represents the different processes that can happen in the system in different model fragments (Figure 2). These model fragments represent the causal dependencies that exists between different quantities by representing a part of the structure of the system as conditions and the causal dependencies as consequences.

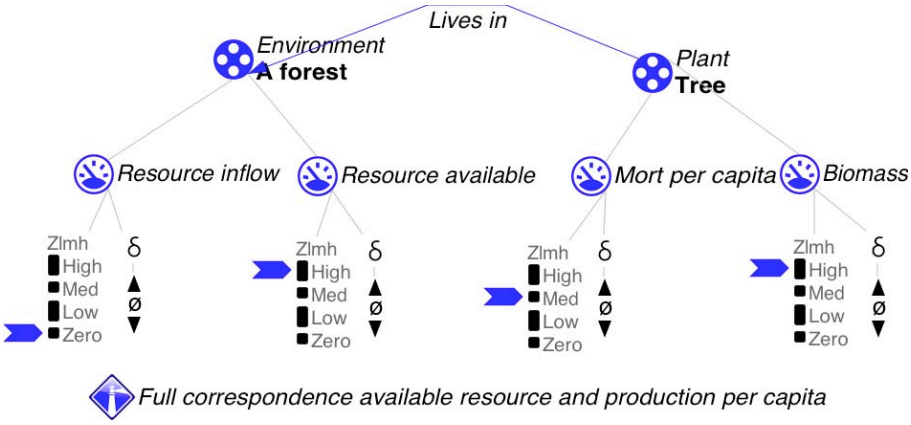


Figure 1: Scenario: production ok, high biomass.

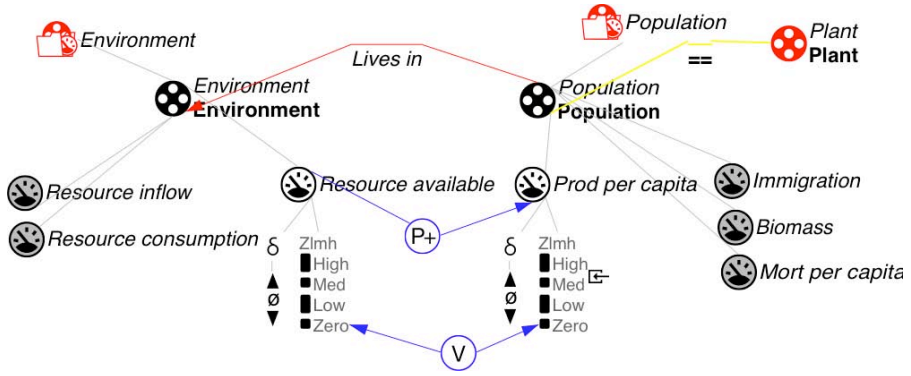


Figure 2: One of the model fragments: Production.

The simulation starts by selecting a specific scenario. The reasoning engine (Garp3) tries to find matching model fragments, i.e. model fragments that have conditions that are fulfilled by the scenario. The consequences of the model fragment are added to the scenario, which potentially allows more model fragments to match. When no more matching model fragments can be found, the derivatives of the quantities are calculated using the derived causal dependencies. Using these derivatives the next possible states of behaviour are determined (possibly multiple due to ambiguity). For each of the new states the same algorithm is used to derive new states. The result is a state graph describing the complete behaviour of a system given the initial scenario (Figure 3). Modellers can inspect the value history and/or the derived inequalities for sets of states (Figure 4), and the dependencies of the system in specific states (Figure 4).

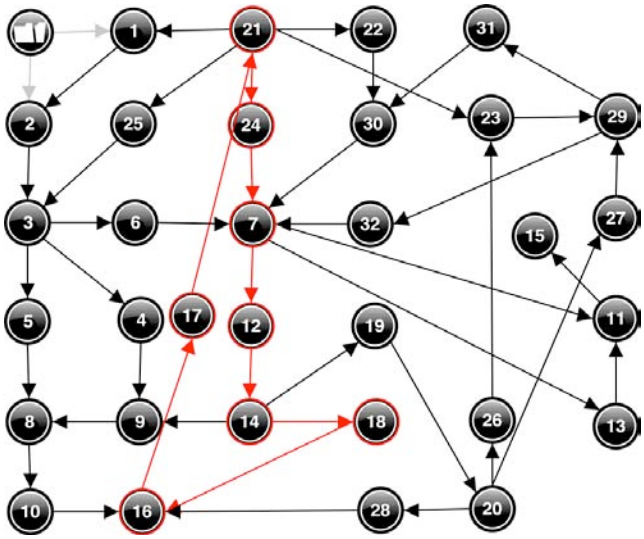
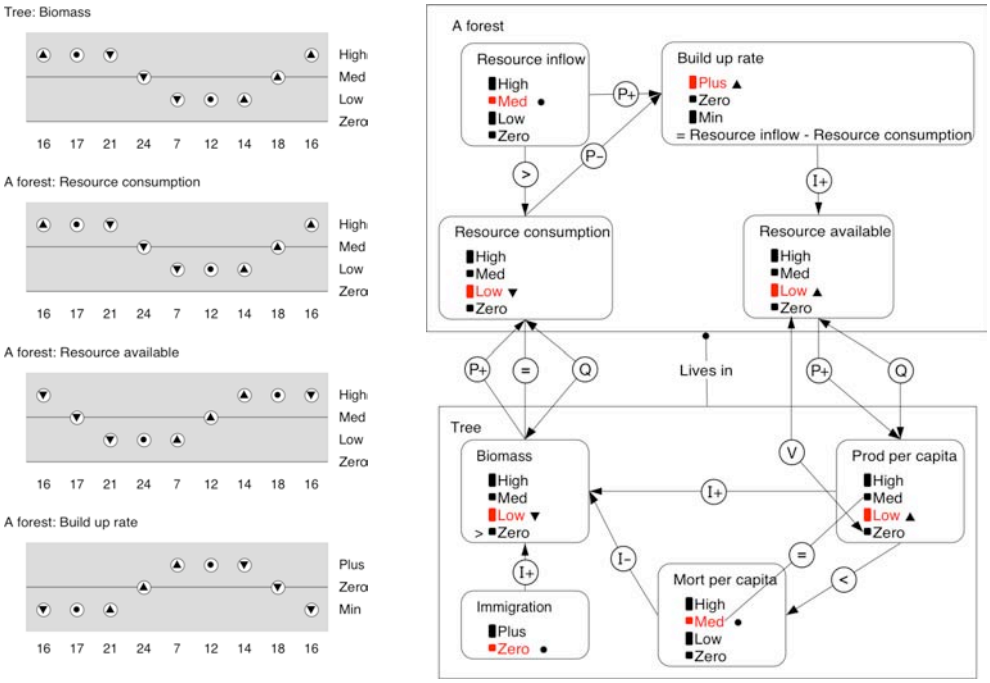


Figure3: State graph after simulating scenario with selected cycle.



(a) Value history of selected states. (b) Dependencies in state7.

Figure4: Value history (a) and Causal model (b).

ACKNOWLEDGEMENTS

The authors wish to thank Jelmer Jellema (Spin in het Web, http://www.spininhetweb.nl/ and Dirk Bertels (Louter vormgeving, http://www.louter.nl/ for their contributions to Garp3.

REFERENCES

Bredeweg, B., Bouwer, A., Jellema, J., Bertels, D., Linnebank, F., and Liem, L. Garp3 - A new Workbench for Qualitative Reasoning and Modelling. *20th International Workshop on Qualitative Reasoning (QR'06)*, C. Bailey-Kellogg and B. Kuipers (eds), pages 21-28, Hanover, New Hampshire, USA, 10-12 July 2006.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
Workshop on Qualitative Reasoning models on Sustainability
Barcelona, July 6-10, 2008

A qualitative model on sustainability of the Riacho Fundo basin (Brasília, Brazil)

Paulo Salles[a], Bert Bredeweg[b] and Ana Luiza Rios Caldas[a]

[a]*Institute of Biological Sciences, University of Brasilia, Brazil*

[b]*Human Computer Studies Laboratory, University of Amsterdam, The Netherlands*

Keywords: Sustainability; Qualitative reasoning; Water resources management; Urbanisation

Urbanisation seems to be the future of the global society (Egger, 2006). Understanding the consequences of land use change and the establishment of urban areas is essential for making sustainable this future. This paper describes an implemented qualitative simulation model about the Brazilian Riacho Fundo case study for the Project NaturNet – Redime[1]. This case study bridges the gap between case studies focussing on sustainability of water bodies in protected areas (River Mesta, Bulgaria, and the Delta of the Danube River, Romania), and case studies that focus on river management and rehabilitation, in areas strongly modified by human influence (River Kamp, Austria, and Rivers Trent and Grand Ouse, England). A complete description of the Riacho Fundo model is presented in Salles (2007).

The Riacho Fundo basin has been one of the most disturbed areas under Brasilia's influence area since the beginning of the new capital construction, in the 1950s. It was progressively impacted by the transformation of natural areas into rural and urban areas. These changes eventually created a densely populated urban structure where few rural properties still keep fruits and vegetables production and food processing industries. In these areas, springs, streams and natural vegetation patches are still protected and keep local biodiversity (see Salles and Caldas, 2006, for an overview of general aspects of the basin). A group of stakeholders prepared a list of relevant problems in the Riacho Fundo basin (Salles and Caldas, 2006). The uncontrolled land occupation and changes in land use were considered the most important problems. Deforestation, biodiversity loss, problems related to water resources management and low community participation in decision making process were also considered highly relevant. Accordingly, the following objectives were set for this modelling effort: (a) to improve understanding of environmental systems and problems that may affect sustainability in the basin; (b) to demonstrate the effects of human actions, both positively and negatively influencing different aspects of the Riacho Fundo basin system; (c) to provide support for stakeholders to learn about sustainability.

Sustainability is addressed in the Riacho Fundo model from three perspectives, *Rural*, *Semi-urban* and *Urban*. A perspective is a set of simulations that address a specific topic and answer a certain type of questions (Falkenhainer and Forbus, 1991). The *Rural* perspective focus on human use of natural resources for agricultural purposes. The *Semi-urban* perspective explores changes in land use that transform natural and agricultural areas into urban areas. The *Urban* perspective addresses the city and its physical infrastructure, particularly for the engineered drainage system for rainfall. These perspectives combine aspects of economy, governance, culture and human well being. The model was implemented in Garp3 (Bredeweg *et al*., 2006) using elements from the Qualitative Process theory (Forbus, 1984), and following a compostional modelling approach (Falkenhainer and Forbus, 1991). Accordingly, the model consists of a library of model fragments, stand-alone units that encode knowledge about structural configurations, situations and the mechanisms of change (processes). Two modelling primitives, direct influences (*I*+ and *I*–) and qualitative proportionalities (*P*+ and *P*–), are very important as they implement causal relations and mathematical relations that determine the qualitative values of the quantities. Table1 presents the main concepts addressed in each perspective.

[1] NaturNet-Redime, EU STREP, project number 004074. Available at www.naturnet.org.

	Perspectives		
Land use	**Urban**	**Semi-urban**	**Rural**
Main problems	Drainage system; flooded areas; transported garbage and damage caused by floods	Urbanization; water infiltration; and soil erosion	Erosion; loss of water resources and biodiversity
Economic features	Services: garages	Industry: textile and food industries	Agriculture: cattle; crops
Soil	Impermeable soil	Soil particle aggregation	Soil fertility
Water resources	Effects of uncontrolled flow of water run off and of the drainage system	Effects of erosion and underground water on springs and rivers	Effects of erosion and underground water on springs and streams
Biological entities	Mosquitos, Pathogens	Vegetation	Vegetation; Vertebrates; Capybara
Human	Economic activities; Human well-being: garbage and water related diseases	Economic activities	Economic activities
Agents	Rainfall	Urbanization	--
Sustainability	Control of diseases; Control of residues	Water quality; Control of residues	Soil fertilization; Reuse of residues

Table 1. Overview of the main concepts addressed by the Riacho Fundo model.

The structure of the Riacho Fundo system involves 32 entitites, and the model consists of 105 model fragments, being 24 implementing the *Rural*, 18 the *Semi-urban* and 30 the *Urban* perspective. Besides that, seven model fragments represent different types of water bodies, as for example, springs, streams and rivers, and 20 model fragments implement a general representation of economic activities, applied to the three perspectives (see Table 1). The current version of the model includes 19 simulations in the *Rural* perspective, 17 in the *Semi-urban* and 18 in the *Urban* perspectives.

Simulations within the *Rural* perspective provide answers to the following question: *what are the effects of soil protection and erosion on agricultural production, on rivers and streams, and on animal biodiversity in the rural area of the Riacho Fundo basin?* A simulation with the scenario 'Crop production influenced by soil nutrients and influences on stream and biodiversity' illustrates how the model provides answers to this question. The scenario is organized around the following entities and configurations: 'Vegetation' covers 'Rural soil'; 'Rural soil' next to 'Stream'; 'Stream' supports 'Capybara'; 'Rural soil' occupied by 'Rural RF'; finally, 'Rural RF' supports 'Crop' and 'Stream' supports 'Crop'. The simulation produces seven initial states and 65 states in total. Figure 1 shows the causal model obtained in state 7. Causality flow starts with changes in the vegetation: if *degradation* is greater than *regeneration*, vegetation *growth rate* is negative and causes *vegetation cover* to decrease. Besides being involved in two feedback relations (a negative influence to stabilize *growth rate* and a positive one to *regeneration*), *vegetation cover* influences *soil structure*, so that when the former decreases, the quality of soil structures also decreases, causing an increasing in the erosion process. Note that the entity 'Rural RF' represents land use, and therefore *eroded land* refers to the whole basin. As *soil structure* is decreasing, *eroded land* increases, and causes *nutrient* and *fertility* to decrease. *Removed soil* is also increasing, and causes the amount of *sediment* to increase. This change causes *depth* and *amount of water* in the stream to decrease. A reduction in the *amount of water* negatively influences the survival of mammals, such as the capybaras, causing a decrease in the local biodiversity. Crop production in the Riacho Fundo basin depends on water and nutrient availability. As both *amount of water* and *fertility* are decreasing, *resource inflow* to the production process decrease, influencing the *production rate*. A negative rate causes agricultural *products* to decrease. Associated to the production process, both *residues* in general and the *pesticides* particular also decrease. This way, the model establishes a causal link between changes in vegetation cover and changes in soil fertility, amount of water, local biodiversity and crop production. As ambiguity is inherent to qualitative

representations, it is assumed that some quantities correspond to other quantities (see Q relations in Figure 1). This modelling primitive implements the notion that two correspondent quantities always have specific pairs of values simultaneously. The effect is a reduction in the number of states in the simulation.

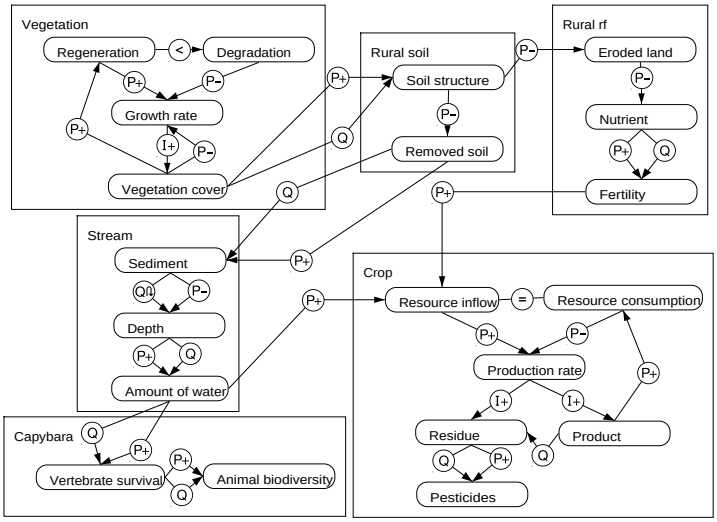


Figure 1. Causal model in state 7 of the simulation 'Crop production influenced by soil nutrients and influences on stream and biodiversity', *Rural* perspective of the Riacho Fundo model.

Simulations within the *Semi-urban* perspective provide answers for the question: *how changes in soil features, under the conflicting forces of urbanization and natural vegetation dynamics, may affect functional springs, rivers and the industrial production in the Riacho Fundo basin?* Finally, simulations in the *Urban* perspective answer the question: *what are the benefits of having a good infrastructure for the drainage of rain water in the urban area of Riacho Fundo with respect to soil protection, economic damages caused by garbage, and the human well being?* Concluding, the model provide answers to relevant questions about sustainability. Ongoing work includes extending the coverage of rural and urban features. From the content point of view, the Riacho Fundo model has the potential for significantly improving stakeholders' understanding about their everyday problems.

REFERENCES

Bredeweg, B.; Bouwer, A.; Jellema, J.; Bertels, D.; Linnebank, F. and Liem, J.(2006) Garp3 - A new Workbench for Qualitative Reasoning and Modelling. *Proceedings of the 20th International Workshop on Qualitative Reasoning (QR-06)*, 21-28, 2006.

Egger, S. (2006) Determining a sustainable city model. *Environ'l Modelling and Software* 21: 1235-1246.

Falkenhainer, B. & Forbus, K. (1991) Compositional Modeling: Finding the Right Model for the Job. *Artificial Intelligence*, 51(1-3): 95-143.

Forbus, K. (1984) Qualitative Process Theory. *Artificial Intelligence*, 24: 85-168.

Salles, P. (2007) QR models and documentation for learning about sustainable development, focusing on basic biological, physical, and chemical processes related to the environment in the Riacho Fundo. Deliverable D6.4.2, NaturNet-Redime Project.

Salles, P. and Caldas, A.L.R. (2006) Textual description of Riacho Fundo case study focusing on basic biological, physical, and chemical processes related to the environment. Deliverable D6.4.1, NaturNet-Redime Project.

Weld, D. & de Kleer, J. (eds.) (1990) Readings in Qualitative Reasoning about Physical Systems. San Mateo, CA: Morgan Kaufmann.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Qualitative modelling of limiting factors for a salmon life cycle in the context of sustainable river rehabilitation

R. A. A. Noble[a], B. Bredeweg[b], P. Salles[c] & I. G. Cowx[a]

[a] The University of Hull International Fisheries Institute, Hull, HU6 7RX UK (r.a.noble@hull.ac.uk and i.g.cowx@hull.ac.uk)
[b] Human Computer Studies, University of Amsterdam, Kruislaan 419 (matrix I), 1098 VA Amsterdam, The Netherlands (bredeweg@science.uva.nl)
[c] University of Brasilia, Brazil (psalles@unb.br)

Keywords: Qualitative modelling; life-cycle; salmon; limiting factors; survival.

1. MODEL BACKGROUND

This paper describes aspects of a Qualitative Reasoning (QR) model which was developed as part of the EU funded NaturNet-Redime (NNR) project (Sixth Framework Program, Project no. 004074 http://www.naturnet.org). This project contributed to the European Union's Strategy for Sustainable Development [European Commission 2001] in the context that the NNR project revolved around developing new technologies (of which QR is one example) for use within a general educational setting, for exploration of sustainable development issues. One of the goals of the NNR project was to improve the Garp3 QR software tool and to construct guidelines for a standardised approach to QR modelling [Bredeweg *et al*. 2007]. The model presented here was developed following this standardised modelling approach and focused on the representation of a species life cycle within the context of river rehabilitation programmes. As such the case study addresses the issues of sustainability of aquatic resource management, an issue that is currently of great concern in the European Union.

River rehabilitation projects in the UK are often targeted to economically valuable fish species (e.g. Atlantic salmon, *Salmo salar* L.). Conservation and management of these species is often based around quantitative life cycle models e.g. [Aprahamian, Wyatt & Shields, 2006] which examine the recruitment of individuals to each consecutive life stage to identify the factors that are impinging on, or limiting the size of, the population. Hence, planning of rehabilitation activities focuses on the key human activities that impact on the life cycle of the fish populations/community being managed. QR modelling has been previously used to examine the functioning of Atlantic salmon redds (spawning areas) [Guerrin & Dumas 2001a,b]. However, this model focused only on one phase during the life cycle. A QR model which encompasses a full life cycle, together with the socio-economic factors of sustainable management, may be an extremely valuable tool. In this context, such a QR model could be used to provide a basis for environmental managers, researchers, stakeholders and students to investigate the potential outcomes and conflicts within a given rehabilitation programme.

2. MODEL STRUCTURE, ISSUES AND SOLUTIONS

The model was developed in the Garp 3 QR modelling environment which allows knowledge about concepts, properties and processes to be integrated in the form of model

fragments (a sub-unit of a model describing a particular aspect of the system) and scenarios which integrate the knowledge from many model fragments in the form of scenarios, simulations and behaviours. The whole scenario of this model follows the concept of the life cycle and the influence human activities have on habitat quality and consequently on the survival of individuals at each life stage (each life stage is considered to be an individual population). In the model each life stage recruits into the next life stage from adult to egg to juvenile to smolt back to adult. Each life stage inhabits a particular river type (eggs and juveniles in upland river; adults migrating through a lowland river) and particular habitat type (eggs in spawning habitat; juveniles in juvenile habitat). The river types flow through catchments (upland and lowland) in which human activity occurs in the form of degradation agents which affect the habitat and river quality for a specific life stage.

The key concepts within the life cycle are the different life stages and the survival of an individual from one life stage to the next is the fundamental process which needs to be represented by the system. In this context each life stage is considered to be an independent population within the model. The basic concepts for a population are described by a number of model fragments which describe the quantities "*Number recruited*" and "*Number surviving*" for a life stage. Within the model the concepts of mortality, survival and the *Numbers surviving* within each life stage, and the factors controlling it, were implemented using the following concepts:

1) The *Numbers surviving* a particular life stage are limited by either or both of the *Numbers recruited* to the life stage and the *Habitat quality* (whichever is lesser) – the maximum size limit of the *Numbers surviving* in any situation can be considered as a quantity, the *Population potential*.
2) The *Numbers surviving* is limited by the *Population potential* and changes in response to being >, < or = to the potential.
3) The changes in the *Numbers surviving* (due to an imbalance with *Population potential*) result from changes in the balance of the level of recruitment and the mortality/survival rates – the net effect of these two rates can be modelled as a single quantity, the *Net survival dynamic*.

Following this *Population potential* is a conceptual quantity that is a combination of the *Number recruited* and the *Habitat quality* which was implemented using complex value correspondences between the controlling variable and the *Population potential* (where the controlling quantity was the quantity with the lesser magnitude). The *Numbers surviving* is then controlled by a directed influence from the *Net survival dynamic* which itself is derived as a calculus *Population potential* minus *Numbers surviving*.

Interrogation of the behaviour paths and dependency diagrams during the model development indicated inconsistent behaviour relating to the derivative behaviour of *Net survival dynamic* when both *Numbers recruited* and *Habitat quality* resulted in a dynamic behaviour of *Population potential*. In particular the inconsistent behaviours were caused in situations when either:

1) *Population potential* > *Numbers surviving* (i.e. *Net survival dynamic* is plus), δ *Population potential* is plus and is bigger than δ *Numbers surviving* which is also plus (due to I+ from *Net survival dynamic*) OR
2) *Population potential* < *Numbers surviving* (i.e. *Net survival dynamic* is minus), δ *Population potential* is minus and is less than δ *Numbers surviving* which is also minus (due to I+ from *Net survival dynamic*).

In these situations the resultant is that *Net survival dynamic* is either 1) plus and increasing or 2) minus and decreasing. The behaviour paths in this situation become inconsistent in a situation where the derivative of *Population potential* becomes steady. At this point the configurations of model fragments indicate that in:

Situation (1) *Net survival dynamic* should be plus and decreasing (as the difference between *Population potential* and *Number surviving* is now getting smaller because the value of *Population potential* is steady and the value of *Number surviving* is increasing due to the I+ from *Net Survival dynamic*), and in;

Situation (2) *Net survival dynamic* should be minus and increasing. In both cases this is an inconsistent behaviour as logically the derivative of *Net survival dynamic* must pass through "steady" to move from increase to decrease of vice versa.

The easiest solution to remove this inconsistent behaviour within the current configuration of model fragments was to eliminate them as possible behaviours in each situation (*Habitat quality > Number recruited*; *Habitat quality < Number recruited*; *Habitat quality = Number recruited*). This was done by constraining the possible derivative values of *Net survival dynamic* in situations for each of its magnitudes (minus, zero, plus) using a complex suite of model fragments.

4. CONCLUSIONS

Whilst the model implemented in GARP 3 to explore the salmon life cycle was able to explore scenarios related to single or pairs of life stages, it is currently unable to explore scenarios relating to the full life cycle. This is due to the complexity and size of simulations required to consider all life stages and their dynamics of changing recruitment and survival even if human activity (degradation and rehabilitation) is only considered for a single life stage in the life cycle. The majority of this complexity is due to the fact that the modelling approach allows some flexibility for the relative rates of changes for the *Number surviving* at each life stage, in response to the changes in the *Population potential* that is generated by the exogenous behaviours of rehabilitation activities, and their effect on habitat quality. The development of this complex model has led to research into potential developments in the Garp 3 modelling environment to aid the construction and simulation of large complicated models. In addition, the model raised fundamental issues in QR for the use of calculus of quantities with dynamic magnitudes. Ideally a general, and simpler, solution to model such a basic calculus concept is required.

ACKNOWLEDGMENTS

This research was funded by the European Commission's Sixth Framework Programme for Research and Development (project number 004074, project acronym Naturnet-Redime - http://www.naturnet.org, see Newsletter ISSN 1801 6480). Thanks must also go to Dr Tim Nuttle (University of Jena, Germany), Jochem Liem, Anders Bouwer and Floris Linnebank (University of Amsterdam) for there assistance in developing the case study above.

REFERENCES

Aprahamian M.W., Wyatt R.J. & Shields B.A., Use of biological reference points for the conservation of Atlantic salmon, *Salmo salar*, in the River Lune, North West England. *Fisheries Management and Ecology 13*, 21-30, 2006.

Bredeweg B., Salles P., Bouwer A., Liem J., Nuttle T., Cioaca E., Nakova E., Noble R., Rios Caldas A.L., Uzanov Y., Varadinova E. & Zitek A, Towards a Structured Approach to Building Qualitative Reasoning Models and Simulations. *Ecological Informatics* 2007, in press.

European Commission., Communication from the European Commission: a Sustainable Europe for a Better World: A European Union Strategy for Sustainable Development. Brussels, Belgium. 2001.

Guerrin F. & Dumas J., Knowledge representation and qualitative simulation of salmon redd functioning. Part I: qualitative modelling and simulation. *Biosystems 59*, 75-84, 2001a.

Guerrin F. & Dumas J., Knowledge representation and qualitative simulation of salmon redd functioning. Part II: qualitative model of redds. *Biosystems 59*, 85-108, 2001b.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Evaluating the potential of qualitative reasoning (QR) models to contribute to a sustainable catchment management

A. Zitek[a], S.Muhar[a], S. Preis[a], B. Bredeweg[b], P. Salles[c] and S. Schmutz[a]

[a] *Institute of Hydrobiology and Aquatic Ecosystem Management*
Department of Water, Atmosphere and Environment
University of Natural Resources and Applied Life Sciences,
Max-Emanuelstraße 17, 1180 Vienna, E-mail: andreas.zitek@boku.ac.at
[b] *Human Computer Studies, University of Amsterdam*
Kruislaan 419 (matrix 1), 1098 VA Amsterdam, The Netherlands,
E-mail: bredeweg@science.uva.nl
[c] *Universidade de Brasilia, Instituto de Ciências Biológicas, Campus Darcy Ribeiro,*
Brasilia - DF, 70.710-900, Brasil, E-mail: paulo.bretas@uol.com.br

Keywords: Qualitative reasoning; Model; Catchment management; Water Framework Directive; Sustainability; Stakeholder; Water abstraction; Energy; Fish; Evaluation.

Within the NaturNet-Redime project (NNR) Qualitative Reasoning (QR) models related to a sustainable catchment management were developed, following a general modelling framework (Bredeweg et al., 2007) and using newly developed Garp3 QR (Garp3, Bredeweg et al., 2008). Two models were developed in the River Kamp (Austria) case study (Zitek, 2006; Zitek et al., 2006): (A) Implementation of sustainable actions in a river catchment (stakeholder integration, quality of sustainability plans, development of ecological integrity and human well being, probability of catastrophic events), and (B) Hydropower production (water storage and release, water abstraction) and its effect on fish.

As a part of training and evaluation activities the potential of the QR models developed within the NNR project were evaluated by experts and students concerning usage in education and decision-making with regard to sustainable management of river catchments.

The models describe basic features related to the sustainable development of the River Kamp Valley in Austria. The catastrophic floods and inundations in August 2002 set new conditions for life and economy in the in the Kamp-valley facing flood control management, landscape architecture and land use planning with essential and future challenges. At the same time the question of an EU-Water Framework Directive (WFD) consistent treatment of the topics flood control/natural retention/prevention was arising. Consequently, the high water event finally represented a chance to develop the riverine landscape together with the local population as well as with the concerned scientific disciplines considering social, economic and ecological claims with regard to the EU-WFD. On this basis an overall integrated concept towards the sustainable development of the River Kamp landscape was developed at the University of Natural Resources and Applied Life Sciences, Vienna. Besides the consideration of the spatial scale (from catchment level up to planning onto municipalities) the interdisciplinary work of the different disciplines biology/nature conservation, landscape planning, water resources management, regional planning, agriculture and forestry and hydropower production is of central relevance for the project. Moreover, planning was conducted in participation with authorities, stakeholders and the local population. The integration of the population into the planning activities exceeded pure information policy with the possibility for the local population to actively participate in developing the future scenarios for their valley. Based on the experiences collected within the Kamp project, two models were developed.

Model A 'Sustainability Management' explores the different aspects of a sustainable catchment management. At the beginning of the process, the catastrophic event increased the fear of the population, which forced the politicians to become active for a more

sustainable development. The development of sustainability plans is modelled to be influenced by stakeholder participation (influenced by public information level, social development and accessibility to the decision making process), inclusion of scientific knowledge and preparedness of planners. The implementation of the plans itself (government action for SD) is assumed to be influenced by the quality of development plans, the resistance against measures (which is influenced by stakeholder participation) and personal interests that might cause a significant pressure to the implementation process. Sustainable actions are suggested to be positively related to human well-being and ecological integrity. The model is organized in 4 Sub-models: 'Community fear caused by a catastrophic event influences government action for SD', 'Stakeholder participation', 'Development and implementation of sustainability development plans', and 'Government influences ecological integrity of river basin and human well being'.

Model B explores important problems related to hydropower use in the Kamp valley and its effect on fish. Additionally the aspect of energy production, consumed energy and energy sold is modelled together with stakeholder satisfaction to represent the causal basis for hydropower use by stakeholders and the driving principle behind the tendency of increasing the amount of abstracted water to increase income and stakeholder satisfaction.

There are mainly two ways of influencing a river by hydropower use: (1) water abstraction and the creation of a residual or minimum flow stretch with the related effects to the physical environment like loss of water, loss of flow velocity reduction of depth and an increase of water temperature, and (2) the storage of water in a reservoir and a constant or peaking release of water from hypolimnetic parts of the reservoir which leads to decreased temperatures below the reservoir favouring cold water species and repressing the reproduction of warm-water species; if the water is on the one hand released at a constant rate this destroys the natural flow regime of a river, if released in a peaking mode ('hydropeaking') it affects fish due to he frequent changes of habitat conditions. Therefore model B focuses on the one hand the exploration of the two ways of hydropower use (direct use via a hydropower plant and abstraction to fill a reservoir with subsequent energy production). Additionally representations are developed that describe the effects of a reduced amount of water in the river (reduced flow velocity and increased temperature) and its effects on the fish fauna; both negative and positive effects of the changed physical environment on fish (favouring fish due to temperature increase or suppressing them; favouring fish with low requirements to flow velocity, the so called indifferent species, or suppressing species with high flow velocity needs, the so called rheophilous species) are captured in model fragments. In total eighth scenarios were developed and allow for a comprehensive analysis of the different effects of the different modes of hydropower production on sensitive guilds of the river type specific fish community of the river Kamp.

Both models were evaluated for their potential use in education and decision-making by students and experts. Generally a model evaluation basically covers 'validation and verification' of the model as well as the 'acceptance of the chosen approach and model' by the addressed stakeholder groups. Validation proves if the contents of the model are acceptable for its intended use, verification proves that the model is correctly implemented by a demonstration of its use. Proving the acceptance of stakeholders typically evaluates the potential of the model and the modelling approach for broader use.

Both evaluations, a general evaluation of Model A 'Sustainability Management' and an expert evaluations of Model A & B 'Water abstraction and Fish' yielded a very positive feedback with regard to the QR approach, the GARP3 software used to build models and the models themselves representing important issues related to the sustainable development of the riverine landscape Kamp. For example, using a 5-point scale (fully disagree, largely disagree, somewhat agree/disagree, largely agree, fully agree) most people 'largely or fully agreed' that QR models represent complex knowledge in an understandable manner and that QR and GARP3 can be seen as a valuable learning tool for understanding real world causal relationships related to a sustainable development of riverine landscapes. Also most people 'largely or fully agreed' that the presented QR models might significantly contribute to the understanding of students and stakeholders which entities and processes drive a sustainable development of a riverine landscape and therefore enhances their capability of making decisions. So the general aim, to produce software and models in QR language that

allow people to interact with and learn about sustainable development is clearly supported by the evaluation results.

Important additional statements related to the QR approach, the software and the models were also collected. Most interesting for the attendees was to see the interrelatedness of the system presented and the use of qualitative 'stock-flow' dynamics known from the System Dynamics approach. Only some added that they sometimes got a bit lost when confronted with the total view of the causal model describing a sustainable development of the Kamp valley. It was also stated that when showing these models to other user groups, their general ability to deal with complexity should be accounted for; meaning that for each user group the way of presenting the model should be adopted. Probably sometimes these models might be too complicated for certain stakeholder groups (people need to have some education e.g. to deal with complexity and causal relationships – to understand I's and P's for example, in a modelling approach like this). A high potential of an application of QR models in various fields, mainly in education but also in decision making and research was suggested by many participants. The potential of the GARP3 software and the QR approach to sustain collective, interactive social learning was clearly pointed out. Particularly, the identification of dependencies and causal relationships was seen as a prerequisite for understanding a system and therefore also for learning and decision making. With regard to a broader use of QR models in society especially for decision making it was stated, that it might take some time and engagement to establish approaches like that in society. (University) education using and teaching such approaches can be seen as an important basis for a further application. Summarizing, the QR approach developed within the NNR project was found to be able to capture and communicate mental models in structured way and therefore has a strong potential be used in education as well in the management processes themselves, especially for integrating different viewpoints of people within a 'mediated modelling process' (Van den Belt, 2004). This shared view can be seen as a pre-requisite for the development of common actions (Rogers & Biggs, 1999) and therefore represents the basis for a sustainable catchment management, the main target of the WFD.

ACKNOWLEDGEMENTS

The research presented in the paper is co-funded by the European Commission within the Sixth Framework Programme for Research and Development (2002-2006) (project NaturNet-Redime, number 004074, http://www.naturnet.org).

REFERENCES

Bredeweg, B., A. Bouwer, J. Jellema, D. Bertels, F. Linnebank and J. Liem (2007). Garp3 - A new workbench for qualitative reasoning and modelling. *4th International Conference on Knowledge Capture (K-CAP 2007)*, D. Sleeman and K. Barker (eds), pages 183-184, Whistler, BC, Canada, October 2007.

Bredeweg, B., P. Salles, A. Bouwer, J. Liem, T. Nuttle, E. Cioaca, E. Nakova, R. Noble, A.-L. Rios Caldas, Y. Uzunov, E. Varadinova and A. Zitek (2008). Towards a structured approach to building qualitative reasoning models and simulations. *Ecological Informatics*, (3)1: 1-12.

Rogers, K. and H. Biggs (1999). Integrating indicators, endpoints and value systems in strategic management of the rivers of the Kruger National Park. Freshwater Biology **41**(2): 439-451.

Van den Belt, M. (2004). Mediated modeling: a system dynamics approach to environmental consensus building. Washington-Covelo-London, Island Press 339.

Zitek, A. (2006). Textual description of the river Kamp case study focusing on basic ecological and socio-economic features for an integrative and sustainable development of the riverine landscape. Deliverable D6.6.1 of the NaturNet-REDIME project. Vienna, Institute of Hydrobiology and Aquatic Ecosystem Management, BOKU. **D6.5.1.**

Zitek, A., S. Muhar, S. Preis, S. Schmutz, B. Bredeweg and P. Salles (2006). QR models and documentation for learning about sustainable development, focusing on basic ecological and socio-economic features for an integrative and sustainable development of the riverine landscape of the Kamp valley. Deliverable D6.6.2 of the NaturNet-REDIME project. Vienna, Institute of Hydrobiology and Aquatic Ecosystem Management, BOKU. **D6.6.2.**

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Collaborative Capture and Reuse of Environmental Knowledge

J. Liem and B. Bredeweg

Laboratory Human Computer Studies Laboratory, University of Amsterdam
Kruislaan 419 (matrix 1), 1098VA Amsterdam, The Netherlands
{jliem,bredeweg}@science.uva.nl

Keywords: ADM1; Benchmark; BSM2; Modelling; Solid Separation Problems; Validation.

Introduction

Searching for existing models concerning a particular topic in environmental science and assessing its relevance is often experienced as being difficult. Actually reusing part of a model in a meaningful way brings along many more problems, and it therefore rarely done. However, an environment that would support such activities would enhance the collaborative capture of environmental knowledge by relieving researchers of redoing modelling work that has been done before by reusing earlier work.

Within the NaturNet-Redime project we have created a working environment around the qualitative modeling and simulation tool Garp3 (http://www.garp3.org) that supports the searching, assessing, and reuse of knowledge formalized in qualitative models. This paper describes the qualitative model library, the Web Ontology Language (OWL) formalization of the models needed for this library, and copy functionality in Garp3, which are part of this working environment. These features should increase the amount of reused model parts in the environmental modeling community.

Garp3 – QR Workbench

The Garp3 workbench implements a diagrammatic approach to modeling and simulating qualitative models. Modeling in Garp3 starts by creating model ingredient definitions. These definitions include entities, agents, assumptions, configurations, quantities and quantity spaces. Entities, which represent the structural objects in a system, are organized in a sub-type hierarchy. They are defined by their name and their position in the hierarchy. Agents and assumptions are defined in the same way. Agents cause influences from outside of the system, while assumptions are labels that indicate that certain conditions are presumed to be true. Configurations are structural relations between entities that are defined by their name. Quantities represent the features of entities and agents that change during simulation, and are defined by their name and a set of possible quantity spaces. Quantity spaces represent the possible values a magnitude (or derivative) of a quantity can have, and are defined by their name and an ordered set of possible values. Quantity spaces are associated to the quantities of entities or agents.

Next to the model ingredients defined by the modeler, there is also a set of predefined model ingredients. These include causal dependencies (proportionalities and influences), correspondences, the operator relations plus and minus, value assignments, and inequalities. All these model ingredient definitions can be used (instantiated) to create model fragments (MFs) and scenarios. MFs can be seen as composite ingredients that incorporate other ingredients as either conditions or consequences. They are organized in a subtype hierarchy, meaning that a child MF inherits the model ingredients of its parents. Furthermore, a MF can incorporate other MFs as conditional ingredients. MFs instantiated

in another MF are called Imported Model Fragments (IMFs). An example MF incorporating another MF twice is shown in Figure 1.

Scenarios are also composite model ingredients. They describe specific system situations. During simulation, MFs are sought which match on the scenario (i.e. the model ingredients fulfill the conditions of the MF). The consequences of matching MFs are merged with the scenario to create an augmented state from which the next states of behavior can be determined.

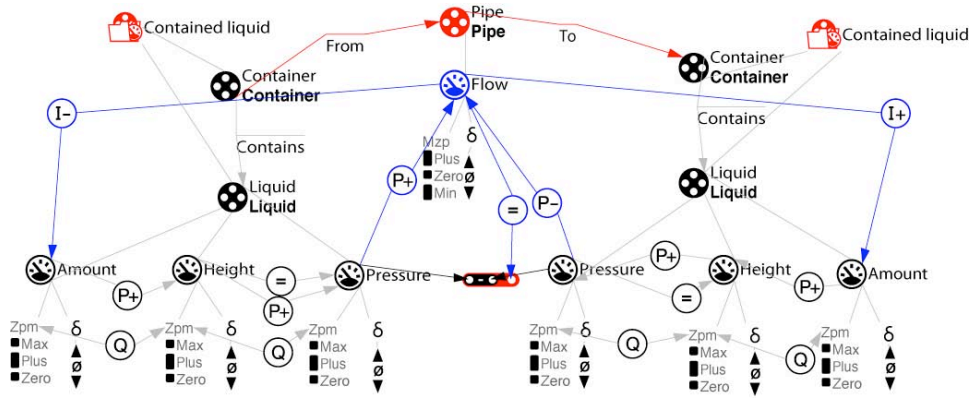


Figure 1. Liquid flow includes 2 Contained Liquid IMFs, a Pipe, and the configurations as conditions, and flow, its calculation and causal relations as consequences.

Model Repository

To reuse models of others, modelers have to be able to share their work and access work of others. This is solved by allowing models to be uploaded to and downloaded from a central online model repository. However, the number of models in the repository can potentially become large, which means that modelers need to be supported by search functionality to find reusable models.

Typically, a modeler will want to search for models which contain a certain entity or quantity (e.g. a model which contains both an entity *population* and a quantity *size*). Normal search engines search for keywords in text and are unable to interpret the explicit knowledge representation in qualitative models. So the search engine is unable to distinguish between different types of model ingredients, or between domain specific and domain independent knowledge (i.e. the QR vocabulary and the knowledge formalized by the modeler). This hampers the search engine's ability to find relevant models. A search solution should make use of the explicit knowledge representation in qualitative models to allow modelers to focus their search using the QR vocabulary.

The Semantic Web initiative proposes that "semantic search" becomes possible by making content machine-accessible [1]. The Web Ontology Language (OWL) is a description-logic based knowledge representation language, which is represented in RDF/XML, and is being developed as part of the Semantic Web initiative. It has become the de-facto standard for the sharing of knowledge on the web in the form of *ontologies*. By formalizing qualitative models as OWL ontologies, the models become interpretable by OWL search engines, and searching for models in which certain model ingredients or certain structures are used becomes possible. Additionally, the formalization of models in OWL opens up the possibility for other QR-tool developers to implement functionality to import these files. This could potentially make models accessible to communities using different QR tools.

A qualitative model repository was implemented (see Figure 2) as a webpage that allows modelers to share their own models as OWL files, and search and download models of others. The search functionality shows the model ingredient definitions of all the models. Selecting a definition reduces the list of matching models, allowing the modeler to iteratively refine the list of potentially useful models. An initial assessment of the

usefulness of the models can be done by reading their metadata. Chosen models can be downloaded and investigated in Garp3.

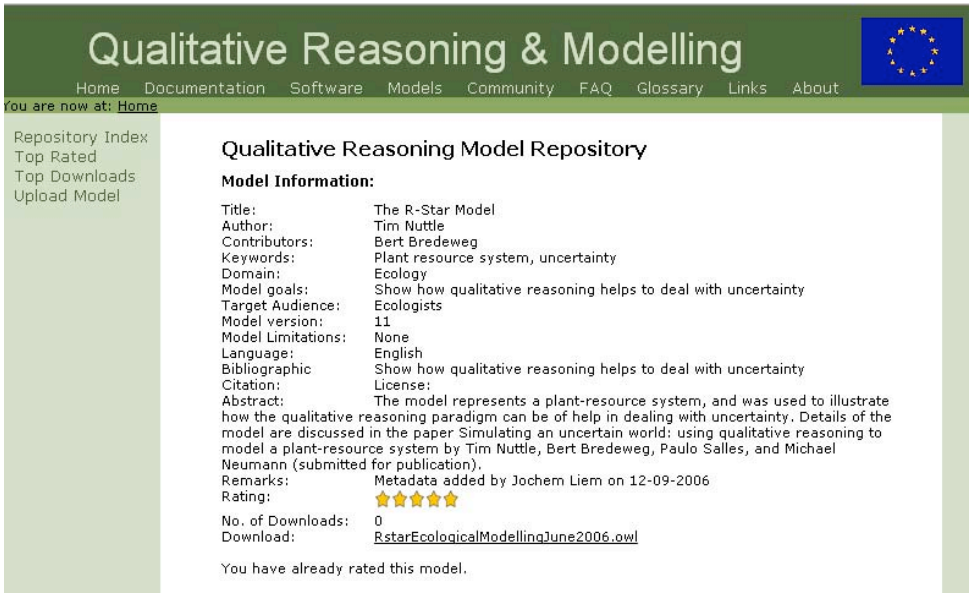


Figure 2: The Qualitative Model Repository.

Reuse through Copy Functionality

The Garp3 workbench allows users to open multiple models at the same time, and copy model ingredient definitions, model fragments and scenarios from one model to the other. There are several issues that have been addressed to make this possible. Firstly, model parts often depend on other model parts without which they are incomplete. Garp3 ensures the completeness of the copied knowledge by creating a minimal self-contained sub-model each time a model part is copied. Secondly, syntactical correctness of models is maintained automatically. Since Garp3 is a diagrammatic tool that automatically ensures syntactical correctness, this correctness is maintained when copying model parts by creating the model parts in the next model as if they would be created from scratch. Thirdly, no redundant knowledge is added when model parts are copied, as similar model parts already existing in a target model are reused. Fourthly, the semantics of the model parts is preserved as much as possible, putting entities, agents and assumptions as close below their parents in the sub-type hierarchy as possible. Finally, existing model parts in a target model are not altered. This assures that alterations to existing simulations of the model can be easily restored by deleting the new model parts. Copied model ingredients with clashing names are created with the suffix '(copy)' behind them. The copy functionality makes it easier to reuse model parts from existing models.

ACKNOWLEDGEMENTS

This work is co-funded by the European Commission within the Sixth Framework Programme (2002-2006), project NaturNet-Redime (http://www.naturnet.org), number 004074.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A new approach for river exploration : River Mesta case study for qualitative modeling sustainable development

E. Nakova[a], B. Bredeweg[b] and P. Salles[c]

[a] *Central Laboratory of General Ecology, Bulgaria, E-mail: elena_nakova@mail.com*
[b] *Human Computer Studies Laboratory, University of Amsterdam, The Netherlands, E-mail: bredeweg@science.uva.nl*
[c] *Institute of Biological Sciences, University of Brasília, Brazil, Email: psalles@unb.br*

Keywords: Qualitative reasoning; Sustainability, Dissolved oxygen, water body.

The general objective of NaturNet-Redime project is to support sustainable development by improving knowledge about all aspects of sustainability and provide education mainly about environmental, social and economic tools for implementation of the EU strategy on sustainable development at both EU and international levels.

This paper describes the progress towards developing a qualitative reasoning model of sustainable development issues in River Mesta Bulgaria, following a standardized framework for conceptual description of qualitative reasoning case studies. (Bredeweg, 2006).

In terms of the Water Framework Directive 2000/60/EC (Annex XIA), River Mesta (RM) is located in the eco-region 7 (Eastern Balkans). It is a transboundary river shared between Bulgaria and Greece.River Mesta has been intensively studied for the last 35 years due to the heavy organic pollution caused by yeast, cellulose and cardboard manufacturing in the town of Razlog. The region is recognized as economically under-developed, with high unemployment. Regional development plans focus on intensifying economic activities based mostly on natural features of the region. This includes further development and diversification of tourism; modernizing and intensifying agriculture and forestry; increasing energy production from hydropower; construction of new roads and streets, and enhancing infrastructure like sewage systems, wastewater treatment plants, and domestic waste landfills.

One of indicative parameters of aquatic ecosystem health is the amount of dissolved oxygen (DO) in the water. Oxygen is an essential component for all living organisms in the aquatic ecosystem. Water pollution, abstraction, erosion and other human activities can disrupt this balance, worsening ecosystem health and decreasing sustainable use of ecosystem services. The fluctuation of DO – natural and anthropogenic, is potentially of great importance for decision making about sustainable and integrated management of aquatic ecosystems. QR reasoning provides a modeling paradigm that allows explicit representation of the various processes that interact in a water body to affect to affect DO(Bredeweg and Struss 2003).

REFERENCES

Bredeweg, B.; Bouwer, A.; Jellema, J.; Bertels, D.; Linnebank, F.; Liem, J. (2006) Garp3 – A new Workbench for Qualitative Reasoning and Modelling. Proceedings of the 20th

International Workshop on Qualitative Reasoning (QR06), p. 21-28 Hanover, New Hampshire, US, July, 2006.

Bredeweg, B., and P. Struss. 2003. Current topics in qualitative reasoning. AI Magazine 24:13-16.

Nakova E., Nuttle T., Bredeweg B., Salles P. - A Garp3 Model of Environmental Sustainability of River Mesta. Proceedings of the 21th International Workshop on Qualitative Reasoning (QR07), ABERYSTWYTH, U.K. , US, June, 2007.

Uzunov, Y., and E. Nakova. 2006.QR models and documentation for learning about

sustainable development, focusing on basic biological, physical and chemical processes related to the environment in River Mesta- Actual STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D6.2.3.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
Workshop on Qualitative Reasoning models on Sustainability
Barcelona, July 6-10, 2008

A library of model fragments on sustainability

P. Salles[a], B. Bredeweg[b] E. Cioaca[c], E. Nakova[d], A. Zitek[e] and R. Noble[f]

[a]*Institute of Biological Sciences, University of Brasília, Brazil*

[b]*Human Computer Studies Laboratory, University of Amsterdam, The Netherlands*

[c]*Danube Delta National Institute, Romania*

[d]*Central Laboratory of General Ecology, Bulgaria*

[e]*Vienna University of Applied Life Sciences, Austria*

[f]*University of Hull, UK*

Keywords: Qualitative reasoning; Perspectives; River basin management; Sustainability

This paper describes the implementation of a Library of model fragments that integrates different models and captures essential concepts for understanding sustainable development related issues, presented in meaningful ways as simulations (Salles and Bredeweg, 2007). Input models to the Library come from the NaturNet – Redime[1] Project case studies and from models about the Millennium Development Goals (Salles, 2005). As the Library includes a broad range of knowledge from different disciplines and a large number os scenarios and model fragments, it is necessary to create handles to explore the Library and to select clusters of concepts that may be used in curricula for learning about sustainability.

Literature on Qualitative Reasoning reports the use of *perspectives*, implemented by means of *modelling assumptions*, as a possible solution for this type of problem (Falkenhainer and Forbus, 1991). A perspective defines a subset of simulation models that can be created to achieve specific goals, that is, to answer questions of a particular type. From this point of view, perspectives serve an organizational function as they guide the modeller in selecting appropriate assumptions, structural relations and scenarios. Perspectives are therefore useful for defining and constraining a simulation, taking into consideration certain aspects of the encoded knowledge while ignoring the rest. When the model user selects a certain perspective, (s)he provides the context for exploring different parts of the library. Depending on which perspective is adopted, different entities, quantities, values, and causal relationships are included in the simulation.

Conceptually, modelling assumptions fall into two categories: simplifying and operating assumptions. *Simplifying assumptions* are used to make explicit how knowledge details such as the underlying perspective, approximations and level of granularity are represented in the model fragments. They provide the vocabulary to be used in the model, explicating what kinds of things exist and what sort of relationships can be held between them, and provides alternative representations for the same thing. *Operating assumptions* are used to manage complexity. In a way, they provide focus to the simulation, by constraining the model to describe relevant behaviour for answering specific questions. Typical operation assumptions include restrictions on quantity values implemented by means of inequalities between quantities and constants (e.g. *number_of* >0) and steady-state assumptions, that determines all derivatives for some class of parameters to have value zero.

Both the NaturNet – Redime case models and the Library were implemented in Garp3 (Bredeweg *et al*., 2006), using facilities for collaborative modelling, such as the sketch environment, and elements from the process-centred ontology (Forbus, 1984). Technically speaking, perspectives were implemented by explicitly representing assumptions and by using other modeling primitives, such as hierarchies of entities and model fragments, attributes, alternative quantity spaces for key quantities and alternative representations of key concepts.

[1] NaturNet-Redime, EU STREP, project number 004074. Available at www.naturnet.org.

The Library of model fragments comprises, in its current implementation, 112 entities, 1 attribute, 60 configurations, 201 quantities, 22 quantity spaces, 24 agents, 45 modelling assumptions, 202 simulation scenarios and 414 model fragments. So far, it is the biggest Qualitative Reasoning model of this kind. The Library can be explored in many ways. Two of them are discussed here: (a) by creating perspectives based in the input models, as they were presented, and (b) by creating thematic perspectives. Table 1 summarizes the perspectives created in this modelling effort.

Types of perspective	Perspectives
Case study-based	Natural ; Rural ; Semi-urban ; Urban ; Natural resources exploitation ; Natural environment rehabilitation ; Social
Thematic based	(I) Natural systems; (II) Natural disasters; (III) Human explores natural resources; (IV) Environmental effects of human activities (in interaction with natural factors); (V) Energy; (VI) Economy; (VII) Education and trainning; (VIII) Science and Technology; (IX) Legislation; (X) Stakeholder participation; (XI) Governmental plans and activities; (XII) Management actions for sustainability; (XIII) Human health; (XIV) Human well being

Table 1. Perspectives created for the Library of model fragments on sustainability.

Case-study based perspectives: Taking the case study models, seven high level perspectives to approach sustainability issues were developed, described as follows (between square brackets, the input model): (a) *Natural*: Basic case models exploring nutrient cycling, food webs, agriculture and industrial pollution, dissolved oxygen in water bodies [*Danube Delta (Dd), River Mesta (Rm)*]; (b) *Rural*: Model exploring deforestation and erosion, soil fertility, water resources, biodiversity and agricultural (crops and cattle) activities [*Rural Riacho Fundo (Rfru)*]; (c) *Semi-urban*: Model exploring urbanization, including soil erosion and the degradation of springs and streams, infiltration and underground water, and water resources in industrial activities (textile and food industries) [*Semi-urban Riacho Fundo (Rfsu*)]; (d) *Urban*: Model compares the urban situation in absence and in presence of an engineered drainage system with respect to the main consequences: floods, economic damage, transported garbage, water related diseases and the human well being [*Urban Riacho Fundo (Rfur*)]; (e) *Natural resources exploitation*: Water abstraction model explores the effects of water abstraction to produce energy on the amount of water, on fish populations; Deforestation model presents the effects of deforestation on the use of biodiversity, erosion and water resources; Global warming model presents the effects of pollution caused by petroleum and solid fuels burning on human health and global warming [*Water abstraction (Aw), Deforestation (D), Global warming (Gw*)]; (f) *Natural environment rehabilitation*: This model explores the dynamics of fish populations and the benefits from improving natural habitats and stocking [*Salmon and Bream (Uk)];* (g) *Social*: This model explores a number aspects related to stakeholder participation in the decision-making process, including legislation, governmental actions, sustainability plans, education and technological solutions [*Stakeholder participation (Sp*)].

Thematic-based perspectives: A thematic-based approach to the Library exploitation was created in order to expand the options available to the users. It is very likely that the users have problems to solve, and the case studies may address only part of these problems. Following this intuition, 14 perspectives were created to address a number of classes of problems. When more than one simulation has to do with the theme, the user would be advised to search, in different input models, the simulations that are associated to the theme by the modellers (Table1). Apparently this is a good option, because this approach favours the integration of knowledge about sustainability. For example, the perspective (I) *Natural systems*, addresses topics such as rivers; vegetation; macrophytes; blue green algae; diatoms; phyto and zooplankton; animals; population growth; biodiversity; food webs; soil integrity; fertility; and nutrients. Simulations, in this case, come from the models [*Dd, Rm, Rfru, Sp, Uk*]. Another interesting example is the perspective (VI) *Economy*: Mechanisms involving resources offer and consumption, production rate, products, residues and jobs in the three sectors (agriculture, industry and services) of economic activities; different types of resources used in

economic activities; different types of residues and pollutants produced by economic activity; (un)balanced proportions of products and residues; influences on GDP of technological innovation exploring biodiversity, uses of water resources and agriculture. The simulations concerning this perspective come from the models [*E, Dd, D, Rfru, Rfsu, Rfur*]. The interested reader can find details of these 21 perspectives and the simulations included in each group in the Deliverable D6.7.1 (Salles and Bredeweg, 2007), available in the NaturNet – Redime portal (www.naturnet.org).

The use of assumptions for reasoning with multiple models has a long tradition in Qualitative Reasoning. Most of the work on this topic focus on algorithms for automatically selecting or changing models according to certain assumptions. Differently from these previous approaches, the work described here uses no numerical information or mathematical functions to define perspectives or to implement assumptions. Using only qualitative representations of concepts, Garp3's representational apparatus creates alternative models according to the perspective taken. The first element used to create a simulation model are the entities. Increasing levels of complexity can be further obtained by means of the inclusion of new entities in the system structure. In fact, control over entities and quantities is a quite effective use of simplifying assumptions to implement perpectives. Considering that each entity is associated to a number of quantities, quantity spaces, qualitative values and qualitative states of the entity, the choice of entities and the associated elements define specific vocabulary for a certain perspective. For example, entities 'Crop', 'Cattle', 'Garage', 'Food industry' and 'Textile industry' introduce vocabulary to describe the type of residues that make economic activities unustainable (pesticides, manure, oil, organic matter and chemical pollutants, respectively). In many situations, alternative representations of the same phenomenum provide different levels of details to simulations (for example, erosion is more detailed in *Semi-urban* than in *Rural* perspective). Operating assumptions reduce complexity and give focus to the simulations. For example, correspondences between quantity magnitudes and derivatives were largely used in all the perspectives. Also, steady state conditions are useful to isolate certain parts of the causal model and, as such, to reduce the size of the simulations. Focus was provided by different means. For example, soil fertility can be determined in three ways: (a) by assuming that the quantity *Fertility* values correspond to *Nutrient* values; (b) by considering that vegetation cover determines the amount of organic matter, and calculating *Fertility* = *Organic matter* + *Nutrient*; and (c) by considering the combination of nutrients and manure, a by-product of livestock. This way, different causal chains can be constructed within the *Rural* perspective, focussing on different aspects of agriculture.

The Library provides a broad coverage to sustainability aspects. From the technical point of view, a perspective-taken approach seems to be effective to organize knowledge and simulations about sustainability. The use of assumptions was essential to integrate independent models. All in all, the Library is a good outcome of NaturNet-Redime. However, this work can be improved in many ways. Ongoing work focus on the optimization of model fragments and entities hierarchies, and in the integration of perspectives, so that new simulations may become possible. A point to be addressed in the future is shifting from one perspective to another. Identifying the requirements for such transitions will lead to better understand the nature of perspective-taking in qualitative reasoning. Concluding, the Library now can be used in the development of curricula about sustainability and, therefore, to meet one of the most relevant goals of NaturNet-Redime, to support stakeholders in finding means to make their activities more sustainable.

References

Bredeweg, B.; Bouwer, A.; Jellema, J.; Bertels, D.; Linnebank, F.; Liem, J. (2006) Garp3 – A new Workbench for Qualitative Reasoning and Modelling. *Proceedings of the 20th International Workshop on Qualitative Reasoning (QR06)*, p. 21-28 Hanover, New Hampshire, US, July, 2006.

Falkenhainer, B. & Forbus, K. (1991) Compositional Modeling: Finding the Right Model for the Job. *Artificial Intelligence*, 51(1-3): 95-143.

Forbus, K. (1984) Qualitative Process Theory. *Artificial Intelligence*, 24: 85-168.

Salles, P. (2005) Qualitative Representations of Indicators of Environmental Sustainability of the Millennium Development Goals. *Proceedings of the Nineteenth International Workshop on Qualitative Reasoning (QR05)*, p. 78-88. Graz, Austria, 18-20 May 2005.

Salles, P. and Bredeweg, B. (2007) *Integrated library of Qualitative Reasoning model fragments*. Deliverable D6.7.1, NaturNet-Redime, EU STREP, project number 004074.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A Qualitative Reasoning–based Curriculum for learning about Sustainable Development

Tim Nuttle[a] **and Anders Bouwer**[b]

[a] *Biology Department, Indiana University of Pennsylvania, 114 Weyandt Hall, Indiana, PA 15212 USA (nuttle@iup.edu)*
[b] *Institute of Ecology, Friedrich Schiller University, Dornburger Strasse 159, 07743 Jena, Germany (andersbouwer@mac.com)*

Keywords: model, simulation, evaluation, sustainability, learning, river.

Qualitative Reasoning (QR) has many qualities making it a useful ontology for education about system dynamics (Bredeweg and Forbus 2003). We implemented an online curriculum for learning about sustainable development using QR in Moodle, an online course management system (http://www.moodle.org). This curriculum is centred on qualitative reasoning models that explore sustainable development situations in a five case studies (Bredeweg et al. 2007), in addition to other models that explore specific sustainability issues. Each model is implemented using the QR model-building and inspecting workbench, Garp3 (Bredeweg et al. 2006). The "lesson" activity in Moodle is used to create a series of questions that guide learners stepwise through different scenarios for each model. Additionally, the "Glossary" resource tool in Moodle is used to provide information that helps learners with help tips and definitions of SD and QR terms. Web pages in Moodle are used to present background information and suggest themes and learning routes for learners to explore the content of the curriculum.

We evaluated the structure of the Lesson activities for one of the modelled scenarios from one of the SD case studies. Each Lesson for each scenario is structured as follows:

1. background about the case study;
2. introduction and learning goals for the scenario;
3. exploration of the system structure;
4. exploration of causality of the system;
5. exploration of the dynamics of the system;
6. application to real world situation and to sustainable development.

Hence, questions are organised to move from structure to causality to dynamics to evaluation, while also achieving the six levels of cognitive ability described in Bloom's Taxonomy (namely: knowledge, comprehension, application, analysis, synthesis, and evaluation; Anderson et al. 2001). Each question is accompanied by a hint link, which instructs the learner how to interact with the model in Garp3 to discover the answer to the question. The Moodle tool collects learners' typed answers to the questions and reports them to the instructor (if desired). Before moving on to the next question, learners are able to compare their answer with an answer provided by the instructor. This allows learners to assess whether they have understood the intended content before moving on to higher-level questions. Specific questions for the evaluated lesson are provided below. Students were also asked about their impressions of the lesson format and using QR models to learn about sustainability.

Setting

We evaluated one of the lessons on the Curriculum page, as implemented in Moodle, with a university Botany class of 19 students. The class was composed of students ranging from second- to fourth-year Biology majors at Indiana University of Pennsylvania, USA. Students gave informed, written consent to participate in the study. None of the students had previously had a course in ecology or sustainable development. Some of the students had worked with concept maps (Cañas et al. 2003), but none had experience with qualitative reasoning models.

Methods

We presented the students with River Kamp Lesson 1 from the online curriculum. Although the curriculum has been designed with the idea in mind of learners working through the material on their own computer, this study investigated the use of the curriculum as standalone teaching material for gauging and expanding learners' understanding and stimulating discussion of problems related to sustainable development. In this case, only the teacher had a computer to show the materials as used the model as a demonstration tool. Hence we displayed the appropriate screen shots as if the students had run the model themselves, using a computer and video projector. Students were allowed to view the diagram and write their responses until it looked like most had finished writing. Then, the facilitator solicited student responses to the question. If it looked like they didn't understand how to answer the question, the presenter showed the appropriate help tip. After a few students had responded, the facilitator presented a prepared "correct" answer before moving on to the next question.

Results

Based on the written responses to questions in the lesson, all students made a good effort to answer the questions to the best of their ability, despite some technical difficulties with reading the displayed diagrams from across a large room.

Some students at first had difficulty grasping the meaning of some of the model diagrams during the first few questions. For example, instead of basing their answers on the diagram for question 2, several students based their answers on their background knowledge about natural catastrophes. This may have been due to their inability to access the help tips that would ordinarily be available if they were working on their own computers, which would guide the students in how to go about answering the question based on what is presented in the model. However, once they had the opportunity to hear responses of other students and the class went over the "correct" answer provided in the Lesson, most were able to catch up and were able to understand and appropriately respond to later questions. Students successfully used the diagrams to reason about a complex system involving causal feedback loops and multiple possible outcomes, including cyclic behaviour. For example, many students correctly predicted that multiple outcomes were possible and that the system would never stabilize (questions 6, 7, and 8). Finally, students responded for questions 9 and 10 that people could learn the government should institute policies that maintain sustainable actions even when community fear level is low. They also named several types of sustainable actions (e.g., maintaining natural ecosystems in the watershed to reduce flood levels) and unsustainable actions (e.g., building too close to the river or in areas that lie too low).

Students were also asked about their impressions of the lesson format and using QR models to learn about sustainability. Many students responded that they thought the simulation model provides useful support for learning about the behaviour of complex systems. Several students suggested more use of colour in the presentation, perhaps to hold their attention better or draw their attention to specific parts of the various diagrams. In fact, colour is used in systematic ways throughout the software to denote the ontological type or status of particular information elements in the diagram (i.e., actual values and derivatives are highlighted in red; for more on this issue, see Bredeweg et al. 2006). However, the students' comments indicate that there may be room for improvement in making the diagrams more attractive to this target group of users, or perhaps in including additional pictures or illustrations that bridge the gap between photographs and the diagrams as currently used. Again, when students interact with the models themselves on their own

computers, this will probably be less of an issue because they will be moving screens around, be able to view photographs from other pages if they want, and can access help tips to draw their attention to the focal points of each diagram.

Concluding remarks

Overall, students responded positively to the experience and several expressed interest (orally or in their written evaluations) in participating in future activities using the online curriculum. We are currently planning additional evaluation of the curriculum using this and other groups of students, where each student will have the opportunity to complete lessons on individual computer work stations.

ACKNOWLEDGEMENTS

The authors wish to thank the student who volunteered to assist in evaluation of the curriculum. This study is a contribution from the NaturNet-Redime project, which was co-funded by the European Commission FP-7.

REFERENCES

Anderson, L. and D. Krathwohl, A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. 2001, New York, New York, USA: Longman.

Bredeweg, B., et al. Garp3: A new workbench for qualitative reasoning and modelling. in Third Model Based Systems Workshop. 2006 (in review).

B. Bredeweg, P. Salles, A. Bouwer, J. Liem, T. Nuttle, E. Cioaca, E. Nakova, R.A.A. Noble, A. L. R. Caldas, Y. Uzunov, E. Varadinova, and A. Zitek. Towards a structured approach to building qualitative reasoning models and simulations. Ecological Informatics, In Press, available online 16 April 2007.

Bredeweg, B. and K. Forbus, Qualitative modeling in education. AI Magazine, 2003. 24(4): p. 35-46.

Cañas, A.J., Coffey, J.W., Carnot, M.J., Feltovich, P., Hoffman, R.R., Feltovich, J.,

Novak, J.D. (2003) A summary of literature pertaining to the use of concept mapping techniques and technologies for education and performance support, Technical report, The Institute for Human and Machine Cognition, Pensacola FL, USA, http://www.ihmc.us (visited August 31st, 2005).

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Danube Delta Biosphere Reserve Model of Environmental Sustainability based on Qualitative Reasoning Concept

E. Cioaca[1], B. Bredeweg[2] and P. Salles[3]

[1]Danube Delta National Institute, Tulcea, Romania (eugenia@indd.tim.ro)
[2]University of Amsterdam, Amsterdam, The Netherlands (bredeweg@science.uva.nl)
[3]University of Brasília, Brasilia, Brazil (paulo.bretas@uol.com.br)

Keywords: Qualitative reasoning, Garp3 workbench, Application of QR technology, Sustainable Development.

The paper presents a Qualitative Reasoning (QR) model of sustainable development issues in the Danube Delta Biosphere Reserve (DDBR) in Romania (Cioaca et al., 2006; 2007). The QR concept stresses the cause-effect relationships among the system components that contribute/hamper the system's sustainable development (Bredeweg et al., 2006). The model construction followed a standardized framework for conceptual description of QR case studies (Bredeweg et al, in press). Based on our expert knowledge, the main cause-effect relationships were identified between the DDBR aquatic ecosystems abiotic (nonliving) and biotic (living) components that hamper its sustainability. Most important is the negative effect of water pollution from the Danube catchment area on aquatic biota and human health in and around the DDBR. We present essential background knowledge about the system, and describe how available knowledge was encapsulated into the QR knowledge structures, including model fragments and scenarios. The model outputs are the simulations (based on the scenarios) and they contribute to understanding the factors that affect the sustainability of the DDBR.

The objective of this contribution is to present a description of the DDBR model system, including the main model goals, the system global behaviour, the main model ingredients as they are implemented into the specialized QR workbench, Garp3 (representative *scenarios* and *model fragments*), and the scenario simulation results. The QR models will be used by end users for learning about specific conditions to be fulfilled by the system (either social, economic, or environmental) in order to contribute to increased public involvement as called for in the Strategy for Sustainable Development.

The DDBR, located at the mouth of the Danube River before it reaches the Black Sea, has been designated as a World Heritage Site and Wetland of International Importance since 1990 (according to The Ramsar Convention on Wetlands, signed in Ramsar, Iran, in 1971). Its area of 5,800 sq. km, making it one of the greatest wetlands in the world, contains 30 types of terrestrial and aquatic ecosystems, of which 23 are natural or artificially modified and 7 are man-made ecosystems, including human settlements (Oosterberg et al., 2000).

The DDBR's status as a biosphere reserve dictates that all social and economic actions must comply with the *biodiversity conservation and protection measures*. Thus, the most appropriate concept of sustainable development for DDBR can be expressed by development through biodiversity, where most flora and fauna species are protected both to meet obligations of international conventions, but also to serve as natural resources for social and economic development of the region.

Scientists from DDBR met with local stakeholders to determine threats to conserve and develop these resources within the DDBR. The stakeholders involved in DDBR management include: nature conservation and protection bodies (DDBRA, NGOs), fishery

and fishing companies, tourism companies, fluvial and marine transport companies, and recreational hunting groups. Stakeholders identified the following threats:

- Decline in biodiversity (number of species) over the last several decades
- Contamination of water and fish from pollutants
- Concern about contamination in humans
- Reduction of fish diversity and abundance.

Contamination of water and fish from pollutants also contributes to health problems in humans. Contaminants come in basically two forms: heavy nutrient loads from agricultural fertilizers and heavy metals from industry. In both cases, most of the pollutants originate from far upstream in the vast Danube River catchments. Heavy nutrient loads lead to algal blooms, which can result in toxic by-products form algae as well as depletion of oxygen in the water when algae die and are degraded by bacteria. This can cause die-offs in fish. Heavy metals in the DDBR waters threaten human populations in two ways, first from direct consumption because many people drink untreated water directly form the DDBR waterways, and second from consumption of fish which bio-accumulate heavy metals (Otchere 2003; Wachs 2000) in their muscle tissues.

Model Goal The main goal of the DDBR model is to understand and emphasize connections between water pollution in the Danube River catchment basin and health of human population living in and around the DDBR.

System Selection and Structural Model The DDBR full structural model contains both terrestrial and aquatic ecosystems. It depicts a broader perspective on the entities and relations between them in the DDBR. The main system entities included in the QR model are *Water*, *Fish*, and *Humans*.

The main physical, physical-chemical and biological processes, influencing aquatic organism group behaviour, in the framework of their Functional Feeding Group relationship, and humans (living in or around DDBR) are: water flow; water eutrophycation - as result of Nutrients increase (mainly Nitrogen and Phosphorous compounds); phytoplankton bloom - overgrowth of algae and cyanobacteria (most of them are poisoning species); water pollution - mainly with nutrients, heavy metals, and cyanobacteria; fish growth, and human being health.

QR model components The QR model consists of modules which structure the modelled system: *Entities*, *Attributes*, *Configurations* (structural relationships between Entities), *Quantities* associated to each *Entity*, *Quantity Spaces* associated to *Quantities*, *Scenarios*, *Model Fragments*, *Agents* (External influences), and *Assumptions*.
The main DDBR QR model components are: 18 Entities hierarchical structured, 17 Scenarios, and 57 Model fragments.

Scenarios A scenario describes the scope of a system to be modelled. It includes the *Entities*/*Agents* involved in the modelled process, *Configurations* between *Entities* /*Agents*, *Entity*/*Agent Quantities* with initial values, and *Assumptions* (if necessary). This structure shows a possible start situation of the modelled process from which changes in the quantity values can be triggered, describing certain behaviours of the system.
The 17 DDBR QR model Scenarios are constructed as to capture:

- The aquatic *Plant* growth process for different initial values of *Nutrient inflow* in the modelled system;
- The aquatic *Animal* growth process;
- Both *Plant* and *Animal* growth process in the framework of the Functional Feeding Group relationship;
- The *Water pollution* process;
- The *Human being* behaviour from *Human health* point of view.

Model Fragments There are three types of Model Fragments (MFs): Static, Process, and Agent. They capture (static) behavioural knowledge about the system, define the system behavioural characteristics related to a process, and introduce the system's external

influences (Agents), respectively. The 57 DDBR QR model MFs implement the structural and behavioural relationships between Danube Delta and both the Aquatic population and Human beings, related to Danube Delta: Water pollution influence on these two system components behaviour.

QR model outputs - Scenario Simulation Results The results of simulating a scenario constitute the model outputs. These are:

- Dependency diagrams. The DDBR QR model Dependency diagrams show the aquatic ecosystem components involved in physical, chemical and biological processes, the relationships among them and the values the Quantities get as result of given Quantity values within initial states. They relate to *Plant* and *Animal Growth* process, *Water pollution* process and its negative effect on both *aquatic populations* and *Human beings* living in and around the DDBR.
- Global State-graphs - as result of full scenario simulation - show all possible system components connections (behaviour) generated by the initial states of the system, by means of paths of states. They show where the process starts (initial state) and where it ends (end-state for a path of states).
- The Value history diagrams of global behaviour paths show the values the Quantities get along the paths of states.
- The Equation history shows inequalities between some system component quantity values, those quantities able to start the process.
- Active MFs in initial and end states.
- Diagrams with Quantities and Quantity values for each state of a process.

These results help end users to understand the modelled system's functional components, causal relationships and the relevant factors affecting the system's sustainability.

REFERENCES

Bredeweg, B., Bouwer, A., Jellema, J., Bertels, D., Linnebank, F. and J. Liem. J. (2006) Garp3 - A new Workbench for Qualitative Reasoning and Modelling. *Proceedings of 20th International Workshop on Qualitative Reasoning*, C. Bailey-Kellogg and B. Kuipers (eds), Hanover, New Hampshire, USA, 10-12 July, pp. 21-28.

Bredeweg, B., Salles, P., Bouwer, A., Liem, J., Nuttle, T., Cioaca, E., Nakova, E., Noble, R.A.A., Caldas, A.L.R., Uzunov, Y., Varadinova, E. and Zitek, A. (in press). Towards a structured approach to building qualitative reasoning models and simulations. *Ecological Informatics* (available online since 16 April 2007).

Cioaca, E. , Bredeweg, B., Salles, P., Nuttle, T. (2007) - A Garp3 Model of Environmental Sustainability in the Danube Delta Biosphere Reserve based on Qualitative Reasoning Concept. *Proceedings of the 21st International Workshop on Qualitative Reasoning*, Chris Price (eds.), Aberystwyth, Wales, U.K., 26-28 June, pp. 7-16.

Cioaca, E., Bredeweg, B., Salles, P., Nuttle, T. (2006) The Danube Delta Biosphere Reserve: a Case Study for Qualitative Modeling of Sustainable Development. *Proceedings of the 20th International Workshop on Qualitative Reasoning*, C. Bailey-Kellogg and B. Kuipers (eds), Hanover, New Hampshire, USA, 10-12 July, pp. 10-12.

Oosterberg W., Staras M., Buijse A.D., Bogdan L., Coops H., Constantinescu A., Ibelings B.W., Hanganu, J., Menting G.A.M., Navodaru I. and Torok, L. (2000) *Ecological Gradients in the Danube Delta Lakes - Present State and Man-induced Changes*. Technical Report, RIZA rapport 2000.015, ISBN 90.369.5309x, Institute for Inland Water Management and Waste Water Treatment (the Netherlands), and the Danube Delta National Institute (Romania).

Otchere, F.A. (2003) Heavy metals concentrations and burden in the bivalves (Anadara (Senilia) senilis, Crassostrea tulipa and Perna perna) from lagoons in Ghana: Model to describe mechanism of accumulation/excretion. *African Journal of Biotechnology*, 2(9): 280-287.

Wachs B. (2000) Heavy metal content in Danubian fish. *Arch. Hydrobiol. Suppl.* 115(4): 533-566.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
Workshop on Qualitative Reasoning models on Sustainability
Barcelona, July 6-10, 2008

A qualitative model about the Island Biogeography theory

Débora Leite Silvano e Paulo Salles

Post-graduate program in Ecology, University of Brasília,
Campus Darcy Ribeiro, Brasília, 70.910-900, Brazil
{dsilvano; psalles}@unb.br

Keywords: island biogeography; qualitative reasoning; ecological models

Proposed by MacArthur and Wilson [1963; 1967], the Island Biogeography is a well established ecological theory used in the account of species diversity in communities. Within this theory, immigration and extinction are seen as the major driving forces in the introduction of new species from a source of colonizers (the 'continent') into a pool of species in the destination (the 'island'). Immigration to the island is determined mainly by the distance of the colonizers source, so the more isolated is the island, the smaller is the immigration rate. Extinction is determined by the size of the island, so that smaller islands have smaller populations, more prone to local extinction. A central point of the theory, the insular equilibrium model, establishes that the number of species in an island is the result of a continuous disappearance of resident species (extinction) and replacement by new colonizing species (immigration).

A quite impressive amount of research was triggered by this theory (recent work includes [Fox and Srivastava, 2006; Kalmar and Currie, 2006; Heaney, 2007; Kadmon and Allouche, 2007]) but no qualitative models of the theory are published so far. This work presents a qualitative reasoning (QR) model [Weld and de Kleer, 1990] developed to support understanding of the most important aspects of the Island Biogeography theory. QR may be a useful approach because causal relations underlying the system are not explicitly demonstrated in mathematical models. Also the use of everyday language facilitates communication, making it easier for non-experts to understand the ecological concepts addressed by the theory. Given that this classic theory has been widely applied to management problems, it is important to have a qualitative model that could be used by students and decision makers, particularly those related to planning and management of protect areas, to learn, understand and make predictions about species diversity in different contexts. The model should be able to answer the following questions: (a) *what are the consequences of distance to the source and island area variation for species diversity?* (b) *How relevant processes influence species diversity?*

The model was implemented in the qualitative reasoning workbench Garp3 [Bredeweg *et al.*, 2006], using the ontology provided by the Qualitative Process theory [Forbus, 1984], particularly with respect to the use of direct influences and qualitative proportionalities to represent causal and mathematical relations. The system structure is modelled around the following entities: *Island*, *Community* and *Population*. These entities capture the notion that immigration and extinction processes operate at the level of the population, and species diversity is a feature of biological communities. Relevant properties of the entities are associated to quantities, and each quantity can assume qualitative values that represent possible states of the entity. The spatial reference for the phenomena described in the model is provided by *Island*, which is associated to the quantities *area* and *area growth rate*. *Community* is associated to the quantity *species diversity*, and holds a structural relation ('contains') with *Population*, which in turn 'occupies the *Island*. The quantities *immigrants*, *emigrants*, *migration rate*, *speciation-extinction rate*, and *establishment-mortality rate* are associated to the entity *Population*.

The model library consists of 21 model fragments that encode descriptive aspects of the three entities and of the structural relations among them. Four processes drive the system dynamics in the model: area growth, migration, speciation-extinction, and establishment-mortality. In fact, they represent seven processes: *area growth* is related to changes in island area size; migration is a combination of *immigration* and *emigration*; *speciation* is the creation of new species, and this process is counter-balanced by *extinction*, the disappearance of existing species; finally, establishment-mortality is a combination of *establishment*, the survival of very young individuals in a particular environment, and *mortality*, their disappearance. All the four rates have quantity spaces {*minus*, *zero*, *plus*} so that the value *plus* is associated to positive effects of the processes (area increasing, immigration, speciation, establishment); *minus*, to the negative effects (area decreasing, emigration, extinction, mortality); and *zero* represents equilibrium between these effects.

The current version of the model has 20 scenarios for simulations that explore the effects of single and combined processes, as migration (to close and distant islands), migration combined to static and changing island areas, extinction and establishment of new species, until complex simulations involving all the entities and processes. The more complex simulation starts with the scenario 'All the processes and area increasing'. Initial values of the quantities are as follows: *area growth rate* is *plus* and the rates of migration, speciation-extinction, and establishment-mortality are *zero*. The quantities *area, immigrant* and *emigrant* have value *medium*, and *species diversity* is *small*. This simulation produces one initial state and 109 states in total, from which 26 are end states that represent steady situations. The causal model and value history diagrams of relevant quantities are shown in Figure 1.

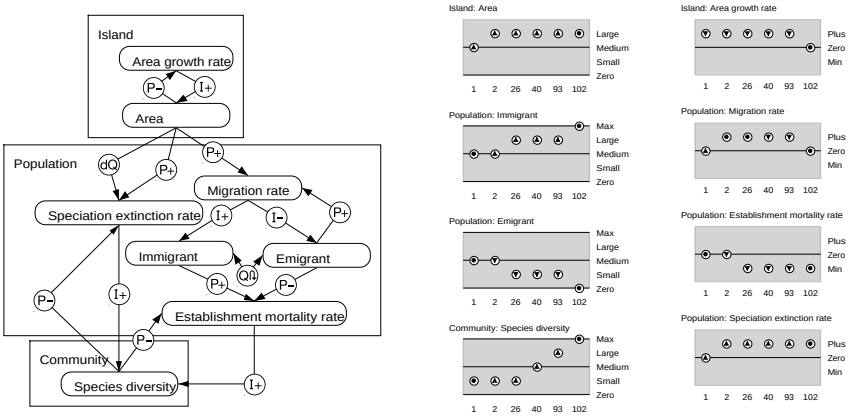


Figure 1. Causal model in state 1 and value history diagrams of relevant quantities.

In this simulation, causality flow starts with the area growth process. A feedback mechanism ensures that the area stabilizes at a certain point. While the area increases, for example, due to habitat regeneration, *migration rate* also increases. Once it assumes the value *plus, immigrant* increases and *emigrant* decreases. The inverse correspondence between these two quantities implements an assumption that keeps a close relation between them and reduces the number of states in the simulation. Changes in these two quantities propagate via qualitative proportionalities to *establishment-mortality rate.* This rate is also influenced by the derivative of *species diversity*. Without further information, this situation is ambiguous: the *rate* may increase (establishment prevails), decrease (mortality prevails) or remain stable, depending on the strength of the two influences. *Speciation-extinction rate* is influenced by changes in *area* and in *species diversity*, but it is assumed that the rate follows the derivative of *area* (see the correspondence between the two derivatives, *d_Q*), leading *speciation-extinction rate* to assume the value *plus*. The magnitude of *species diversity* is eventually determined by four processes: speciation, extinction, establishment and mortality, aggregated into two rates. The final outcome is not ambiguous, because although *speciation-extinction rate* and *establishment-mortality rate* are equal in state 1,

from state 2 onwards speciation prevails and its rate becomes greater than *establishment-mortality rate*. As a consequence, *species diversity* increases, as shown in Figure 1. The behaviour path [1 →2 →26 →40 →93 →102] shows that *area* increases, *immigrant* increases and *emigrant* decreases; *migration rate* starts in zero, increases to *plus* and stabilizes and decreases until it reaches *zero* again. It is interesting to note all the quantities are stable in state 102: *area* is *large*, *immigrant* is *maximum* and *emigrant* is *zero*; *area growth rate* and *migration rate* are *zero*; *speciation rate* is *plus* and *mortality rate* is *minus*. *Species diversity* reaches the value *maximum*, and all the derivatives equal *zero*. This way, state 102 can be seen as an extreme case of insular equilibrium.

The model was able to answer the questions formulated above and produced interesting results. Explicit causal relations may be used to support predictions and explanations about the relations between distance to the source and island area to species diversity, and the model outputs are in accordance to the Island Biogeography theory. This way, distant islands are not colonized, and the bigger is the island, the smaller is the extinction rate and the bigger is species diversity.

However, during the implementation of the ideas developed in the theory some difficulties were found. It became clear that the mathematical model is far too aggregate as a knowledge representation of the ecological processes underlying the Island Biogeography theory. In fact, immigration and extinction alone cannot provide causal account for species diversity change. The solution implemented in the model was to include emigration and speciation and, for the sake of completeness, establishment and mortality (of young organisms), as species diversity does not increase only due to new species arrival, but because of they become well established. The current version is currently being extended in many aspects. For example, to create a better representation of the concept of 'isolated', by including the possibility of intermediate refuges between the source and the island, and to include the concept of 'habitat' in the model. In fact, this would scale up the theoretical coverage of the model, as recent studies are merging concepts from the original Island Biogeography theory and the niche theory to account for species diversity [Kadmon and Allouche**,** 2007]. Concluding, qualitative representations of ecological theories prove to be efficient to define the current understanding of causal relations and to make explicit hidden assumptions about the objects included in the model and to be clear about ecological processes involved in the dynamics of the system of interest. Using an everyday vocabulary to express these ideas has the potential to make ecological theories accessible to students and non-experts.

REFERENCES

Bredeweg, B.; Bouwer, A.; Jellema, J.; Bertels, D.; Linnebank, F. and Liem, J. Garp3 - A new Workbench for Qualitative Reasoning and Modelling. *Proceedings of the 20th International Workshop on Qualitative Reasoning (QR-06)*, p. 21-28, 2006.

Forbus, K.D. Qualitative process theory. *Artificial Intelligence* 24:85-168, 1984.

Fox, J.W. and Srivastava, D. Predicting local-regional richness relationships using island biogeography models. *Oikos* 113:2 376, 2006.

Heaney, L.R. Is a new paradigm emerging for oceanic island biogeography? *Journal of Biogeography* 34(5):753-57, 2007.

Kadmon, R. and Allouche, O. Integrating the Effects of Area, Isolation, and Habitat Heterogeneity on Species Diversity: A Unification of Island Biogeography and Niche Theory. *The American Naturalist* 170:443-454, 2007.

Kalmar, A. and Currie, D.J. A global model of island biogeography. *Global Ecology and Biogeography* 15(1):72–81, 2006.

MacArthur, R.H. and Wilson E.O. An equilibrium theory of insular biogeography. *Evolution* 17:273–387, 1963.

MacArthur, R.H. and Wilson E.O. *The theory of island biogeography*. Princeton University Press. Princeton, 1967.

Weld, D. & de Kleer, J. (eds.) *Readings in Qualitative Reasoning about Physical Systems*. Morgan Kaufmann, San Mateo, CA, 1990.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Simulating an uncertain world: using qualitative reasoning to model a plant-resource system

Tim Nuttle[a], Bert Bredeweg[b], Paulo Salles[c], and Michael Neumann[d]

[a] *Biology Department, Indiana University of Pennsylvania, 114 Weyandt Hall, Indiana, PA 15212 USA (nuttle@iup.edu)*
[b] *Human Computer Studies, University of Amsterdam, The Netherlands*
[e] *Biological Sciences Institute, University of Brasilia, Brazil*
[d] *Department of Biometry, Justus-Liebig-University of Giessen, Germany*

Keywords: Artificial intelligence; resource competition; qualitative reasoning and modeling, simulation

Uncertainty is part of everyday life both for scientists and for decision makers. Nevertheless, such professionals build up theories, develop technological applications, and (try to) make informed decisions. According to Pielke (2001), uncertainty means that more than one outcome is consistent with our expectations. As such, uncertainty is not just some feature of the natural world waiting to be revealed but is instead a fundamental characteristic of how human perceptions and understandings shape expectations.

Much ecological knowledge is incomplete, qualitative and fuzzy, often expressed verbally and diagrammatically. More specifically, these characteristics are often related to the following aspects of uncertainty:

- *structural uncertainty* related to the system structure, which includes knowledge about objects, relations between entities, and what quantities should be included in the model;
- *uncertainty in explanation*, which refers to lack of knowledge to explain why a certain result or behavior appears;
- *vagueness* of the value of a quantity, a reference to the lack of knowledge about possible values of relevant quantities;
- *unknown functional relations* involving quantities included in the model;
- lack of knowledge about *exogenous processes*—those that can affect the system, but are not affected by it;
- uncertainty about the *range of possible behaviors* of the system given a particular set of conditions.

Such kinds of uncertainty associated with incomplete knowledge of a phenomenon—and incomplete knowledge of the limits of one's knowledge—are referred to as epistemic uncertainty (Pielke 2003). In contrast, aleatory uncertainty describes the kind of uncertainty that results from natural variation or stochasticity. Here, we focus on describing ways to deal with epistemic uncertainty in modeling ecological systems.

The characteristics of ecological data just described can make analytical or numerical solutions difficult or impossible to achieve, but nevertheless ecologists need ways to make their substantial qualitative knowledge explicit, well organized, computer processable (Rykiel 1989). In many cases ecologists, managers, and decision makers are interested primarily (or only) in qualitative aspects of ecological systems. For example, questions related to whether a set of conditions will improve a situation or make it worse, what happens if a system component is missing, whether certain outcomes are possible, or what

combination of factors caused a given outcome to occur can all be addressed (to a certain degree) with qualitative answers and qualitative data.

In this paper we discuss the use of Qualitative Reasoning (QR) (Weld and de Kleer 1990, Bredeweg and Struss 2003) to cope with epistemic uncertainty in ecological modeling. To do this, we describe a model of a simple yet fundamental ecological process, uptake of resources by plants. The model is based on resource competition theory (Tilman 1982). Our goals are thus to build a qualitative model that (1) captures the mechanisms of plant growth based on extraction of resources from the environment and that (2) illustrates how epistemic uncertainty can be addressed in QR models. Furthermore, by describing this fundamental interaction in the QR ontology, we construct a building block for more complex QR models of related systems.

QR provides a useful framework for expressing uncertainty due to inexact knowledge about parameter values and for exploring the consequences of different understandings of system structure. Uncertainty in parameter values can be expressed and managed using different representations of and constraints on the parameter's quantity space. Compositional modeling supports the creation of alternative models representing different system structures. Finally, QR models allow the creation of a full envisionment of all possible outcomes given a set of causal processes, a particular system structure, and starting values.

ACKNOWLEDGEMENTS

This research was funded by the European Commission's Sixth Framework Programme for Research and Development (project number 004074, project acronym Naturnet-Redime). Information on the topics being studied in this project can be found at http://www.naturnet.org.

REFERENCES

Bredeweg, B., and P. Struss. 2003. Current topics in qualitative reasoning. AI Magazine **24**:13-16.

Pielke, R. A., Jr. 2001. Room for doubt. Nature **410**:151.

Pielke, R. A., Jr. 2003. The role of models in prediction for decision. Pages 111-135 *in* C. D. Canham, J. J. Cole, and W. K. Lauenroth, editors. Models in Ecosystem Science. Princeton University Press, Princeton, New Jersey, USA.

Tilman, D. 1982. Resource competition and community structure. Princeton University Press, Princeton, New Jersey.

Weld, D., and J. de Kleer. 1990. Readings in Qualitative Reasoning about Physcial Systems. Morgan Kaufman, San Mateo, California, USA.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A Simple Tool for Preference Rating or Multi-criteria Assessment

Peter K A Barz (MIEMA, CEnv)
IEMA -Institute for Environmental Management & Assessment
info@iema.net

***Keywords**:* Qualitative decision-making; Multi-parameter iterative assessment; Relative State of Sustainability.

1. INTRODUCTION

Many assessment methodologies use a range of assessment parameters and also rely on qualitative decision-making - which may be supported by quantitative findings, as the case may be. In order to assist the organisation, integration and presentation of such qualitative decisions, the 'MyMatrix' software tool has been developed by three partners of the NATURNET-REDIME team [2007] as a simple template for use in diverse practical applications.

The following presents the basic 'MyMatrix' software tool as well as a further development for a specific application.



Figure 1. MyMatrix.

1.1 The MyMatrix Tools

As Figure 1 shows, the 'MyMatrix' software prototype is provided in the form of a simple on-line table that allows the user to enter a) the title or subject of assessment; b) any number of different parameters that influence that subject (the present example only allows for a maximum of 12 parameters); and c) the user's decision or 'rating' of the expected 'impact' of each of these parameters. The latter is achieved by clicking the respective cell

in the column of the impact scale that ranges from 'Zero' over 'Low' and 'Medium' to 'High'. On completion of this process and by clicking on 'Make myChart', a standard polar graph (for examples see Figure 3) is produced that is arranged in concentric bands representing the given impact scale - whereby the 'High' impact band (red) forms the centre of the circle and the 'Zero' impact is the band (green) with the largest diameter forming the circumference of the circle. The parameters are identified as headings around the circumference of the polar graph and the individual impact ratings are located opposite to these within the circle on the edge of their respective bands.

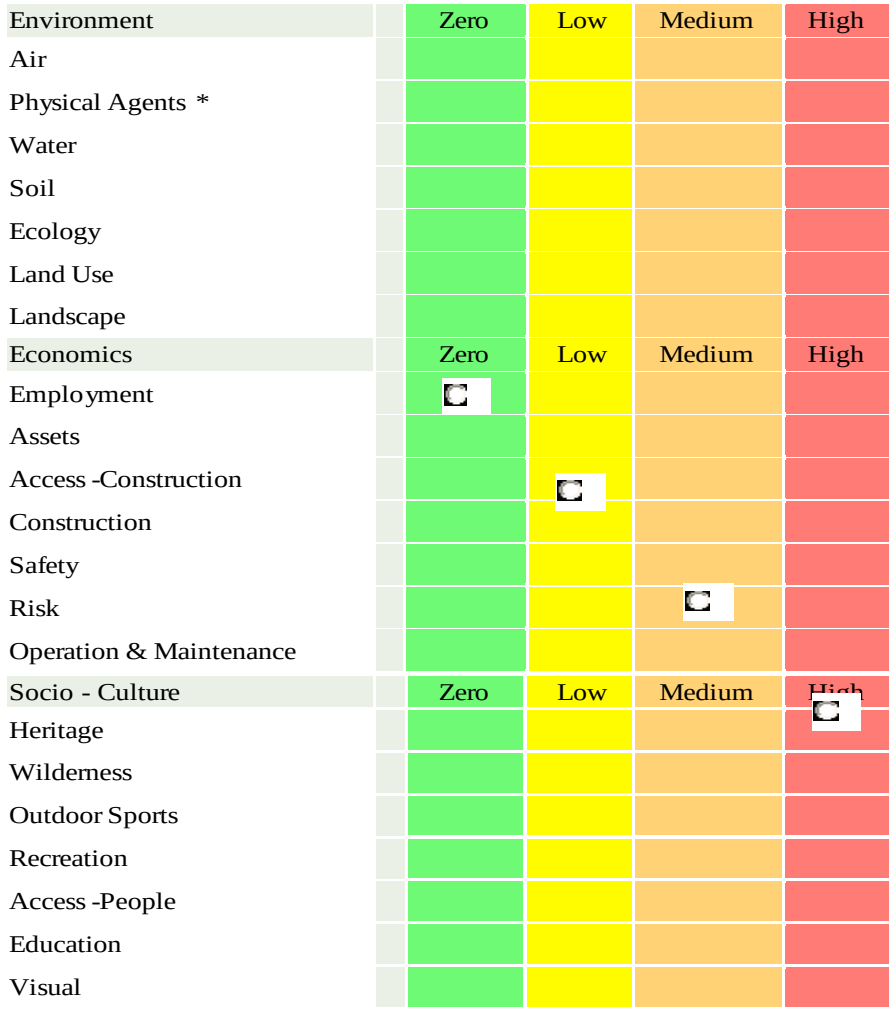


Figure 2. The SIA Impact Matrix.

The individual impact ratings are joined by a line thus forming an enclosed area which now represents the combined impact of all parameters as a percentage of 'zero impact' which represents the full area of the circle. In this manner the relative impact of the parameters can be determined. This process is iterative and may be repeated – e.g. with different impact ratings many times.

1.2 The SIA Impact Matrix

The 'MyMatrix' has been further developed as a tool for Sustainability Impact Assessment (SIA) - called SIA Impact Matrix (see Figure 2). In this case the matrix shows the parameters that have been selected as an example of the assessment of a proposal for the construction of a wind farm.

Based on the model of the 3 Pillars of Sustainable Development as promoted by the European Council [2006], the parameters are derived here from the distinct categories of

assessment, i.e. environmental impact assessment, as well as existing and user-defined socio-cultural and economic/construction parameters characteristic of such a project proposal. Furthermore, the SIA Impact Matrix has extended functionality, i.e. on completion of the parameter impact ratings for each category the user may view the resulting assessments for each of the 3 Pillars as distinct polar graphs. A further graph shows the combination of the ratings of the parameters of the three categories as an integrated picture – whereby the area covered depicts the overall 'Relative State of Sustainability' of the construction proposal. This process is iterative and may be repeated as many times as required to obtain a particular outcome – e.g. to obtain a specific (minimum) overall relative rating (percentage) – or to view and change specific ratings in an overall context.

Both the Assessment Matrix and SIA Matrix may be found on http://www.env-net.com/.

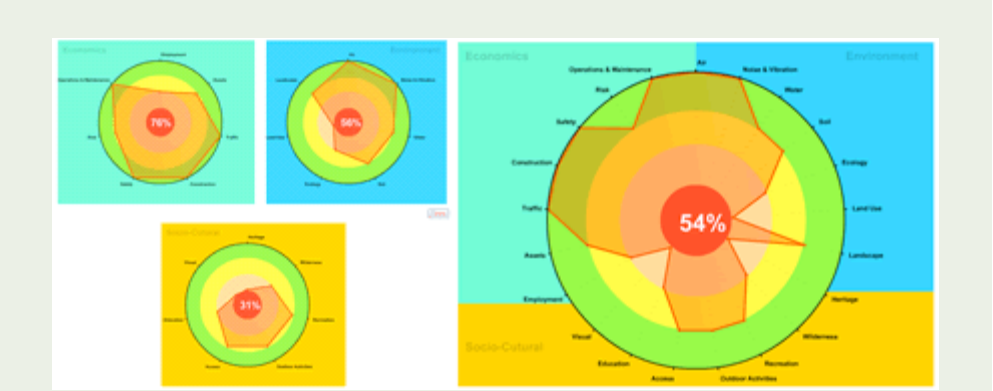


Figure 3. Example of Presentation of Multiple Impacts via Polar Graphs.

REFERENCES

European Council, Sustainable Development Strategy (EU SDS); *European Union*, Brussels 2006.

Naturnet-Redime, New Education and Decision Support Model for Active Behaviour in Sustainable Development Based on Innovative Web Services and Qualitative Reasoning; Priority Sustdev-2004-3.viii.2.e; *European Commission, Sixth Framework Programme*; IPR owners: Environmental Network Limited (ENL), United Kingdom; Department of Remote Sensing and Landscape Information Systems (FeLIS) - Albert-Ludwigs-Universitaet Freiburg, Germany; Ceske Centre for Science & Society, Czech Republic; CZ); Prague 2007.

ACKNOWLEDGMENTS

Polar Graph; Open Software; http://www.maani.us/xml_charts/

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Integrated Solution Support System for Water Management

A. Kassahun[a], M. Blind[b], A. Krause[c], O. Roosenschoon[c]

[a] *Information Technology Group, Department of Social Sciences, Wageningen University, The Netherlands (ayalew.kassahun@wur.nl)*

[b] *Rijkswaterstaat Centre for Water Management, The Netherlands*

[c] *Centre of Geo-Information, Environmental Sciences Group, Wageningen University, The Netherlands*

Abstract: Solving water management problems involves technical, social, economic, political and legal challenges and thus requires an integrated approach involving people from different backgrounds and roles. The integrated approach has been given a prominent role within the European Union's Water Framework Directive (WFD). The WFD requires an integrated approach in water management to achieve good ecological status of all water bodies. It consists amongst others of the following main planning stages: describing objectives, assessing present state, identifying gaps between objectives and present state, developing management plan, implementing measures and evaluating their impacts. The directive prescribes broad participation and consultation to achieve its objectives. Besides the obvious desktop software, such an integrated approach can benefit from using a variety of support tools. In addition to tools for specific tasks such as numerical models and questionnaires, knowledge bases on options and process support tools may be utilized. Water stress, defined as the lack of water of appropriate quality is one issue related to, but not specifically addressed by the WFD. However, like in the WFD, a participatory approach could be used to mitigate water stress. Similarly various tools can or need to be used in such a complex process. In the AquaStress Integrated project the Integrated Solution Support System (I^3S – I-triple-S) is developed. One of the cornerstones of the approach taken in AquaStress is that organizing available knowledge provides sufficient information to improve the possibility to make a water stress mitigation process truly end-user driven, meaning that dedicated local information is only collected after specific need is expressed by the stakeholders in the process. The novelty of the I^3S lies in the combination of such knowledge stored in knowledge-bases, with adaptable workflow management facilities and with specific task-oriented tools – all originating from different sources. This paper describes the I^3S.

Keywords: Integrated approach, Water Framework Directive, Water management, Participatory process, Water Stress

1 INTRODUCTION

Today's water and environmental management frequently requires a participatory approach, since the solutions affect many stakeholders and a variety of policy fields are involved. Europe's Water Framework Directive (European Commission, 2000) provides a good example of this complexity: The directive prescribes broad participation and consultation to achieve good ecological status of all water bodies in the European Union by 2015. From 2009 onwards Integrated River Basin Management plans need to be implemented. The types of measures in such plans are of very diverse nature, for example: upgrading waste water treatment plants, improving the morphological structure of rivers, regulating fertilizer use and many more economic (support) measures, such as taxation and subsidies, and educational measures may also contribute to reach the directive's objectives. Authorities on

different levels, e.g. local, regional and national authorities, and of different domains, e.g. ministries of infrastructure, water management, environment, spatial planning, and economy need to collaborate since the measures affect different policy fields and spatial scales.

Finding an appropriate set of measures to reach the objectives very much depends on the successfulness of the process leading to this selection. Besides the need to involve aforementioned authorities, this also implies involvement of various stakeholder groups and possibly the public. In the WFD, active stakeholder participation is prescribed. In such a process achieving common understanding, trust and confidence are notably challenging. Experts have a particular role in such a process: They need to deliver most of the knowledge and information that is used. If the experts or their information are not trusted this may cause a participatory process to fail [Pahl-Wostl and Hare, 2004].

Though the WFD also deals with water scarcity, water scarcity does not have a prominent role in the WFD implementation. Water stress, defined as the lack of water of appropriate quality is however of growing importance. Since water stress affects many different groups and people, a stakeholder-driven approach may be important. In the AquaStress project the participatory development of water stress mitigation plans and options is the main subject of research.

The aim of the work block four (WB 4) of the AquaStress Project is to develop a system which supports the entire participatory process, by providing a suite of integrated software tools and a knowledge system that allows having both experience and scientific information at the fingertips. The system under construction is called the I^3S (I-triple-S). The tools are of quite different nature, in functionality, user groups, and their moments of use within a participatory process. This paper describes the I^3S.

2 STORY LINE

This section provides a narrative description of the process the I^3S intends to support. An extended narrative description will be used in section 4, integrating and describing actual tools of the I^3S. The phases in this story-line are based on Maurel [2003]: “

1. Starting organization – The objective of the starting phase is to create a process design and achieve clear agreements with both the client and administrator, who may be one and the same person. A preliminary plan on who needs to be involved when should be part of this phase. Other activities are: determining the type of process, a design of the process, boundary conditions and announcement of the process. This first step does not yet include external actors. It involves the core group of persons involved executing / commissioning the participatory process.
2. Actor analysis, context – The objective of this stage is to get a full overview of actors (stakeholders) and fine tune idea’s on whom to involve when. This may be done in a participatory setting.
3. Diagnostic of the current situation – The objective of this step is to achieve a broadly agreed upon assessment of the current situation and problem identification. Activities are preliminary investigation; collect knowledge and insights, analysis and ordering of information and knowledge and informing the decisions makers.
4. Search of solutions – The objectives of this step are to search for useful solutions and realistic alternatives. It should also result in design and synthesis of these potential solutions, and transparent choices. The activities include ‘ordering and analysis of potential solutions’, determining effectiveness, involvement/participation and informing the decision makers.
5. Implementation, evaluation – The key objective is to create an implementation plan and implement it subsequently. Both the previous process and the effects of implementation should be evaluated.”

Imagine a certain region water stress occurs. Both the availability of water and its quality are of concern. Water is used by a variety of users, such as farmers, households, tourists and

nature, the latter meaning that a certain amount and quality of water is required to maintain a good ecological status. There are many possible measures to resolve the problems, e.g. improving irrigation, developing reservoirs, desalinisation, reducing water loss in distribution systems, installing water saving tabs, etc. There are also many ways to support the implementation of such measures, such as pricing and subsidies, regulations, education etc. Obviously, many people may be affected by the choice of measures and mechanisms, and financial resources, political and societal support is required to mitigate the water stress situation.

2.1 Starting Organisation Phase

To tackle the problem, the water authority appoints a project manager or project team. The task is to develop a broadly supported water stress mitigation plan.

The project team is convinced that this can only be achieved if relevant stakeholders are actively involved and can influence the decision. They are also aware that the different stakeholders have different perceptions about the problem, and may not always understand the perspectives of others. There is also some distrust due to previous experiences, where they were confronted with problem analysis and mitigation plans. The stakeholders did have little say and had no insight in the validity of the analyses and solutions developed by the experts.

Hence, the project team decides to start an open process, in which stakeholders are the main driving force in the development of the water stress mitigation plan. Given the fact that there will be many people involved, and planning and recording of activities is of paramount importance to achieve trust and transparency, they intend to use a web-based, participatory workflow management tool.

The team also realizes that the information required throughout the process should be demand (stakeholder) driven. To avoid that the process slows down each time information is required, e.g. during a meeting, the team decides that information on other sites can be used as an approximation at such a meeting. After such a meeting local information can be collected and presented if so desired. Due to the effort carried out in AquaStress, information on similar sites that dealt with water stress and potential mitigation options can be found on in a knowledge base accessible though the Internet.

2.2 Actors Analysis, Context

It is now important that the team identifies the stakeholders and actors – the societal context of the water stress situation. Different methods are available to support this activity. In this case the actor analysis is carried out via a desktop study without dedicated software support and a number of meetings with those identified. The process and results of the study needs to be stored for future use.

2.3 Diagnostic of Current Situation

It is important that the team reaches common understanding about the problem. This can be achieved via a number of meetings where the different stakeholders meet and discuss. The process can be facilitated through questionnaires, ‘water stress games’ and numerous other methods which all aim to increase the common understanding about the problems at hand and their implications for the different stakeholders. The development of appropriated indicators to describe the water stress situation within the context of the site is a major task. Agreement on a measure that best describes the situation and which can be used to evaluate the effectiveness of measures is a major output of this stage.

2.4 Search of Solutions

A number of meetings, discussions and iterations are organised in which the virtual game is replayed with different information and options. The selection of options is based on amongst others extensive information on the effectiveness of options elsewhere. After some pre-selection of promising measures in terms of effect and acceptability, it is decided to use more advance modelling tools to achieve a more accurate estimate of effects.

Based on the outcome a final session is being organised by the team. In this meeting a final deliberation takes place, resulting in prioritisation of options and finally an advice to the water authority.

2.5 Implementation, Evaluation

At this stage the competent authority implements the plan and evaluates the effects. In principle the stakeholders may be involved in this phase as well, but this is beyond the scope considered here.

3 THE I^{3}S' OBJECTIVE, ARCHITECTURE AND COMPONENTS (TOOLS)

3.1 I^{3}S' Objective

According to the mission of work block four (WB 4) of the AquaStress Project key science and knowledge outputs have to be brought together and integrated in a computer based infrastructure [Gijsbers, 2007a, b]. This system must put these outputs at the disposal of the user community finally assisting stakeholders to resolve problems arising from water stress. In that regard the support system's objective is "to enhance the selection process of water stress mitigation options by providing a suite of tools that can effectively support the participatory development of a water stress mitigation plan":

- Selection process of water stress mitigation options in this context includes all the required steps to reach agreement in a participatory setting, starting from nothing and leading to a broadly supported set of options.
- A suite of tools in this context means computer based tools (software) that can support one or more of the steps in the process. Examples are: workflow managers, knowledge bases (KBs), models, databases etc.
- Effectively means that it allows to access knowledge already available and produced within AquaStress, that the tools interact whenever appropriate, reducing tedious work on getting tools to work properly and also allows alternative and new tools to be more easy used in combination with other tools.
- Support means that it helps carrying out a task or tasks. Depending on the task, support is provided to different types of users. For example: A project manager is supported by providing a tool that helps him organize the process; A starting meeting involving different types of stakeholders may be supported by a gaming/social learning tool; A selection between alternative option-sets may be supported by a multi criteria tool.

3.2 Conceptual Overview

According to the ANSI/IEEE Std 1471-2000 [IEEE, 2000] the term architecture is defined as the fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution.

In principal the architecture of a software-system serves as fundamental description to understand or design a system consisting of several components. Hence architecture describes the collaboration or interaction amongst these components in terms of data and control-flows, and constraints of conditions, rather than the design of the individual

components. Moreover the software architecture discipline is centred on the idea of reducing complexity through abstraction and separation of concerns.

Within AquaStress a variety of tools and models are used to collect data and knowledge and to conduct research. The I^{3}S aims to bring this scientific output together in a complex support system that comprises of web-based and stand-alone applications. With that in mind, the architecture is implemented as a combination of:

- A 3-tier web-based Client-Server architecture with "thin-clients" which means most part of the application logic is implemented on the server side; and
- A 2-tier Client-Server architecture with "fat-clients" which means most part of the application logic is implemented on the client side.

Figure 1 illustrates the overall architectural design and the components of the I^{3}S:

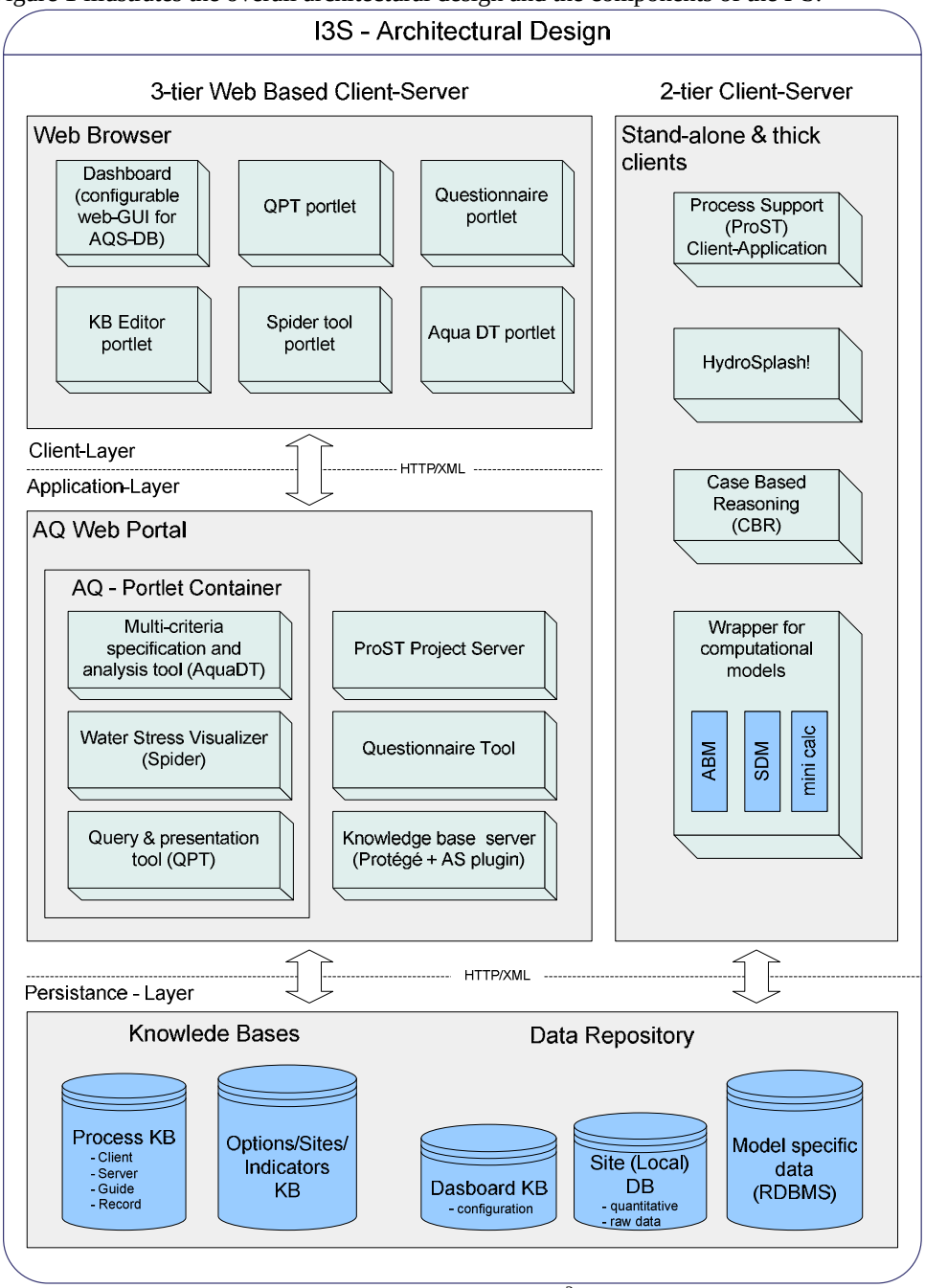


Figure 1: Conceptual view on the I^{3}S Architecture

Many of the tools are interrelated with each other or share a common data. To provide an integrated solution it is necessary to determine the kinds of dependency that may exist amongst the tools. One main characteristics of the I^{3}S design is that it is implemented as a (meta-)data-centred architecture, which means that the AquaStress knowledge bases play a central role in linking and integrating the tools. Additionally integration is achieved by using the Process Support Tool (ProST). Besides providing guidance to project teams during all the relevant steps, ProST also provides team members with the necessary information needed to execute the tasks of the participatory process including data and settings for the tools thereby reducing the tedious work of setting up tools.

The I^{3}S accommodates various levels of use:

- Browsing static fact sheets and downloading the relevant component for stand-alone applications;
- Browsing dynamic fact sheets with HTTP-based data retrieval from AquaStress knowledge bases using prescribed XML-formats, and
- Use of standalone tools that communicate with the knowledge bases using HTTP-based communication protocol using prescribed XML-file formats.

The left hand side of Figure 1 represents the 3-tier web based client server architecture. **The client-layer** represents both dynamic and static web contents. Dynamic contents are generated on the server side and provided by the corresponding portlets. Static contents are mainly links to case specific (tailored) HTML-pages. Some of these tools are directly available through the portal. Others tools however are only presented at the website and should be downloaded separately. They are described in detail in section 3.3.

The application layer is implemented as a web-portal that provides the entry point for the various components of the system. The AquaStress web portal is a site that functions as a single point of access to information. It presents information from diverse sources in a unified way and provides a way for stakeholders and users to obtain a consistent look and feel with access control and procedures for most applications. It presents information via fact sheets or dynamic data retrievals on the test sites, water stress mitigation processes, options, indicators, tools etc. Dynamic information retrievals are typically handled via HTTP Post/Get messages. Some of the information may be static HTML pages that contains a link to standalone application (or its contact point). In terms of future scalability purposes it will directly connect to specific web-services.

The persistence layer contains all persistently stored information, including knowledge items, metadata and raw data.

> ***The knowledge bases are*** at the heart of the system. They contain the knowledge which components in the system can utilize. The I^{3}S knowledge bases hold primarily textual (meta-) information on processes, sites experiencing water stress, mitigation options, indicators, etc. This information can be used by tool components to present the information to the user, to populate applications e.g. for gaming and modelling studies or to retrieve information about the location of raw data. It provides a shared definition of concepts and knowledge items within the project. To enable interaction with knowledge bases, a standardized XML-based data exchange protocol, composed of an XML Schema Definition to standardize the data exchange format, and a set of calling methods have been defined. Components that exchange (numerical) data in real-time should implement OpenMI (www.openmi.org) as a data exchange protocol.
>
> ***The database*** contains raw data such as time series required for indicator calculation, but (potentially) also for model inputs and parameters. It can serve as a common data-layer. The database can hold any type of numerical values and associated uncertainties, ranging from meteorology via ecology and hydrology to socio-economic values.

The right hand side of Figure 1 represents a conventional 2-tier client server architecture. The tools are coherent stand-alone applications that are loosely coupled through the shared

information in the knowledge bases. In doing so the desktop-applications can derive meta-information from the knowledge bases including settings for retrieving data from a shared database.

3.3 I^{3}S Components (tools / functionalities)

A number of tools has been developed or used in the AquaStress project to help with identifying relevant issues and solving problems. Some of these tools are tightly coupled with the knowledge bases, databases and other tools. Others are loosely coupled and can function independently. In this section we give a short description of the tools which have actually been used within AquaStress or which have been identified as being useful in the process of solving water mitigation problems. These descriptions consists of 1) a short description of the functionality, and 2) the application objective of the tool.

Table 1. A list software tools and functionalities beneficial for a participatory process of water stress mitigation planning.

Tool / functionality	Starting organization	Actors analysis, context	Diagnostic of current situation	Search of solutions	Implementation, evaluation	Availability in I^{3}S L = Linked A – Available but not linked
e-Glossary	☑	☑	☑	☑	☑	L
Workflow / Process management support tool	☑	☑	☑	☑	☑	L
Knowledge base (KB)	☑	☑	☑	☑	☑	L
Database (DB)			☑	☑	☑	L
Data visualization and processing (including GIS)			☑	☑	☑	L/A
Cognitive map (CMap)		☑	☑	☑		
Uncertainty assessment		☑	☑	☑	☑	A
Questionnaire (Q-tool)		☑	☑	☑	☑	L
Actor analysis		☑				
Computer supported gaming		☑	☑	☑		L
Group support system		☐	☐	☐	☐	
Cost-effectiveness analysis				☐		
Multi criteria analysis (MCA)			☑	☑		L
Case-based reasoning (CBR)		☑	☑	☑		A
Medium to complex modelling, e.g.			☑	☑		
Integrated assessment			☐	☐		
Resource flow modelling			☑	☑	☑	A
Integrated, complex modelling			☑	☑		
System dynamics models (SD)			☑	☑	☑	A
Agent Based Modelling			☑	☑		A
Mini-models – knowledge rule based modelling			☑	☑	☑	L
Querying and presentation tool (QPT)			☑	☑	☑	L

Taken from Blind et. al. [2007] Table 1 shows which software functionalities were considered useful at the outset of the AquaStress project and at what stages in the

participatory process they could be of use. The stages were chosen according to Maurel [2003]. When interpreting the table, one must account for the definitions of the stages; in principle some of the tools may be used in stages other than the ones they are marked for in the table. Moreover, not all the tools indicated are not used in practice. Whether or not a functionality is used depends on the actual process and the needs of its participants.

In the following section a detailed description of the functionality is provided.

3.3.1 e-Glossary

An electronic glossary in general is an invaluable tool for complex participatory processes. It helps to gain common understanding. The I^3S contains a glossary which can be used easily by different tools. It contains information on: topic (a list of topics or term groups have to be defined), definition; domain of applicability (e.g. Groundwater, Hydrodynamics, etc.), source, authors, references and language.

3.3.2 Workflow / Process Management Support Tool (PRoST)

Complex (participatory) processes can be supported by workflow management tools. In AquaStress such a tool has been developed: ProST. It consists of scientific and technical guidance specifying managed process steps on how to carry out various tasks to achieve reliable and reproducible results. “Managed” means that tasks within a managed process are described unambiguously, are scheduled and monitored. An effective process management support system distinguishes different types of users, identifies their interests and information needs. ProST allows multiple actors to work on a project. ProST becomes more effective when knowledge bases and tools used to carry out specific tasks are linked, e.g. can export output to the ProST Tool.

The objectives of the process management and support tool within AquaStress are:

- To describe the process flow;
- To describe tasks to be done and any relevant information that can be of use in executing the tasks defined in the process. See also “Storage of knowledge – options, etc; Knowledge bases”;
- To leave audit trail of process instances;
- To provide effective reporting facility for finished process instances and process instances in progress.

3.3.3 Knowledge Bases (KBs)

Knowledge bases here stands for a structured collection of relevant pieces of knowledge, preferably to be used by persons and by machines / computers. In AquaStress knowledge bases and the associated knowledge base editor is developed for:

- Miscellaneous Tools;
- Numerical models;
- Sites where water stress occurs;
- Water stress mitigation options;
- A decomposition of what multidisciplinary teams (including stakeholders) have to do to mitigate water stress problems.

The AquaStress e-glossary is also an example of a knowledge storage. By linking the knowledge bases to a workflow system, the knowledge becomes easily available at different steps and can be directly stored for process reporting purposes.

3.3.4 Data Storage / Databases (DB)

A data-storage or data-storages ideally allow storing all different types of data (spatial data, time-series, different domains such as hydrology, economy, ecology etc) in a transparent way such that different tools can easily link to it to retrieve numerical data, but also to store numerical results. In AquaStress the generic database developed in the Harmoni-RiB (http://harmonirib.geus.info) project has been used. It allows to store uncertainty information. For AquaStress, a new data-dictionary, covering data-needs of different tools is developed.

3.3.5 Numerical Data Visualisation and Processing (including GIS)

Complex, participatory processes require visualisation of (usually large amounts) of data (monitoring data, modelling results, spatial data, etc). Human beings are generally not geared to grasp raw data. For decision support with natural resources, spatial and temporal variability is particularly important to visualize. Furthermore, different scenarios can best be compared graphically.

A wide variety of visualisation tools are generally required to present and summarize data graphically, to best suit the human perception. In the I^3S various tools have visualisation capabilities. For further data exploration additional tools are required which are currently not embedded in the I^3S.

3.3.6 Cognitive Map/Fuzzy Cognitive Mapping (CMap)

Cognitive mapping software has the functionality to store for example one's own perception of the reality in a mental map, achieving a mental models (mental models) [Lasut, 2005]. In cognitive maps, the stored concepts are decoded, analyzed and clearly structured through cause and effect relationships, and this is significantly useful both for individual aims (thorough understanding of complex issues) and in group situations. A cognitive map is constituted by nodes, which represents the concepts and are connected to each other by links (also called edges). The edges are directed to show the directions of the cause-effect relationships. There are three main application objectives:

- Elicitation of Stakeholders' knowledge/perspective. Moreover, through questioning on the map structure, the Stakeholders' are also encouraged to find new relationships and solutions and to reach a better issue understanding.
- Improving the communication between Stakeholders'. In depth reflection upon each other maps give also the possibility of finding alternative ways of understanding the problem. Cognitive Maps serves as basis when policies and management options are discussed: the decision-making process is facilitated and conflict solutions are encouraged.
- Options analysis: cognitive mapping techniques allow several options to be examined to see which are the most beneficial and whether more detailed one need to be considered.

Currently cognitive mapping tools used in AquaStress are not integrated within I3S.

3.3.7 Uncertainty Assessment Tools

Trust in complex, participatory processes and in results of modelling and simulation can amongst others benefit from uncertainty analysis. Uncertainty assessment functionality includes all types of functionality that allows uncertainty to be explicitly addressed within AquaStress, including:

- Tools and methods that help to determine sources of uncertainty qualitatively;
- Tools and methods to quantify uncertainty in
  - data,
  - models,
  - both in data and models;
- Tools that visualize uncertainties;

- Tools that help to identify sources of uncertainty within a participatory process (e.g. evaluation of the completeness of stakeholder analysis, discussion on unknown futures.

The objective of applying such tools is to determine, evaluate and communicate the quality of results (either qualitatively or quantitatively) and as such to provide a quality mark on and increase trust about the results expected from water stress mitigation options.
The I3S includes possibilities to use the data-uncertainty engine developed within the HarmoniRiB project (http://harmonirib.geus.info).

3.3.8 Questionnaire

A questionnaire or a list of questions that is used to gain information from certain persons is an important tool in complex participatory processes, where opinions and views of participants are important input to the process and evaluation means.

In the I^3S a web-based questionnaire tool is available which supports the development, web-based filling, and evaluation of questionnaires. This tool is linked to the process support tool ProST.

3.3.9 Actor Analysis

Actor analysis results in a list of all relevant people and groups which affect or are affected by a certain problem and mitigation options, the case here being water stress and water stress mitigation options. Relevant actors are: Actors that have an interest in the decision making; Actors that can hinder the decision making; Actors that can enrich the decision making; Actors that has to be involved on moral arguments. [De Bruijn, et al., 2002]

The objectives of applying actor analysis tools are:

- To identify all relevant actors, so that you know whom to involve in your process.
- To learn about different problem perceptions of actors/stakeholders and about different content aspects of the problem situation.
- To estimate how the network of actors will participate and to determine who the 'enemies' are and who your 'friends'. Whom do you need or don't need, etc.

Currently the I3S does not contain a dedicated software tool for actor analysis.

3.3.10 Computer Supported Gaming

There are many definitions of gaming, but based on a system perspective, gaming can be seen as a communication mode that contains a game-specific model, appropriate communication technologies and a multi-player or multi stakeholder interaction pattern. There are also many different game genres, for example:

- Role Playing Games (RPGs) are gaming situations in which players take on their own or other people's roles or behavioural patterns in a real or imaginary context.
- Strategy games cast a player with the ability to manipulate the environment through path finding and simulating the effects of decisions.

In participatory processes games may be used, for example, to raise individual and group consciousness, raise motivation to solve problems, develop knowledge, learning skills, learning, experiencing unknown reality etc. Each game genre has it's own strong points related to the above objectives. In the domain of AquaStress, new information is exponentially generated, the problem situations are integrated and complex and the process is interactive and participatory with stakeholders involved.

The I^3S offers the Splash game which will be linked to various I^3S information sources.

3.3.11 Group Support System

De Vreede and Muller [1997] define Group Support Systems (GSS) as "... information systems that aim to make group meetings more productive by offering electronic support for a variety of meeting activities." An example is the Group Decision Room (GDR), which is "... a meeting environment in which electronic meeting support is used to help groups address complex problems collaboratively. The GDR consists of a normal meeting room in which every work space is equipped with a computer, these enable meeting participants to work together using an electronic meeting system."

An electronic meeting system or Group Support System (GSS) helps people to generate new ideas (brainstorming), to define concepts, to organize ideas into categories, and to evaluate ideas using various criteria and voting techniques. Groups can use a GSS to perform activities such as project evaluations, strategic planning, work process analysis and design, crisis management, budgeting, and group training.

Though a GDS would be extremely valuable, such a tool is not available within the I3S.

3.3.12 Cost-Effectiveness Analysis (CEA)

Cost effectiveness analysis is a technique in which the cost and effects of an intervention and an alternative are presented in a ratio of incremental cost to incremental effect. The method permits the comparison of alternative interventions (or programmes) in which costs are measured in monetary units and effects (outputs) are measured in non-monetary units. The output can be any indicator addressing quantity or quality aspect of water stress. Since multi-criteria analysis is the preferred method in AquaStress, the I^3S does currently not include a (an integrated) cost-effectiveness analysis component.

3.3.13 Multi Criteria Analysis (MCA)

Multi-criteria analysis is a set of procedures of analysis of complex decision problems involving non-commensurable, conflicting criteria on the basis of which alternative decisions are evaluated. MCA is used amongst others to

- Comparison, interpretation of information on outputs and costs (both monetary and other) of different combination of management options;
- Evaluating and ranking of alternative measures on the basis of their costs and effectiveness as a basis to formulate a water stress mitigation plan of measures;
- In AquaStress it is also used to select indicators for water stress based on participants preferences.

.
A participatory MCA too "AquaDT", is incorporated in the I^3S.

3.3.14 Case-Based Reasoning (CBR)

Case-based reasoning or learning by analogy is a methodology

- For modelling of human cognition.
- Of artificial intelligence (AI) used for electronic encapsulation and reuse of knowledge.
- For development of intelligent computer systems.

It can for example be used to find similar sites to a site under consideration and hence help to find and transfer knowledge from one site to another. Case-based reasoning is hence an important aspect of the I^3S since it improves the possibilities to extract specific knowledge from the knowledge bases. A case-based reasoning tool is developed and loosely linked to the I^3S.

3.3.15 Medium to complex modelling

For the purpose of this position paper it would be to far to discuss all different types of modelling. Within AquaStress we distinguished the following models and classified them to be medium to complex models:

- Integrated Assessment Models: Computer simulation programs representing a coupled natural system and a socio-economic system, modelling one or more cause-effect chains including feedback loops, and explicitly designed to serve as a tool to analyse policies in order to guide and inform the policy process, mostly by means of scenario analysis.
- Resource flow models: In general a resource flow model represents the flow of matter in a system within a defined time period of a spatial unit. It is not geographically explicit.
- System Dynamics Models: System Dynamics Models (SDM) and System Thinking is a methodology for studying and managing complex feedback systems. It started from the idea of applying concepts from feedback control theory to the study of industrial systems. The system dynamics approach is typically used where no formal impact assessment (i.e. simulation models) exist, but could be developed by linking a number of feedback mechanisms.
- Agent based models: An Agent Based Model is a specific individual based computational model for computer simulation extensively related to the theme in complex systems, emergence, Monte Carlo Method, computational sociology, multi agent systems, and evolutionary programming [http://en.wikipedia.org/wiki/Agent_based_model].
- (Integrated) Simulation models: This comprises a range of domain specific (e.g. hydrology, economy, ecology) models.

The different types of models have their merits. They are usually used to increase system understanding and cause effect relationships, making them very valuable to assess the effectiveness of changes to the system (e.g. implementing mitigation options). In AquaStress different models are used in different sites. These models are only applicable in a stand-alone way. The I^3S does point towards the different models.

3.3.16 Simple Modelling - Knowledge Rule Based Modelling

Simple models (mini-models) are mathematical equations, which in AquaStress are used to represent the effects an option. The input parameters represent site-specific data from site knowledge base and the dependent variable represents the effect as cost or any other indicator value. In I^3S simple modelling is fully integrated within QPT and the knowledge bases.

3.3.17 Querying and Presentation

In AquaStress a querying and presentation tool (QPT) has been developed that acts as an interface for browsing the knowledge bases and execution of simple modelling rules. Users can use this tool to calculate and present the effect of an option to mitigate water stress on a case study at a test site in terms of water stress indicators and indices and show the effect in the i³S web portal.

3.4 System integration

System Integration refers to the practice of combining individual software components into one system. The I^3S system brings together a diverse suite of software tools for diverse user profiles. To enable I^3S to function as one system and support the entire participatory process we identify the need for three types of system integration. The first type of integration is based a web-portal which provides a single point of access to information and tools to all users. The second type of integration links the various tools through common knowledge bases so that data, metadata and knowledge can be shared. The third type of integration focuses on supporting the participatory process. By enabling integration along

time axis this last integration enables output from one tool to be used as input to another tool at a later stage of the participatory process.

The AquaStress web-portal, as an integration platform, provides access to information as dynamic fact sheets. It also provides static links to the tools. The core of the web portal is the AquaStress querying and presentation tool which together with the AquaStress knowledge bases provides the following functionalities:

- Query options: simple queries to the knowledge base enable the user browse the water stress mitigation options;
- Query indicators: simple queries of to the knowledge base enable user to browse through the AquaStress water stress indicators;
- Query the AquaStress test sites and case studies: simple queries to the knowledge base enable the user to browse through the AquaStress test sites and case studies. More complex queries enable users to calculate the effect of options to mitigate water stress on a case study at a test site in terms of water stress indicators and indices.

A common problem in integrated water management is that users waste valuable time in exporting, converting and importing data from different tools. Therefore, over the years, a number of modelling frameworks for linking models have been developed (refer, for instance, to Gijsbers et al. [2002], Rahman et al. [2003], Argent and Rizzoli [2004] and Denzer [2005]). These solutions are, however, intended for data exchange among models in real-time and not meant for sharing of meta-information amongst different users using the same or different tools. The second type of integration, therefore, focuses on making information available to users and tools using common knowledge bases. The knowledge bases are composed of ontologies and associated instances, but the tools are not required to represent the ontologies internally. Therefore, the primary source of data exchange between the tools and the knowledge bases is based on XML-files following an agreed upon data-exchange schema and HTTP-based communication. The basic exchange format is simple: all tools are required to 'understand' the XML elements *list*, *item* and *category* (*type* of the item being exchanged) and *attribute-value* pairs (see Figure 2). The interpretation of complex value types, for instance *rules* for calculating water stress indicators, are done entirely by those tools that can interpret the *rules*. Other tools are either supposed to render the information to the user or ignore it all together.

The AquaStress process support tool forms the third type of integration. Traditionally integrated support systems are developed either from scratch or by integrating existing tools through major software modifications. Workflow based applications enable reuse of existing tools through removal of data and process flow dependency [Leymann and Roller, 1997]. A process support tool not only guides team members and monitors their activity but it also enables users to launch the tools required for a given task with the necessary data and settings that were made available during previous tasks or during defining the participatory process thereby reducing the tedious work of setting up tools.

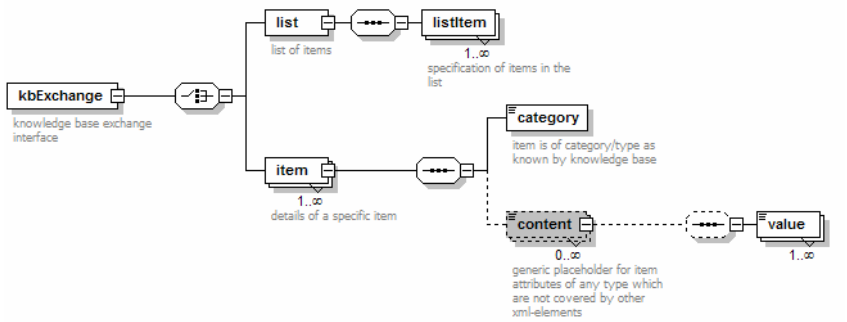


Figure 2. XML schema for data exchange.

4 SOLVING WATER STRESS PROBLEMS WITH I^{3}S

In paragraph 2, a short story line is represented in which the process is described that the I^{3}S intends to support. In paragraph 3, the overall architecture and the possible tools are brought into the spotlight. In this section, these parts will be combined into an abstract use case (Table 2) based on the needs of AquaStress test sites [Ferrand and Blind, 2007]. The intention of the I^{3}S is that it will be available in the future. Therefore, in AquaStress only a start can be made using the information available for a selected number of sites. Note that not all tools can or will be used in all situations, nor are all tools described previously included in the detailed use-case shown in Table 2.

Table 2. An abstract use-case for the I^{3}S

(Abbreviations: AP = Actors panel; KB = Knowledge base; NGO = Non-governmental organisation; PM = project manager; ProST = Process support tool; AT = Competent authority; CT= Citizens; PT = project team; QPT: Query and Presentation Tool; SG = Technical steering group; SA = System analyst)

Step	Tools used	Action	Target group
General monitoring of the water situation is organized with the pre-existing tools.		AT	AT
Non-governmental organisations (NGOs), citizens and NGOs have their own monitoring capacities based mainly on field observation and naive evaluation of the situation.		CT, NGOs	
Based on information from various monitoring systems, water stress becomes an issue in an area; which also becomes clear from citizens feedbacks. Competent authorities decide that something needs to be done.		AT	
PHASE 1: Starting Organisation			
The competent authority appoints a project team and a project manager who are responsible for defining and implementing a strategy.		AT	
The project manager is aware that different groups have different views on the severity of the problem and especially on the solutions. S/He is (vaguely) aware of political, economic and social issues. S/He finally decides to settle a technical steering group with the aim of defining a participatory process.		PM	SG
The steering group looks at the existing water stress mitigation processes available in the KBs. They discuss and proposes a first version of the participatory process. The project team collaboratively refines the process.	QPT, KB editor	SG, PT	
The project manager opens the given process in the process support tool ProST to set-up the process – enlist users, set deadlines, etc. and launches the process in ProST. At later stages the process will be modified as key actors are identified. From now on ProST will be used to guide and log all activities.	ProST	SG	PT
PHASE 2: Actor analysis / context			

Step	Tools used	Action	Target group
The project manager starts a process to identify the key actors. S/He contacts them and validates their participation.		PM	AP
To establish the current state steering group carefully designs a questionnaire.	Questionnaire -tool	SG	AP
The technical steering group processes the questionnaires to establish a first set of indicators and options.	KB editor	PM + SG	
The technical steering group establishes the basic cognitive maps based on the questionnaire collected.	CMap-tool	PM + SG	
Phase 3: Diagnostic of the current situation + Phase 4: Search of solutions			
The project manager calls the first meeting of the actors panel to actually start the participatory process.		PM	AP
At the first meeting s/he presents results including indicators, options and cognitive maps. Participants discuss the remainder of the process.	QPT, ProST	PM + SG	AP
Actors panel members can discuss the cognitive maps and establish new ones, giving their visions of the overall situation.	CMap-tool	PM + SG	AP
The project manager organizes a first multi-criteria assessment about the various issues in order to extract the priorities and actions. Conflicts and coalitions can emerge.	AquaDT, KB editor	PM + SG	AP
The project manager also presents a gaming too (Splash) to roughly explore interactions and cause effects relationships. The game will play a central role in the follow up meeting if agreed by actors panel.	Splash	PM + SG	AP
The project manager asks specialists to tune a first system model, based on the first results.	SD tool	PM + ST	SA
The combines Phase 3/Phase 4 steps may be repeated a number of times.			
The project manager calls a meeting.		PM	SG, AP
At this meeting previous discussion on indicators is re-iterated. The set of selected options is reduced and a common strategy emerges. Using the case based reasoning tool, the strategy is compared to previous cases. The meeting finally decides on the options which need further technical evaluation. The meeting decides if detailed models are required.	CBR-tool, QPT, KB editor	PM + SG	SG, AP
Additional modelling and expertise gathering are carried out. An economic analysis can be started.		SG	
Uncertainty assessment is made on the results.	Uncertainty analysis tool	SG	
Results are communicated to participants. All participants and the authority can exchange detailed information about their preferences.	QPT, AquaDT, KB editor	PM + SG	AP, SG, AT

Step	Tools used	Action	Target group
An operational water stress mitigation plan is instantiated.		PM	AT
The project manager finally reports to the authorities and finalizes the project	ProST	PM	
Phase 5: Implementation, evaluation			

5 CONCLUSIONS AND DISCUSSION

The development of the I^{3}S is a challenging activity, since it attempts to develop integrated software support for complex processes. This novel approach of integrating a diverse suite of software tools for diverse user profiles requires, among others, a high degree of knowledge on software engineering and an organisational capacity to get tool developers to participate in finding a solution to integration that has acceptable properties for 'their' tools.

Initially the process of integration started with identifying use cases and gathering information about the independent tools that are used in water stress mitigation. Next, based on the use cases and the list of tools, tool developers and technical experts determined 1) dependency among the tools in terms of data and knowledge items, 2) potential end users for each tool, 3) how and when users will need the tools, and 4) which new tools are required for integration.

Though such an integrated system is common in the industry, for instance Enterprise Resource Planning (ERP) systems, integration of such a diverse collections of tools is unknown in water management. Building new tightly integrated system is not only not feasible within the capacity and the budget of the AquaStress project but also not desirable. The challenge was, therefore, to build an integrated system using already existing tools. To achieve that (meta-)data centric architecture was chosen with three types of integration: a web-portal, common knowledge bases and process support. The web portal provides a single entry point for the various tools and presents information in a unified way. The knowledge bases enable sharing of meta information and knowledge items among the tools. The process support tool enables the tasks of the participatory process to be scheduled and monitored there by potentially allowing output from one tool to be used as input to another tool.

We started the development of the I^{3}S with the vision that in the future complex participatory processes could be supported by an information system which delivers knowledge and tools quickly for an end-user driven process in an integrated manner. Knowledge bases, a workflow manager, and several tools which can support dedicated tasks should work together in such a way that the process remains transparent and trustworthy, without the time-consuming need to develop much software support during the participatory process itself. Based on the accomplished work we conclude that developing such a system to support complex, participatory processes is feasible. Especially from the technical point of view there are no major barriers. But, there are several challenges ahead; we list the major ones below:

1) Within the project we managed to combine a set of tools. However, adapting them to work properly within the system requires work. If this work will be carried out depends most of all on the benefits the process leaders see in using the system. If the vision that process support is required for the process at hand and/or on the longer term (future participatory processes) is not shared, the willingness to adapt existing tools will be low.
2) We used open standards for connecting the different component in the system, such as XML and OpenMI. However, we needed to define data exchange formats, e.g. using dedicated XML schemes, to define content and semantics. This makes linkages of tools to the system quite dedicated, i.e. I^{3}S specific. Hence, for seamless and more generic integration we need widely agreed upon schema definitions – as OpenMI widely supports linking models in real time, there is a

need for a meta-model or ontology to exchange meta information amongst tools used in participatory water management.

3) In the project we learned that there is a bigger gap than expected between end-users (stakeholders and project managers) knowledge about potential IT support and what we intended to develop. We must recognize that mainline thinking and knowledge is about individual, sectoral tools, models, and may be Decision Support Systems, but there is less awareness on the potential of knowledge bases and workflow management support. We think that by demonstrating the system awareness will be rising on the benefits of integration and use of functionalities that may help bringing participatory water management forward.
4) One key added value of I^3S is the knowledge bases. It should be noted that I^3S' success strongly depends on easy accessible information in knowledge bases. Though these knowledge bases were easy to develop, structuring knowledge and populating the system have posed a major challenge. Partly this is due to the over-commitment of staff in other areas within the AquaStress Project; partly due to the lack of knowledge about the usefulness of populating the systems and possibly due to the fear of information being taken out-of context, fear of 'plagiary' and lack of 'credits', compared to writing a report. There are few true drivers to share information.
5) A particular challenge in developing the I^3S was the involvement of end-users. In our vision (see the use-case), much generic expert knowledge and some tools required early in the process would be available at the start: work-flow management tools, virtual games and questionnaire functionalities. In the case of AquaStress these need to be developed (and populated), while the participatory processes started immediately. Hence the I^3S was not yet available at the beginning of local site studies. We advice that in end-user driven projects serious thought must be given whether or not available information needs to be organized prior to starting the participatory process.
6) In AquaStress, the role of a process manager was not specifically specified at the outset of the project. Much responsibility was put on what we could refer to as 'self-organizing' teams in test sites. The workflow manager however requires different roles to be specifically allocated to individuals or groups of people. We expect that this will not be a main problem in real-life, since responsibilities are usually clearly defined.

In conclusion, in the AquaStress project we managed to develop a suite of linked tools in such a way that we believe they can help making participatory processes truly end-user driven. Technology-wise there remain issues to be tackled, but this will not be the main barrier for further development, acceptance and use of such a system. However, acceptance and use of the system will rely much more on the willingness and capability of people to use such integrated systems, the willingness and availability of resources to populate knowledge bases and most importantly on the willingness of water authorities to invest in transparent participatory approaches in which the authorities will lose some control on the participatory process due to shift of control towards stakeholders.

ACKNOWLEDGEMENT

The present work was carried out within the project AquaStress funded by the European Commission in FP6 under Contract 511231.

REFERENCES

Argent, R. M., An overview of model integration for environmental applications - components, frameworks and semantics, *Environmental Modelling and Software*, *19*(3), 219-234, 2004.

Blind, M., F. Dirksen and P. Sahota (editors), D4.2-4, Report: Functionality Fact Sheets, Unplanned AquaStress Deliverable, 55 pp., 2007.

Bruijn, H. de, E.ten Heuvelhof and R. in't Veld, *Process management: why project management fails in complex decision making processes*, Kluwer Academic Publishers, 180 pp., Dordrecht, 2002.

European Commission, Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, *Official Journal of the European Communities 327*, 1-73, 2000.

Ferrand, N. and M. Blind, D4.1-4, Report: Integrating needs and requirements from the sites' case studies and partners' activities in operational scenarios for the I^3S infrastructure, 22 pp, 2007.

Gijsbers, P. J. A., R. V. Moore and C. I. Tindall, HarmonIT: Towards OMI, an Open Modelling Interface and Environment to harmonise European developments in water related simulation software, Hydroinformatics 2002. Fifth International Conference on Hydroinformatics: Cardiff, UK, IAHR, 7 pp, 2002.

Gijsbers, P. J.A., D4.3-2, Report: Detailed Architecture Design – I^3S Integrated Solutions Support System, 21 pp, 2007a.

Gijsbers, P, J.A., D4.3-3, Report: Architecture for the I^3S Integrated Solutions Support System, 31 pp, 2007b.

IEEE, IEEE 1471:2000 - Recommended practice for architectural description of software intensive systems. Los Alamos, CA, 2000

Lasut, A., Creative Thinking and Modelling for the Decision Support in Water Management, FEEM Working Paper No. 81.05, 2005.

Leymann, F. and D. Roller, Workflow-based Applications. *IBM Sys Journal*, 36, 102-123, 1997

Maurel, P.(editor), Public Participation And The European Water Framework Directive-Role Of Information And Communication Tools, 86 pp, 2003. For download: http://www.harmonicop.info/_files/_down/ICTools.pdf.

Pahl-Wostl, C. and M. Hare, Processes of social learning in integrated resources management. *Journal of Community and Applied Social Psychology*, *14*(3), 1-14, 2004.

Rahman, J. M., S. P. Seaton, J.-M. Perraud , H. Hotham, D. I. Verrelli and J. R. Coleman, It's TIME for a new environmental modelling framework. In: Post, D. A., (Ed.), MODSIM 2003 International Congress on Modelling and Simulation: Townsville, Modelling and Simulation Society of Australia and New Zealand Inc., 1727-1732, 2003.

Vreede, GJ. de, Why some GSS meetings just don't work: exploring success factors of electronic meetings, Proceedings of the 7th European Conference on Information Systems (ECIS), vol. III, 1266-1285, 1997.

Supporting deliberation processes within complex decisions in natural resources management

Máñez M[1] and S Panebianco.[1]

Abstract

Decisions regarding water management issues are by nature complex; they involve balancing the three interconnected pillars of sustainability: social welfare, economic needs and the needs of an ever-changing environment. Moreover, such decisions may include controversial or poorly understood issues, such as the role of equity or property rights, as well as conflicts of interest. Due to the uncertainty about facts of matter or simply due to differences in goals, different stakeholders (people or interest groups that are directly affected by or affecting a management problem) often have quite different perspectives on what should be done. Only by taking into account these different perspectives and trying to balance the needs of society, environment and economics, can a manager reach any type of sustainable decision: one generally acceptable and beneficial to all sectors.

Sharing perspectives, i.e. finding a group consensus on weights and scores before entering them into the model is an ambitious goal but sometimes not feasible in conflicting situations. We need tools able to incorporate inputs from different people in order to compare and aggregate different perspectives. Thus, we aim to support stakeholders in making just such a complex decision.

We can start off the process by stating that the goals of such a tool include

- solutions to problems of resources use.
- solutions that are generally acceptable (for which there is a consensus amongst participants,
- solutions that are relevant worldwide and thus adaptable to other regions
- solutions that can be carried out within timelines appropriate to the length of a project
- sustainable solutions.

We suggest a methodological approach based on deliberative valuation and supported by the software tool called, *AquaDT that* facilitates the process of decision making when adaptive, participatory and multi-level management arrangements have been or are likely to be critical to success.

[1] Institute of Environmental Systems Research, University of Osnabrück, Barbarastrasse 12-G66. 49076 Osnabrück. Germany

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

System Dynamics Modelling: A Tool for Participatory Simulation of Complex Water Systems within AquaStress

L.S. Vamvakeridou-Lyroudia[a] and D.A. Savic[b]
[a]*Centre for Water Systems, University of Exeter, Harrison Building, North Park Road, Exeter EX4 4QF, UK (L.S.Vamvakeridou-Lyroudia@ex.ac.uk)*
[b]*Centre for Water Systems, University of Exeter, Harrison Building, North Park Road, Exeter EX4 4QF, UK (D.Savic@ex.ac.uk)*

Keywords: System Dynamics Modelling; Water Systems; Simulation; Participatory process; Decision making

System Dynamics Modelling (SDM) is a methodology for studying and managing complex feedback systems, typically used when formal analytical models do not exist, but system simulation can be developed by linking a number of feedback mechanisms. Forrester (1961) introduced system thinking and SDM in the early 60's as a modelling and simulation methodology for long-term decision-making in dynamic industrial management problems. Since then, SDM has been applied to various business policy and strategy problems (Barlas 2002). Consequently it has proven to be very useful for the simulation and study of complex environmental (Ford 1999, Mulligan and Wainwright 2004) and water systems (Simonovic 2003) in an integrated way.

Constructing, examining, and modifying System Dynamics Models (SDM) follows an iterative approach. Starting from conceptual qualitative models, simple quantitative models with few feedback loops and little detail are built, so as to allow the construction of an initial working numerical simulation model (Ford 1999). The working model can then be modified and improved as necessary to show the desired level of detail and complexity.

AquaStress-Mitigation of Water Stress through new Approaches to Integrating Management, Technical, Economic and Institutional Instruments (AquaStress 2005) is an EC FP6 IP project (2005-2009), comprising 35 international partners and 8 case studies, which differ considerably both in technical and spatial/societal terms. Some of the case studies involve agriculture/irrigation/water allocation issues, while others focus on urban/industrial water quantity/quality problems, the locations ranging from Northern Europe to Northern Africa. Most case studies involve re-cycling/re-use and/or re-allocation of water, whose water quality in turn varies over time and space. Additional complications arise by occasionally incomplete data and by the need of introducing interactive participatory processes at every stage, leading to application-driven investigations. Each case study is considered to be a highly complex water system, where water management options technically available (e.g. water re-use, groundwater management, desalination etc), links, interactions and consequences, all aiming at water stress mitigation, are interlinked and hard to simulate separately, i.e. without taking into account their internal connections.

Therefore in order to establish a common ground as well as a common approach from the engineering point of view, conceptual modelling and SDM have been applied. Consequently a SDM software tool has been included in the AquaStress "Integrated

Solutions Support System" (i3S), a software environment that brings together data, information, knowledge and tools being used and developed within the AquaStress project by different partners and stakeholders (AquaStress 2006a).

The purpose of this paper is to describe the SDM tool selected for the project, the participatory process for its application, as well as a characteristic prototype case study.

Mathematically all (or most) existing SDM visual environments are similar (AquaStress 2005). Here SIMILE (Muetzelfeldt & Massheder 2003, www.simulistics.com) has been selected as the primary software platform for implementing the quantitative (numerical) model for the case studies, for two reasons: (a) it efficiently supports breaking the model into sub-models, thus facilitating the development process of very complex systems, and (b) it can automatically produce model documentation (code) in C++, thus making the model potentially re-usable for further specialized applications, if necessary. Additionally the VENSIM (Simonovic 2003, www.vensim.com) graphical environment has been used for developing the qualitative conceptual/causal loop diagram of the system.

Within AquaStress a continuous interactive process, requiring mutual interdisciplinary communication and understanding at all stages is applied, between two different groups of participants. The first group consists of interdisciplinary "experts", who define, describe and suggest various "technical options" to be potentially applied for solving a problem (i.e. mitigating water stress) in specific case studies. The second group comprises local "stakeholders", who present the case study to the experts, together with initial suggestions for solving the problem, listen to further suggestions, react to them by accepting/ rejecting/modifying them and finally implement the "solution". Consequently SDM has been applied, according to the following step-by-step procedure (Vamvakeridou-Lyroudia et al 2007):

1. Initially all (project specific) technically available water management options available for improving conditions, are defined and formally conceptualized in SDM terms (i.e. as stocks, flows and converters) by the "experts". These generic conceptual diagrams are considered to be the "building blocks" for each water system within the project.
2. Problem identification at system level for each case study follows, together with the description of a dynamic hypothesis explaining the cause of the problem. This step requires interactive cooperation between "experts" and "stakeholders".
3. The conceptual (qualitative) model (diagram for each system/case study is built, by combining and linking several technical options, again through an interactive procedure (AquaStress 2006b).
4. Consequently case study specific initial quantitative models (including numerical data and parameters), are being developed, in SDM visual environment/mode by the "experts", revised and approved by the stakeholders.
5. SDM models are updated and improved interactively, i.e. by presenting and discussing intermediate models to the "stakeholders", gradually modifying and improving them through discussion and consent.
6. The models are tested and evaluated by the "stakeholders", to make certain that they represent the behaviour seen in the real world (model evaluation). If necessary further improvements are carried out (return to Step 5).
7. The final SDM models are being used for generating alternative scenarios, exploring factors, policies and impacts, aiming at supporting the decision making process.
8. "Solutions" are selected and implemented by the stakeholders, while at the end of the project, the SDM models remain with the stakeholders for further use.

The prototype for the application of SDM within AquaStress, presented in this paper, is the case study of the Kremikovtzi water system. The Kremikovtzi plant, near Sofia, Bulgaria, initially constructed in 1963, was designed to support a complete metallurgical cycle, contributing nowadays to about 2% of the Gross Domestic Product of Bulgaria and over 10% of the Bulgarian exports to the EU. In terms of water use the plant is one of the biggest

water consumers in the country, its average yearly water consumption being roughly equivalent to a city of 600000 inhabitants (Dimova et al 2007, Vamvakeridou-Lyroudia et al 2007).

The Kremikovtzi industrial water supply system is extremely complex and consists of both fresh water and reused water sources. The fact that some of the system fresh water sources are also used by urban and agricultural water users in the Sofia region, leads to regulations for priorities and upper limits to water consumption for industrial use, as well as water stress situations arising in times of drought. The purpose of the SDM model is to improve water management within the system, by (a) studying the effects of increased water re-cycling rates (b) estimating the effects of drought, by simulating "normal", "dry" and "very dry" years. The SDM application has now been completed (Steps 1-6), 23 different operational scenarios have been evaluated, while the final selection (Step 8 of the procedure) as well as the formal evaluation of the SDM tool by the stakeholders, took place in May 2008, while full documentation, is available for downloading (Vamvakeridou-Lyroudia and Savic 2008).

Within AquaStress SDM has been designed as a stand alone tool. Communication with other tools exists, as follows: The Questionnaire tool can provide an initial approach to the case study, by reflecting the stakeholders' views on critical local issues (Steps 1 and 2 in the procedure). In a similar way the Knowledge Base provides useful information about technical options available (Step 3), while numerical data for the quantitative model can be retrieved from the Data base (Step 4). However in the prototype application this has not taken place, because inevitably, the development of each i3S tool took place in parallel during the project. Consequently, the SDM prototype was for instance between Steps 4 and 5, before the questionnaire was available as a tool.

ACKNOWLEDGMENTS

The research presented in this paper was funded by the EC FP6 project AQUASTRESS, IP grant 511231. The authors would like to acknowledge the assistance provided by other AquaStress partners from the Rheinisch-Westfaelische Technische Hochschule Aachen in Germany (Katharina Tarnacki, Thomas Wintgens and Thomas Melin) and the University of Architecture, Civil Engineering and Geodesy in Sofia, Bulgaria (Galina Dimova and Irina Ribarova). They would also like to thank the authorities of Kremikovtzi Ltd, Botunetz, Sofia, Bulgaria for their cooperation and Mr. Nikolai Savov, in particular, for the data provided and his personal assistance.

REFERENCES

AquaStress, Deliverable D4.2.1., Review criteria and case study selection report, http://environ.chemeng.ntua.gr/aquastress, 2005.

AquaStress, Deliverable D4.2.4., Report: Functionality fact sheets, http://environ.chemeng.ntua.gr/aquastress, 2006a.

AquaStress, Deliverable D3.x.2., Background information and functional relationships for multisectorial assessment of options, http://environ.chemeng.ntua.gr/aquastress, 2006b.

Barlas,Y., *System dynamics: systemic feedback modelling for policy analysis in knowledge for sustainable development – an insight into the encyclopedia of life support,* Paris France, Oxford UK: UNESCO Publishing, Eolss Publishers, 2002.

Dimova, G., Tarnacki, K., Melin, T., Ribarova, I., Vamvakeridou-Lyroudia, L., Savov, N., and Wintgens, T., The water balance as a tool for improving the industrial water management in the metallurgical industry – Case study Kremikovtzi Ltd., Bulgaria, Proc. 6th IWA specialty conference on wastewater reclamation & reuse for sustainability, October 9-12, 2007.

Ford, A., *Modeling the Environment: An Introduction to System Dynamics Modeling of Environmental Systems*, Washington DC, U.S.A: Island Press, 1999.

Forrester, J., *Industrial Dynamics*, Waltham, MA, U.S.A.: Pegasus Communications, 1961.

Muetzelfeldt, R. & Massheder, J., The Simile visual modelling environment, *Europ. Journal of Agronomy*, *18*, 345-358, 2003.

Mulligan, M. and Wainwright, J., Modelling and Model Building, in Wainwright J. & Mulligan M. (eds.), *Environmental Modelling: Finding Simplicity in Complexity,* West Sussex, England, UK: John Wiley & Sons Ltd., 7-73, 2004.

Simonovic, S., Assessment of Water Resources Through System Dynamics Simulation: From Global Issues to Regional Solutions, Proc. 36th Hawaii International Conference on System Sciences IEEE (HICSS'03), Hawaii: USA, 2003.

Vamvakeridou-Lyroudia, L.S., Savic, D.A., Tarnacki K., Wintgens T., Dimova, G. and Ribarova I., Conceptual/System Dynamics Modelling Applied for the Simulation of Complex Water Systems, in *Water Management Challenges in Global Change*, Taylor & Francis Group, London UK, 159-167, Proc. Int. Conf. CCWI 2007 and SUWM 2007, Leicester UK, September 3-5, 2007.

Vamvakeridou-Lyroudia, L.S., and Savic, D.A., System Dynamics Modelling: The Kremikovtzi Water System, Report No.2008/01, Centre for Water Systems, School of Engineering, Computing and Mathematics, University of Exeter, Exeter, U.K., 132p, http://www.centres.ex.ac.uk/cws/downloads/cat_view/43-aquastress, 2008.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

An experimental protocol for extracting and exploring knowledge from citizens interviews for participatory water management

G. Abrami[a], N. Ferrand[a] and J. E. Rougier[a]

[a] *Cemagref, UMR G-eau, 361rue JF Breton, BP5095, 34196 Montpellier cedex 05, France (abrami/ferrand/rougier@cemagref.fr)*

Keywords: knowledge engineering, concept maps, citizen interviews, ontology, OWL, participatory water management

Participatory water management over large territories implies the generation of large amount of data that needs processing and integration. Computing tools such as simulation models or statistical tools can be used to explore quantitative data, but for qualitative data such as questionnaires, interviews or meeting minutes, most of the time the integration is done manually by experts. The work presented here aims at experimenting how knowledge engineering techniques could help in designing a protocol that could facilitate, speed-up and generalize the processing of interviews. A very important point is that we consider the variety of viewpoints contained in a set of interview as a prominent feature of the included knowledge. Special care is given to keep the "signature" of all pieces of knowledge so that different viewpoints can be compared or compatibility of strategic options challenged with these various viewpoints.

The experimental protocol we are designing consists in two main parts: 1. extracting knowledge from interviews and storing it in a computer-processable format; 2. designing tools that make it easy to (a) visualize and (b) question this knowledge. OWL ontologies are used to store knowledge [Polhill et al., 2006] and concept maps are used to visualize it [Castro et al., 2006]. Then logics (with OWL) or graph theory (with concept maps) can be used to question it. We use the following softwares : COE-CMap (Concept Map building with OWL import-export possibilities); Sphynx for lemmatisation; Protégé for ontology building and manipulation. The main phases of our protocol are:

- Data collection : at this stage of experiment, the idea is to use structured questionnaires that fit in a predefined ontology. More work has to be carried on text analysis techniques and software to use an unstructured corpus, following [Dray et al., 2006]. However special attention has to be carried and care has to be taken in questioning the quality of the initial knowledge when analysing the protocol's results.
- Ontologies building: different levels of ontologies are merged into our structured database so that the knowkedge can be easily questionable. The issue here is to find a way to get a semi-automatic classification of concepts.
- Knowledge extraction and visualisation: sub-ontologies can then easily be extracted by performing queries on the knowledge base. These sub-ontologies are converted into concept maps and can be visualised using the software CMap
- Knowledge processing: the knowledge stored in the database represents viewpoints of different actors on their system. Deductive logic and graph theory techniques will be challenged and explored to perform automated comparison of these viewpoints or evaluate the way they are compatible with management options.

The result of this work will consist in a set of methodological guidelines to produce an OWL knowledge base from a corpus of citizen interviews and a set of tools for its exploration for challenging stakeholders viewpoints and management strategies.

ACKNOWLEDGEMENTS

This work is part of the European Aquastress project and experimentation is done using data from the Bulgarian Iskar test case.

REFERENCES

Castro, A. G., Rocca-Serra, P., et al., The use of concept maps during knowledge elicitation in ontology development processes--the nutrigenomics use case, BMC bioinformatics [electronic resource]. 7: 267, 2006.

Dray, A., Perez, P., Jones, N., Le Page, C., D'aquino, P., White, I., Auatabu, T., The AtollGame Experience : from knowledge engineering to a computer-assisted role playing game, Journal of Artificial Societieds and Social Simulation 9(1), http://jasss.soc.surrey.ac.uk/9/1/6.html, 2006.

Polhill, G. and Ziervogel, G., Using ontologies with case studies: an end-user perspective on OWL, National Centre for e-Social Science Annual Conference 2006, Manchester, UK, 2006.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A Multi-Agent system for water management in river basins

[a]T. Rendón-Sallard, [a]M. Sànchez-Marrè, [b]M. Aulinas, [b]J.Comas

[a] *Knowledge Engineering and Machine Learning Group, Universitat Politècnica de Catalunya c/Jordi Girona 1-3, E08034, Barcelona, Spain {trendon,miquel}@lsi.upc.edu*
[b] *Laboratory of Chemical and Environmental Engineering, University of Girona, Catalonia,Spain montseaulinas@gmail.com, quim@lequia.udg.es*

***Keywords**:* Multi-Agent systems; Wastewater management; Modelling

Wastewater management in river basins is becoming increasingly complex. Whilst there is an urge to reduce the ecological imbalances in fluvial ecosystems, more wastewater has to be treated because of both demographic and industrial growth. Furthermore, given the intrinsic multidimensionality of river basins, its management must take into account all the agents that affect and are affected by the wastewater Aulinas *et al.* [2007].

Traditionally hydraulic infrastructures for sanitation have been managed separately, taking into account only the characteristics of the water at the entry and exit points of each installation. The current tendency promoted by the Water Framework Directive is to treat the hydrographic basin as a single area of operations within which hydraulic infrastructures have to be managed in an integrated manner, taking into account the condition of the receptor environment Devesa [2006].

In our work, we are implementing this integrated approach with a Multi-Agent system (MAS) at catchment scale that integrates the main water systems involved (*i.e.,* sewage system, Waste Water Treatment Plants and River) by means of agents who model real environmental situations and work collaboratively for achieving the system's goals.

The aim of our MAS is to cope with the complexity related to the decision-making processes found in the river basins management (*e.g.*, uncertainty, intrinsic instability, approximate knowledge, multidimensionality of river basins). The MAS can simulate different scenarios for evaluating the consequences of critical processes in a river catchment system in order to recommend alternative courses of actions and thus provide support in the decision-making of river basins management Rendon *et al.* [2006].

Our software tool is an extension of the agent platform Jadex. We incorporated the artificial intelligence technique *Rule-based reasoning* as the inference mechanism so we can have rule-based agents capable of representing domain knowledge. In the near future we plan to incorporate *Case-based reasoning* as well. For the design and development of a prototype for our case study we used the Prometheus methodology Padgham and Winikoff, [2004] and the agent platform Jadex Braubach *et al*. [2004] respectively.

The case of study is the *Besòs river basin*, located on the North East of the Mediterranean coast of Spain. The catchment area is one of the most populated catchments in Catalonia, having more than two million people connected. The scope of the study area is around the final reaches of the Congost River. The river sustains, in an area of 70 km^2, the discharges of four towns which are connected to two Waste Water Treatment Plants (WWTP).

The MAS features the main elements of the hydraulic infrastructure of our case of study, *the Besòs river basin*, and aims to manage the environmental system as a single area, integrating the two sanitation systems (La Garriga and Granollers) with their respective sewage systems and WWTP's, as well as the Besòs river as the receptor for their waste water. Other elements are rain control stations, river water

quality control stations, flow retention and storage tanks. There also is a sewer channel that connects both WWTPs, allowing to by-pass the flow from the La Garriga-WWTP to the Granollers-WWTP. To integrate knowledge and control the management of the Storage tanks we use decision trees defined in Devesa [2006]. The set of actuation rules represented in the trees aims to minimise the impact on the water quality of the river.

ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish project TIN2004-01368 and European project Provenance IST-511085.

REFERENCES

Aulinas, M., Tolchinsky, P., Turon, C., Poch, M., and Cortés, U., Is my spill environmentally safe? towards an integrated management of wastewater in a river basin using agents that can argue. In The 7th International IWA Symposium on Systems Analysis and Integrated Assessment in Water Management, 2007.

Braubach, Pokahr, and Lamersdorf. Jadex: A Short Overview, in: Main Conference Net.ObjectDays, AgentExpo 2004.

Devesa, F., Development of an Environmental Decision Support System for the management of hydraulic infrastructures, with the objective to guarantee the water quality at the Besós river basin. Phd Thesis for University of Girona, Spain ,2006

Padgham L. and Winikoff M.,Developing Intelligent Agent Systems: A Practical Guide. John Wiley and sons, 2004.

Rendon-Sallard, T. and Sànchez-Marrè , M., Simulating scenarios for decision-making in river basin systems through a Multi-Agent system. In the 3rd biennal meeting of the International Environmental Modelling and Software Society; iEMSS 2006.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

WARGI: a Tool for Water System Drought Mitigation within EU and National Projects

Giovanni M. Sechi [a] and Andrea Sulis [a]
[a] *Dept. of Land Engineering, University of Cagliari, 09123 Cagliari, Italy*
(asulis@unica.it)

***Keywords**:* WARGI-DSS, Complex Water Systems, Drought, SEDEMED II, PRIN 2005, PON-CyberSar.

Even if Mediterranean regions have frequently experienced drought events in last 20 years, the definition of a comprehensive drought management plan with clear objectives and tools is still far to be achieved. Currently, the common response to droughts includes expensive measures to minimize its impacts once the drought occurs instead of a planning strategy to reduce system vulnerability with a set of structural and non structural measures, and to improve the effectiveness of these measures to face drought. Likely, the achievement of a large consensus on a drought plan is the main goal for those responsible for managing the resources in those systems. In our view, the definition and implementation of severe measures in advance is a good way to reduce conflicts among users. Nevertheless, when pre-emptive measures are taken before the drought events the level and the timing of these measures are critical aspects. Computer modeling provides a way of predicting the future behaviour of water systems (Loucks, 1992), and to reach a common understanding and agreement on what should be done when drought occurs.

A methodology to reach the best combination of mitigation measures has been developed by the University of Cagliari (Italy) within the SEDEMED II Project (2003) operating in the INTERREG IIIB MEDOCC framework. SEDEMED II has involved 13 international partners from Southern Europe and Northern Africa with the aim of defining and spreading shared standards and techniques which permit to prevent and mitigate drought events on different spatial scales.

The proposed methodology is a full integration of optimization and simulation model implemented in the WARGI (WAter Resources Graphical Interface) (Sechi and Sulis, 2007) decision support system. The exploratory power of the optimization model allows to dynamically defining a set of pre-emptive measures under a large set of future hydrological and demand scenarios that the water system might experience. The optimization model assures a good balance between the effectiveness and costs of these measures and avoids the implementation of unnecessary early restriction on demands or too late and no effective measures. The simulation model analyses these set of measures and adopts additional measures during the drought event particularly in the case of a drought more severe than in the past. The proposed approach has been under test in the South Sardinia water system within the SEDEMED II Project and in the Agri-Sinni water system (South Italy) within a national research project (PRIN, 2005).

Particularly in complex systems as the South Sardinia with a large set of representative hydrological and demand scenarios, the search of the best combination of mitigation measures worthy to be simulated might be time consuming. In such cases, the proposed approach might require more advanced computing resources.

GRIDs allocate the computational power of a large number of computers to solve large-scale problems and involve sharing and virtualizing computing resources located in

different places and connected by a network. The goal is to scale the problem so that each computing resource can make a useful contribution. When WARGI is interfaced with a GRID, each deterministic optimization problem based on a particular demand and hydrological scenario is analyzed by a different computing element of the GRID. Problem submission, and retrieving information and outputs are performed by GRID services in a user friendly manner. In the search of the best combination of drought mitigation measures, preliminary applications to the South Sardinia water system show:

1. Significant time and resource saving;
2. Higher resilience in the case of failure of a single computing element;
3. Easy access to distributed information stored in different places.

ACKNOWLEDGMENTS

These researches were developed with the financial support of the European Union, under the Project SEDEMED II "Drought and desertification in Mediterranean basins, INTERREG IIIB MEDOCC", and the financial support of the Italian Ministry of Education, University and Research, under the PON-Cybersar Project (2006) "High Performance in Data Process and Simulation" and the PRIN 2005 Project "Decision Support Model in complex system management under shortage conditions".

REFERENCES

Loucks, D.P., Water Resources Systems Models: Their Role in Planning, *Journal of Water Resources Planning and Management, 118(3)*, 214-223, 1992.

PON-Cybersar Project, National Operation Programme Scientific Research, Technological Development, Higher Training – Activity 11.2: High Performance in Data Process and Simulation, 2006.

PRIN, Decision Support Model in complex system management under shortage conditions, prot. 2005082390, Italian Ministry of Education, University and Research, 2005.

Sechi, G.M. and Sulis, A., Mixed simulation-optimization technique for complex water resources systems analysis under drought conditions, Chapter 11, in: *Methods and Tools for Drought Analysis and Management*, Rossi G., Vega T. and Bonaccorso B. (eds.), Springer, Berlin, 217-237, 2007.

SEDEMED, Drought and desertification in Mediterranean basins, INTERREG IIIB MEDOCC Axe 4 - Ref. 2003-03-4.4-I-010, 2003.

iEMSs 2008: International Congress on Environmental and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental and Software Society (iEMSs), 2008

Community modelling, and data-model interoperability

Alexey Voinov[a], Ilya Zaslavskiy[b], David Arctur[c], Chris Duffy[d] and Ralf Seppelt[e]

a – Chesapeake Research Consortium Community Program, Johns Hopkins University, USA, alexey.voinov@uvm.edu
b – San Diego Supercomputer Centre, CA, USA
c – OGC Interoperability Institute, Austin, Texas, USA
d – Penn State University, PA, USA
e – Centre for Environmental Research Leipzig-Halle, Germany

Abstract: Community modelling is a promising paradigm to develop complex evolving and adaptable modelling systems that can share methods, data and models more easily within specialized communities. Why then are cooperative modelling communities still quite rare and do not propagate easily? Why has open source been so successful for software development, yet open models are still quite exotic? One difference between software and models is that software shares some common language. Models often use very different principles, theories, and semantics. For example hydrodynamic models, ecological models, and decision support models may have limited commonalities, In these cases, the disciplinary problem being solved may be the impediment to communication and to development of effective community tools these principles to another; it becomes difficult for one model to talk to another one. Similar problems prevail in data operations, when data sets (which are also models of sort) are hard to integrate with other data.

An issue of contemporary interest is how will community data and models be implemented within environmental observatories. The environmental observatory may are become the ultimate driver for advancing research with a clear need for interoperability standards and functionality.

There are at least three facets to the problem:
- Lack of common modelling and software tools to enable modularity and connectivity;
- Insufficient community understanding or access to basic tools;
- Lack of social motivation and communication skills to enable communal work and sharing environments.

The goals of this paper are to explore these areas with respect to the following points:
- Understand the interoperability needs of the community for data and models within a participatory and collaborative framework;
- Discuss research scenarios that would benefit from interoperability and explore interoperability architecture and standards supporting these scenarios;
- Explore environmental system observatory ontologies, with particular attention to mapping variables to concepts;

- Discuss common access protocols, enabling models to automatically search for data needed and link to data servers. Design data interoperability for model input/output to help link models.

Keywords: Modularity, observatory, culture of sharing, ontology, open source, environmental management

1. INTRODUCTION

There is an increasing consensus that value can be derived from integrating different models and data sets for predictive understanding of environmental problems and for operational forecasting.

Linking models and coordinating data sets has become a leading driver for research at environmental observatories. The nation's ability to monitor, understand and forecast environmental change are a part of NSF strategy for building capacity for research communities engaged in advanced continuous monitoring and modelling of the environment, e.g. CLEANER [1], CUAHSI [2], NEON [3], GEON [4, 28], LTER [5]. They are envisioned as collaborative interoperable systems for synthesizing environmental data from data repositories and real time monitoring into a shared data integration and analysis environment that supports multidimensional and multifaceted modelling of environmental processes, linkage or coupling to environmental sensors/data, and accessible to a broad community of environmental researchers. However they have not been doing much in terms of model integration.

There is also considerable interest in making models "talk" to each other. OpenMI is one such approach.

Other approaches allow data and environmental models to be loosely coupled, tightly coupled, or embedded. Bhatt et al. (this volume) describe how a hydrodynamic model for watershed hydrology greatly improves the modelling process when tight coupling between the GIS tool and the physical model is implemented. Embedded approaches will allow the user to "steer" the physical model in real time such as for development of climate scenarios for long term water resources decision making. Traditional approaches where the GIS tool is independent of the physical model application will continue to be important for simple models with limited data demands, but coupled approaches are likely to be of greater interest researchers and managers.

Development of an integrated data sharing infrastructure to facilitate multidisciplinary collaborative analysis and modelling in the context of an environmental observatory is a pressing need. With such infrastructure, researchers should be able to publish and document their data, discover what information is available based on agreed-upon metadata descriptions, retrieve the information over common data access mechanisms, understand and resolve semantic discrepancies between datasets, integrate them for use in analysis and modelling codes, and share research findings with community members. In this research cycle, information sharing and re-use are the major underpinnings in reducing fragmentation in environmental research, and engaging a broader research community from the environmental science and related domains in advanced data collection, analysis and modelling.

The need to focus on the common data foundation for the communities involved in environmental monitoring, analysis and modelling, is underscored by the following observations:

- Research groups and communities are not isolated but have "neighbours" that are interested or in need of the same data sets;
- Processes and problems are not isolated but have an inherent complexity that requires multi-disciplinary teams need to understand or solve;

- Increasing size and complexity, and hence information needs, of environmental models mean pressure to find automated ways to interpret larger and more complex sets of data across disciplines;
- Besides re-using data collected in neighbouring research domains, it becomes common to include longer time series data in environmental models, which requires that data are archived in such formats and preservation environments that support long-term storage and retrieval (i.e., "data interoperability in time", in addition to interoperability across domains);
- Interoperability focus is a way to consolidate the entire community around common data interchange issues, including data interchange standards and standards governance, heterogeneities across datasets and techniques for resolving such heterogeneities, leading to formulation of new research questions that could not be addressed over a fragmented infrastructure.

The range of interoperability challenges, derived from differences in structure and semantics of datasets, data publication, discovery and access mechanisms, as well as in modelling approaches, have been described in recent literature on environmental observing systems [60]. Technical interoperability issues, such as those related to common procedures for real time data management, integration of streaming data with data archives, and technologies for expressing and resolving well-understood structural and semantic heterogeneities, have been the focus of NSF attention over the last several years, within such initiatives as CEO:P (Cyberinfrastructure for Environmental Observatories: Prototypes), Geoinformatics, CLEANER/WATERS [1, 6] and NEON [3]. However, purely technical solutions for interoperability are insufficient for establishing a shared interoperable infrastructure for environmental observatories. For the software infrastructure to be truly useful for empowering the entire research community with data sharing and collaborative research capabilities, several additional challenges must be addressed. They include:

- making the community aware of the available data and software resources, common component software, and advances within other disciplines (e.g. high performance computing, visualization, etc.)
- building consensus about information models used in the community,
- understanding and harmonizing data structures and data access mechanisms and formats used within the community, and
- building support for modelling applications that take advantage of the infrastructure.

The initial experience of infrastructure development and adoption within GEON [4] and CUAHSI [2] HIS (Hydrologic Information System) projects provides ample evidence of heterogeneities in data and resources needed to compute watershed and estuary models in particular, and of the range of interoperability challenges stemming from different data structures and semantics adopted within the community. Similar interoperability issues are being addressed, in a cross-country setting, within the European "Water Information System for Europe" (WISE) project [27]. At the global scale, the issues of standard formats and exchange protocols for observations data, environmental observations in particular are being addressed by the Open Geospatial Consortium [38,59, 61, 62-65].

Development of ecosystem models in general has been limited by the ability of any single team of researchers to deal with the conceptual complexity of formulating, building, calibrating, and debugging complex models. The need for collaborative model building has been recognized in the environmental sciences. The current-generation models tend to be "idiosyncratic monoliths that are comprehensible only to the builders" [26]. Communicating the structure of the model to others can become an insurmountable obstacle to collaboration and acceptance of the model. The interoperability functions that we propose to develop will be the core of a system of

middleware that would allow integrating existing models and will provide for easy integration of new models.

An additional reason for the current disconnect is the mismatch between the generality and scalability expectations of modeling applications, which typically aim at providing the most accurate description of physical processes to support understanding, forecasting and/or decision-making for a fairly narrow well-defined case, vs. development of common software tools, where attaining a generic treatment applicable to a large class of use cases is a necessary ingredient.

Moreover there is a clear need to learn to communicate not just the models, but also be able to dig inside the model structure and have the functionality to extract the most successful pieces and modules to use them in other models. Otherwise we constantly run into the "reinventing the wheel" syndrome, when almost similar functions and algorithms are use to describe ecological processes simply embedded in different data and user interfaces. We could imagine much more advances in modelling if instead of rebuilding and reprogramming the same modules over again in new models, we could focus on further exploration of new ecological phenomena and modelling methods simply plugging in the already tested and approved modules from already existing models.

The models provide ample evidence for heterogeneities that need to be resolved in order to obtain more detailed and accurate model results. There is already a significant community building effort going on. For example, within the Chesapeake Bay area there is the CCMP (Chesapeake Community Modelling Project), which is building reliable working relationships among all participating modellers who are willing to share their model and data structures and needs to develop the important interoperability functions. This will include the integration of the very large existing and continuously increasing watershed, hydrodynamic, and biogeochemical (water quality) databases. This effort is led by the Chesapeake Research Consortium (CRC), which organizes universities, as well as government and non-profit research organizations around common problems. In addition, the CCMP can serve as the body responsible for governance of data exchange standards within the Bay area, providing linkages between various community strata grouped according to:

- organizations: government, industry, research universities and students, environmental and other non-profit groups (including standards groups), citizens, schoolchildren, etc.;
- domains: atmospheric, surface water, soils/geology/landscape, ground water, oceanic, socio-economics, demographics, etc.;
- paradigms: from data collectors/archivists/publishers to modellers, analysts, and decision makers.

Other examples are the NSF-funded CSDMS (Community Surface Dynamic Modelling System - http://csdms.colorado.edu), the EPA-funded CMAS (Community Modelling and Analysis System - http://www.cmascenter.org/), the DoD-funded CSTM (Community Sediment Transport Model - http://woodshole.er.usgs.gov/ project-pages/sediment-transport/) and others.

2. INTEROPERABILITY IN ENVIRONMENTAL OBSERVATORIES, AND COMMUNITY BUILDING

Our experience developing data infrastructure and modelling framework highlighted the following issues which are critical for successful model development and environmental forecasting, and – as we believe – reflect general deficiencies in data integration frameworks for environmental observatories. These include data availability, metadata and catalogues, differences in information models and data access protocols, differences in model requirements

with respect to available data, differences in semantics among datasets, and difficulties in consensus building. These are each discussed briefly in turn.

2.1. Data availability, metadata and catalogues

Creating easy-to-use, uniform and scalable data and services publication and discovery mechanisms, and helping community members familiarize themselves with the available resources, are the basis for engaging community members into an efficient interoperable network. At the moment, the many environmental stakeholders maintain their own data archives and data access systems. For the Chesapeake Bay, for example, several community resource repositories and discovery interfaces are being developed. They include:

- The Chesapeake Information Management System (CIMS), which brings together multiple datasets assembled by federal, state and local agencies. CIMS datasets have consistent metadata descriptions based on FGDC Content Standard for Digital Geospatial Metadata and the metadata content is searchable via text search, with many datasets available for download (the search interface is at [29]).

- The GEON-based CBEO portal, developed within the on-going NSF-supported CEO:P project. Within this portal, users can publish their datasets and services and search for registered resources of different types, including shapefiles, raster images, Excel spreadsheets, relational databases, WMS/WFS services, web services, documents, etc. A subset of CIMS water quality database is already registered through this portal [30]. Catalogue search is available via the portal, and via the GEONSearch web service.

- Hydrologic Information System (HIS Server) being developed within the NSF-supported CUAHSI HIS project, which supports publication, discovery and retrieval of observations data via WaterOneFlow web services [31]. The services follow CUAHSI WaterML [6] specification which has been adopted as an Open Geospatial Consortium's (OGC) Discussion Paper. The data discovery services available within the HIS Server are tuned to observations data series [32] and support data search and retrieval across data repositories via desktop and online (Data Access System for Hydrology, DASH [33]) interfaces. Under the aegis of the NSF WATERS program, the software is now installed at ten WATERS Hydrologic Observatory Testbed sites across the country, three of which are affiliated with the Chesapeake Bay program [35]. In addition, a HIS Server node is established for the CBEO project, and available within the GEON-based portal mentioned above.

Despite recent integration efforts, such as making the CIMS data, and the HIS Server for the Chesapeake Bay, available via the GEON-based portal, the above information management frameworks remain largely disconnected. Additional work is required to reconcile metadata and information discovery protocols across the repositories, to enable users query and explore available data for the area regardless of the repository and the adopted data management framework. Such disconnect is indicative for the current state of environmental observing systems, and should be addressed via the creation of a common set of re-usable cyberinfrastructure modules supporting integration of disparate data and models.

2.2 Differences in information models for observations data, and data access protocols

Our earlier attempt to import CIMS observations data into CUAHSI ODM (Observations Data Model, [36]) uncovered several differences in metadata used by both systems. In turn, both information models are different from the Open Geospatial Consortium's (OGC) "Observations and Measurements" [37] model and their Best Practices specification for Earth Observation Products [38].

In addition, each of the nationally hosted environmental data sources, such as hydrologic data repositories at USGS and EPA, have different data access interfaces. The CUAHSI WaterOneFlow services provide a simplified, consistent way of accessing data from a combination of these sources. While similar in approach to the OGC web services

specifications, the CUAHSI web services are not OGC compliant at the moment, though initial harmonization steps are outlined in the CUAHSI WaterML specification [39]. As a result users wishing to access these sources can only do so using CUAHSI-compatible client software. Other non-OGC, non-CUAHSI data servers require a still greater degree of data customization software. Other domains than hydrologic data have similar issues, and may not even have web-based access.

It should also be mentioned that current approaches do not yet address the problem of model-data coupling which is likely to be necessary in the next generation of complex physical, ecological, and chemical environmental models. Furthermore it is not yet clear what will be the data needs of the next generation of modelling tools. It is likely that at least some clarity on this issue will come from the Environmental Observatory initiatives currently underway.

2.3 Differences in model needs with respect to the available data

Models require increasingly large volumes of input data, which raises a performance problem for accessing the data via web services during model runs. It is necessary in many cases to download these large volumes of data from the national portals and transform the data in certain ways to prepare for use with the desired numerical or analytical models. Spatial data may be initially available as irregular grids, regular grids, or vector form such as points, lines, and polygons, while the model may only work with the data in a different form. The scale or resolution of the initial data may not be appropriate for a given model execution, and must be transformed to the correct scale. The temporal resolution and reference system of time-series observations can differ from that expected or needed by models. Each model may have its own specific data requirements, resulting in increasingly painstaking manual effort to locate, obtain, and transform the data in preparation for modelling applications. At a minimum a new data model is needed that reflects the relationship between modelling and current and future data streams. It is clear that the data and modelling process must be considered together if progress is to be made on either front.

Participatory Modelling [40-44] is becoming a recognized approach to modelling complex systems for decision-making. However, there are no agreed standards and platforms for data sharing and group model development available so far.

2.4 Differences in semantics between datasets available for different times and assembled in different research domains.

Another significant issue is the difference in nomenclatures and semantics among data sources for a given type of data, usually as a function of the compilation or hosting organization. The simplest differences may be in terms of language, keywords, metadata tags, etc., especially across subject domains and time periods. For example, the depth of a stream may be referred to as "gage height," "stage," or "waterlevel." A similar example is that the classification codes (identifiers) for variables such as nutrients (e.g., nitrogen) will vary, depending on the data producer's conventions. Semantic reconciliation is a well-recognized component of the data interoperability challenge [37, 45, 46, 47, 48, 49, 50, 51, 52, etc.]. Within both GEON and CUAHSI HIS, several technical approaches to ontology management and semantics-based integration have been explored. Beyond the technical issues, significant effort is required to make common nomenclatures and ontology translations accepted as the basis for community lingua franca.

2.5 Difficulties in consensus building.

The realization that interoperability is desirable and important is relatively recent. Most of the interoperability difficulties just described have arisen due to institutional and cultural conventions, rapid advance in open software, sensors and communication, all of which have evolved over long periods of time, from independent research communities, and without the priority among data producers for the need or benefits to coordinate efforts.

It is clear now that interoperability is not just a technical or technological problem, but also one of building trust and consensus within and between subject domain communities. This is not an activity that will arise on its own within any single agency, but requires a determined effort at multiple levels—from the leaders to the workers, across numerous organizations—to develop consensus-based approaches. Consortia such as the W3C (World Wide Web Consortium), OASIS (Organization for Advancement of Structured Information Systems), OGC, and many others have emerged over the last three decades to develop tools and techniques for interoperability at the level of basic information exchange mechanisms using the Internet. At the same time, consortia of scientific research centres have emerged, such as NCAR (National Centre for Atmospheric Research), EarthScope, CENS (Centre for Embedded Networked Sensing), NEON, GEON, CUAHSI, CBEO, and many others dedicated to enabling integrative, interdisciplinary research. Each of these major centres seeks to achieve some degree of harmonization and standardization of information models, semantics, ontologies, and accessing methods across multiple domains. Various governance patterns and structures are now emerging: interoperability standards that support identifier governance, vocabularies and structural definitions, and are supported by protocols, persistence implementation, and vocabulary publication and governance. However, the growing complexity of the standards themselves raises still more issues in terms of consistent interpretation, implementation, and usage.

In addition, significant institutional and cultural issues still remain that limit the efficacy of interoperability technologies and policies.

- (Institutional) Agreements may be needed between data providers and data users, concerning the scope of use, limiting liability, establishing means of cost recovery for source data compilation, addressing privacy and security concerns, and otherwise establishing trust between parties.

- (Institutional/Cultural) In many cases, established policies and procedures inhibit or even preclude interoperability. This can happen when a given community has invested substantial time and effort in developing data models, semantics, and ontologies that are in conflict with related models from other communities, leading to a sense of competition between the communities rather than cooperation and collaboration.

- (Cultural) Interest and willingness to share data is a pre-requisite to achieving interoperability, but field scientists can often be reluctant to share their data in useful ways, even when required by the funding agency. Developing trust and motivating data sharing "in the proper spirit" may require a substantial investment of resources for outreach. Resistance to sharing will be exacerbated if this issue is discounted or ignored.

The most straightforward aspects of interoperability are in the design and development of technologies and practices for cyberinfrastructures supporting federated data sources linked to users through service-oriented architectures and web services. This approach serves a wide range of government, business, and scientific application areas, and is the subject of nearly all the current standards work in progress at OASIS, OGC, and other IT standards organizations. OGC has evolved a process for testbed pilot projects that has achieved high levels of sponsored participation and collaboration by government and industry data providers and users. As a result of OGC testbed projects which use rapid prototyping as a way to accelerate the design, testing, and adoption of interoperability tools and standards, numerous national agencies in the US, UK, Europe, and Australia now routinely require OGC specifications to be applied in RFPs for enterprise and federated geospatial information portals and other systems. This is an important step toward alleviating the institutional barriers to interoperability. By engaging both the leadership and the programming staff of government agencies and commercial entities in the activities of defining and adopting geospatial standards and best practices, many of the trust issues at the institutional level are addressed, and organizational policies evolve accordingly. Now the issues of most concern within the IT standards consortia are not so much about whether to share data, but how to do so while managing data security, intellectual property

rights and licensing, personal privacy, cost recovery, and adequate functionality and performance of transactions for data creation and update.

Much work is also being done in core technology for bridging and mediating between different semantics and ontologies. The Geography Markup Language (GML) [63] was initially developed by OGC members in 2000, and has steadily improved to support the representation of essentially any type of geographic feature, but it is too abstract and general for direct use within a subject domain (user communities are encouraged to adopt an application schema based on some reasonable subset of GML to precisely characterize a given information model). Another broadly applicable standard more recently adopted has to do with webs of sensor data: SensorML [64], TransducerML [65], Sensor Observation Service [66], and other related specifications, were jointly developed by OGC and IEEE (Institute of Electrical and Electronics Engineers), and are beginning to be used in environmental observatory context.

The IT standards organizations have not, in general, been involved in the harmonization of semantics and ontologies within specific subject domains. Each distinct information community needs to develop its own information models and metadata catalogues, simply because these must reflect the deep science within each subject area. These developments, including GeoSciML (GeoSciences Markup Language, based on GML), EML (Ecological Metadata Language), CML (Chemistry Markup Language), WaterML (for hydrologic applications), and other domain-specific ML's, represent important steps in the maturing of community attitudes and understanding toward sharing information models and data in effective ways. At the same time, the choice, or development, of a relevant data model and markup language adequately describing the variety of data presentation and transformation needs of a research domain, remains a challenge. Furthermore, because these developments involve agencies, universities, and other research centres around the world, there is growing pressure to overcome many of the institutional and international barriers to interoperability, which have existed.

3. THE VISION AND THE APPROACH

Addressing the above issues requires a comprehensive approach to data and model interoperability in the course of community modelling effort. It has the following components:

- Understand the interoperability needs of the community in a participatory and collaborative effort;
- Develop a “common component” approach to data/model tools and architectures;
- Develop research scenarios that would benefit from interoperability, and build research groups around them (e.g., hypoxia). Build consensus about interoperability architecture and standards supporting these scenarios;
- Expand on environmental system observatory ontologies, in particular for mapping variables to concepts;
- Bring together modellers and data providers to agree on common access protocols, enabling models to automatically search for data needed and link to data servers to download the information required for modelling. Design data interoperability for model input/output to help link models;
- Provide an environment to support participatory modelling efforts;
- Design and implement a pilot version of software tools required for model data integration over the web;
- Leveraging existing data repositories, create an environment where researchers can discover, visualize and retrieve available information in a standard uniform manner; and
- Understand and apply more of the existing OGC standards and best practices.

4. TOWARDS AN INTEGRATED COMMUNITY MODELING INFRASTRUCTURE

There are two major complementary components in this kind of research: software research and development, and community building and integration. Both are equally important. Within the software-related category, there are a number of distinct types of components, interfaces, and web services, which can be considered.

4.1 Data availability, and harmonizing data discovery interfaces

We envision that easy query and browse access to a community information system with multiple community-contributed resources, coupled with the ease of publishing data of common interest, is a critical component of community infrastructure. The task of **unifying data discovery and exploring data availability** has the following sub-tasks:

1) Register community-generated observations datasets via existing online systems. Identify the most promising systems and focus on them.

2) Develop a uniform data discovery services system. Explore data discovery standards being adopted by the OGC. In particular, OGC has developed a set of Catalogue Services [53], common specifications for discovery, browse, and metadata query interfaces that support uniform search against heterogeneous data and service catalogues. OGC's most recent work has focused on the ebXML Registry Information Model (ebRIM) profile of the Catalogue services, as a way to accommodate various application schemas and coordinate reference systems.

3) Construct data availability research, and develop guides for data discovery and interpretation. Data discovery and integration are greatly aided by the abilities to explore and visualize large volumes of observations data and catalogues efficiently and in greater detail. Visualizing such dynamics, and interpreting it, is critical for better understanding of the entire observations history and geography for various regions.

4.2 Harmonizing information models and data access protocols

This will require development of translators, cross-walks, and other "infrastructure glue" for components supporting structural and semantic interoperability across data. This could include:

1) Exploring the differences and harmonizing proposed information models for observations data, including the OGC's "Observations and Measurements" [61] and the CUAHSI's ODM [36] and verify that the information models are complete for the observation datasets available. For example, the benefits of structural harmonization and crosswalks, can be illustrated by the Australian WRON project that assembled OGC-compliant WFS (Web Feature Service) services to expose observations data for New South Wales. Creating XSLT transformations that expose the Australian WFS services as CUAHSI WaterOneFlow services [54] allows one to quickly tune the CUAHSI HIS online data access system (DASH) to New South Wales hydrologic observations [55].

2) Developing converters and translation services for datasets and models. The CUAHSI WaterML development and its harmonization with the O&M specification is of great interest to OGC and the WRON project. The US Department of Energy, CCA program (Common Component Architecture) is a case in point.

3) Expanding the technologies developed for observations data within existing observatories, including biogeochemistry data, land use, atmospheric data, etc.. Arriving at data interchange standards and ontologies applicable across communities is a challenging task. It requires explicating and comparing semantic frameworks used in the neighbouring fields, outlining information models for commonly used data sets, standardizing data discovery and access mechanisms, and assessing common integration scenarios, e.g., in support of comprehensive modelling on watersheds and estuaries.

4.3 Matching model needs with the available data

By applying the same formalism that treats data sets as independent modules that can be accessed just like the modelling modules, we can integrate the data sets into the modelling system as well.

We need model pre-processors that automatically find model inputs from datasets registered to data catalogues, retrieve available data via web services, and assemble them in formats ready for modelling. Assembling model inputs is one of the most time consuming tasks in comprehensive environmental modelling, and requires knowledge of available datasets from several related domains (hydrologic, biological, atmospheric, social-economic, and demographic data, etc.). Making data discovery and assembly dynamic, even for a fraction of model inputs, has the potential of improving the efficiency and relevancy of environmental models, at the same time reducing proliferation of multiple versions of identical or almost identical datasets. For example consider the inputs and data flows for the CBP Phase 5 model [9,10] or similar data for PSU's PIHM model [11,12] (Figure 1).

The data structure should be designed to be flexible enough for modification and customization of the model, and rich enough to represent complex user defined spatial relations, and extensible to add more software tools as the need arises. For this we may need to explore new integration methodologies for linking a GIS framework with physical models that enables users to take advantage of object oriented programming (OOP) with direct access to the GIS data structure, which support efficient query and data transfer between the model and GIS.

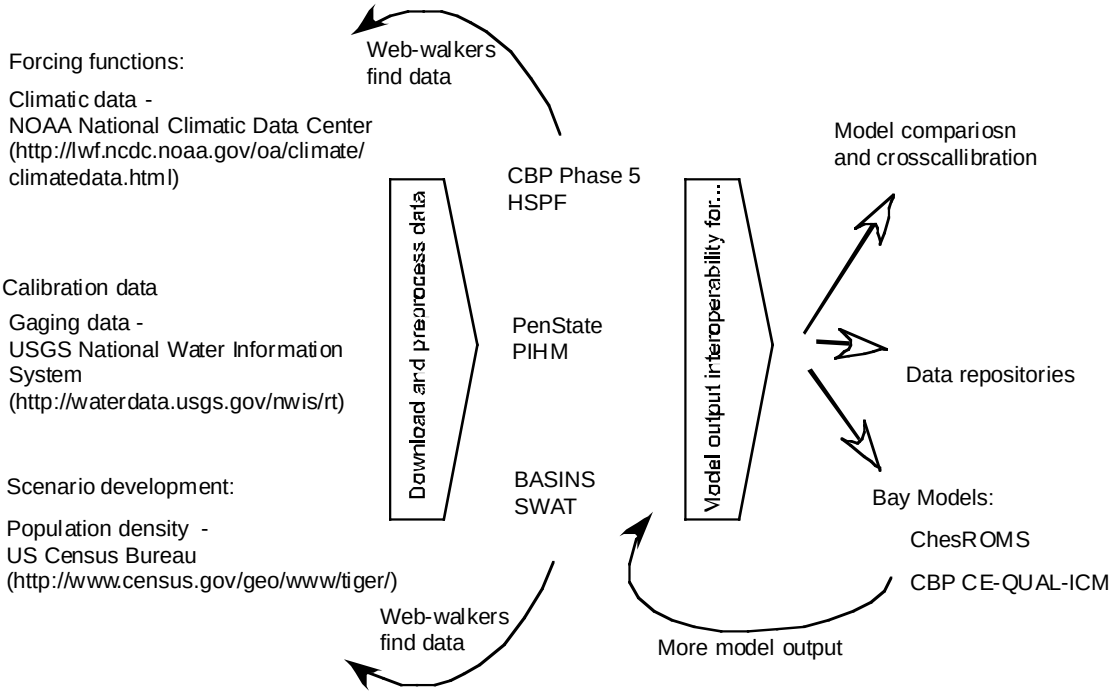


Figure 1. An example of data flow between some data sets and models. There are many more models and data sets, but these will be primarily targeted for this pilot study [11-25].

Suppose we have a watershed model such as the HSPF [7,8] model that is core of the Chesapeake Bay Program (CBP) [9,10]. Like other watershed models, it requires information about climatic conditions and flow data for streams. These data sets are available from the web; however, substantial effort is required to download all the information needed and convert it to the proper format for landscape modelling. Each time we move from one sub-watershed to another, this effort must be repeated.

It is essential to have the right software tools and consensus between data providers (in this case, NOAA and USGS) and modellers (in this case, CBP) to make sure that the watershed model can access the data needed as a standard pre-processing, setup routine, when the data will be found and downloaded for further model runs.

In addition to climatic and water data, the watershed models are linked to socioeconomic information required for landuse coverages and calculation of loading factors. This leads to further exploration of linkages to census data available from the US Census Bureau [56]. The census data is organized according to census blocks and tracts, which have nothing in common with the watershed and subwatershed spatial structures assumed in watershed modelling. We need additional preprocessing to reorganize and resample this data to make it available for the model.

Additionally, these standardized protocols for data access should be available to any models in the CCMP directory and beyond. For example, the PSU PIHM [11,12] watershed model has similar data requirements, but runs over a triangular spatial grid. Resampling procedures should provide similar data access for a different geometric structure used in this model. The output from these models can be then piped into Bay models such as the CBP-QUAL-ICM or ChesROMS, which is part of the CCMP open source distribution.

4.4 Ontology and semantics

Exploring differences in vocabularies for variable names, units, collection methods, sampling mediums, etc., used by local data collection agencies and projects, and developing semantic cross-walks, comprise a critical interoperability component for building community networks. Ontologies can be harmonized effectively via ontology-focused “variable tagging” workshops, in which measured variables are mapped to broader concepts, which in turn belong to even broader concepts eventually ending at the "top concept." The key to this effort is that at lower levels a certain degree of heterogeneity must be accepted because experience shows that it is extremely difficult to reach consensus on what to call a certain variable. It is more prudent (and in a way provides a much higher degree of harmony and as such acceptance) to permit variety at this level to obtain buy-in because a high degree of flexibility is maintained, but then merge different variables at higher concepts and use the higher (or more general concepts) as an entry for search queries. Furthermore, the effort must reflect the continuing advances in cyberinfrastructure tools and resources which will require a flexible or evolutionary approach if these advances are to reach the domain community. This work will leverage the initial set of tools for registering ontologies, ontology tagging, and ontology-based search developed within GEON and CUAHSI HIS. For hydrologic data, these tools include HydroSeek [57], an ontology-aided search engine developed at Drexel university that can search over USGS, EPA, and several local observation networks including the CIMS databases, and the ontology tagger tool [58] which is a graphical user interface that helps users associate variable names in an ODM database with concepts in a hydrologic ontology, to make the database available for ontology-based search. The underlying ontologies are currently focused on physical and water quality parameters that are of primary interest to hydrologists or environmental engineers, hence they need to be build up to encompass a larger spectrum that is also of importance to for example oceanographers and ecologists.

4.5 Serving data for participatory modelling efforts

In recent years, there has been a shift from top-down prescriptive management of water resources towards policy making and planning processes that require on-going active engagement and collaboration between stakeholders, scientists, and decision-makers. Participatory modelling is the process of incorporating stakeholders, often including the public, and decision-makers into an otherwise purely analytic modelling process to support decisions involving complex ecological questions. It is recognized as an important means by which non-scientists are engaged in the scientific process and is becoming an important part of watershed

planning, restoration, and management. The development of unique, practical, and affordable solutions to ecological problems is often best accomplished by engaging stakeholders and decision makers in the research process. These group modelling efforts require specific types of models and data to be successful. These modelling tools are usually simpler than what we find in full-fledged research models, and for example might use Excel, Stella, MATLAB, or Mathematica as a means of joint learning and system representation. The scoping models that are produced are designed to gain shared experience about the system and build consensus among stakeholders. However, they also require data to make them run. Moreover, they need a lasting web presence that would support group interactions, and link directly to diverse data and modelling tools.

4.6 Community consensus-building and testbed processes

Various communities need to be engaged in as many ways as possible. This can be achieved by the following means:

1) On-going "cyber-seminars" for which selected participants would submit white papers in advance for consideration by the community. On specific days, discussions of the content could take place using simple web conferencing tools such as Skype, Webex, etc. Follow-up discussion would take place on a twiki over a prescribed period of time (one to two weeks). The results of discussion could be incorporated into the initial white paper(s) and re-posted to the community, which could then decide on subsequent action: e.g., allow the paper to form the basis for subsequent research and development work; redirect research currently underway; submit for presentation to a conference; etc. The twiki could be used for further feedback and subsequent results as needed.

2) Provide means for a Web 2.0 approach to obtain community rating and annotation of submitted resources. Each key resource, such as source data sets, derived data sets, metadata catalogs, software tools developed to search the metadata and actual data sets, etc., will have a commonly identified means in the portal for users to rate the resource and enter comments explaining their ratings. The comments are essential to enable the resource developers to understand and respond to any problems or other issues. These comments and follow-up will be posted to resource-specific twikis for tracking and archive.

3) Support Inter- and intra-community ontology development with a Web 2.0 approach, along the lines of urbandictionary.com. Terms of reference, classification systems, valid values for coded domains, etc., will be posted to twikis allowing all users to see and vote on competing definitions. As specific issues arise that require more extensive discussion for resolution, the community will be called together for a cyber-seminar.

Solving the cultural issues mentioned previously is more difficult than the institutional issues, because these very often come down to working with attitudes, opinions, and beliefs of strong-willed individuals who may resist external directives. But here too, lessons are being learned and progress is being made. An individual field scientist can hardly be faulted for reluctance to share her data when she doesn't know the motives and applications a given user might have in mind. She may not see or accept the value to her of allowing her data to be accessed and used without direct involvement. But if, say, someone from the cyberinfrastructure community were to take a personal interest in the work, and earn trust based on a relationship of shared understanding, then the perspective is more likely to open. One needs to be able to see and trust ways of benefiting from the synergistic implications of contributing one's data to a greater body of knowledge. (It is also likely that some form of external rewards could help to further encourage this outcome.) Certainly, even with personal attention and effort to create such bridging relationships with the IT community, some scientists will still reject the importance of sharing their data, and of working with others to use community-based information models and data exchange mechanisms. This is just part of human nature. Such scientists may yet respond to firm directives from funding agencies to adhere to community information models and best practices, but the difference between motivated and unmotivated adherence to recommended standards and practices can make a big difference in the fitness of data for other uses. It should

also be mentioned that funding resources and funding agencies should better reflect the changing environment described above, and institutions must find new metrics for evaluating excellence of scientists who commit to community data and model development.

In summary, while the focus to date on interoperability standards and tools has been largely technical, we need to acknowledge and work with the social aspects of it now. In order to move forward, we must now focus on forming consensus around information and modelling requirements and architectures, making the differences between current information models and workflows explicit, helping to build solid relationships between IT and subject domain scientists, and otherwise engaging the various stakeholder communities in these activities as much as possible.

REFERENCES

[1] Collaborative Large-Scale Engineering Analysis Network for Environmental Research (CLEANER): An Engineering Cyberinfrastructure "Test Bed" (CLEANER), project office (http://cleaner.ncsa.uiuc.edu/home/)

[2] Consortium of Universities for the Advancement of Hydrologic Science, Inc. (http://www.cuashi.org)

[3] National Ecological Observatory Network (http://neoninc.org/)

[4] GEON; The Geosciences Network (http://www.geongrid.org)

[5] The Long Term Ecological Research Network (http://www.lternet.edu/)

[6] The WATer and Environmental Research Systems Network (WATERS Network) (http://watersnet.org/)

[7] Donigian, A.S., J.C. Imhoff, B.R. Bicknell, and J.L. Kittle. 1984. Application Guide for Hydrological Simulation Program - FORTRAN (HSPF). Athens, GA, U.S. EPA, Environmental Research Laboratory.

[8] Donigian, A.S., Jr., B.R. Bicknell, and J.C. Imhoff. 1995. Hydrologic Simulation Program - FORTRAN (HSPF). Chapter 12 in Computer Models of Watershed Hydrology, V.P. Singh, Ed., Water Resources Publications, Littleton, CO.

[9] Donigian, Jr., A.S., B.R. Bicknell, A.S. Patwardhan, L.C. Linker, C.H. Chang, and R. Reynolds. 1994. Chesapeake Bay Program Watershed Model Application to calculate bay nutrient loadings. U.S. EPA Chesapeake Bay Program Office, Annapolis, MD.

[10] http://www.chesapeakebay.net/temporary/mdsc/community_model/about.htm

[11] Duffy, C.J. 2004. Semi-discrete dynamical model for mountain-front recharge and water balance estimation, Rio Grande of southern Colorado and New Mexico, Pages 255-271 in: J.F. Hogan, F.M. Phillips, and B.R. Scanlon (eds.), Groundwater Recharge in a Desert Environment: The Southwestern United States. Water Science and Applications Series, vol. 9, American Geophysical Union, Washington, D.C.

[12] Qu, Y., and C. J. Duffy (2007), A Semi-Discrete Finite-Volume Formulation for Multi-Process Watershed Simulation, Water Resour. Res., doi:10.1029/2006WR005752

[13] http://www.epa.gov/waterscience/basins/index.html

[14] http://ccmp.chesapeake.org/CCMP/models/ChesROMS/index.php

[15] Cerco, C., and Cole, T., 1994: Three-dimensional eutrophication model of Chesapeake Bay, Technical Report EL-94-4, US Army Engineer Waterways Experiment Station, Vicksburg, MS.

[16] http://el.erdc.usace.army.mil/elmodels/icminfo.html

[17] http://www.ccalmr.ogi.edu/CORIE/modeling/elcirc/

[18] Zhang, Y.-L., Baptista, A.M. and Myers, E.P. (2004) "A cross-scale model for 3D baroclinic circulation in estuary-plume-shelf systems: I. Formulation and skill assessment". Cont. Shelf Res. 24: 2187-2214.

[19] Cerco, C.F., 1995: Simulation of long-term trends in Chesapeake Bay eutrophication, Journal of Environmental Engineering, vol. 121(4), p. 298-310.

[20] Sheng, Y.P., 1986: A Three-dimensional Mathematical Model of Coastal, Estuarine and Lake Currents Using Boundary Fitted Grid, Report No. 585, A.R.A.P. Group of Titan Systems, New Jersey, Princeton, NJ, 22 p.

[21] http://www-nml.dartmouth.edu/circmods/gom.html

[22] http://chartmaker.ncd.noaa.gov/csdl/op/c3po.html

[23] Xu, J. and R.R. Hood. (2007). Modeling biogeochemical cycles in Chesapeake Bay with a coupled physical-biological model. Coastal Shelf Est. Sci.

[24] http://www.hydroqual.com/wr_rca.html

[25] Di Toro, D.M. 2001. Sediment Flux Modeling. J. Wiley and Sons. New York.

[26] Acock, B. and J. F. Reynolds (1990). Model Structure and Data Base Development. Process Modeling of Forest Growth Responses to Environmental Stress.. R. K. Dixon, R. S. Meldahl, G. A. Ruark and W. G. Warren. Portland, OR, Timber Press.

[27] http://wise2.jrc.it/wfdview/php/index.php

[28] Baru, C. (2005) GEON: The GEON Grid Software Architecture. ESRI International User Conference, San Diego, August 2004.

[29] http://www.chesapeakebay.net/data/index.cfm

[30] http://geon16.sdsc.edu:8080/gridsphere/

[31] http://www.cuahsi.org/his

[32] http://river.sdsc.edu/wateroneflow/

[33] http://river.sdsc.edu/DASH

[34] http://watersnet.org/wtbs/wtbs08/index.html

[35] http://watersnet.org/wtbs/wtbs10/index.html

[36] CUAHSI Observations Data Model (ODM), (http://www.cuahsi.org/his/odm.html)

[37] Ouskel, A. and A. Sheth. 1999. Semantic Interoperability in Global Information Systems. ACM SIGMOD Record Vol. 28 No. 1 pp. 5-12.

[38] GML 3.1.1 Application schema for Earth Observation products, 2007. Editor Jerome Gasperi, OGC Best Practices Document, OGC 06-080r2 (http://portal.opengeospatial.org/files/?artifact_id=22161).

[39] [WaterML07] CUAHSI WaterML, Editors Ilya Zaslavsky, David Valentine, Tim Whiteaker, 2007, OGC Discussion Paper, OGC 07-041r1 (http://portal.opengeospatial.org/files/?artifact_id=21743)

[40] Brown Gaddis, E. J., Vladich, H., and Voinov, A. 2007. Participatory modeling and the dilemma of diffuse nitrogen management in a residential watershed. Environ. Model. Softw. 22, 5 (May. 2007), 619-629.

[41] A. Castelletti and R. Soncini-Sessa, 2006. A procedural approach to strengthening integration and participation in water resource planning Environmental Modelling & Software, Volume 21, Issue 10, Pages 1455-1470

[42] Joanne Tippett, 2005. The value of combining a systems view of sustainability with a participatory protocol for ecologically informed design in river basins Environmental Modelling & Software, Volume 20, Issue 2, Pages 119-139

[43] Roberts, N., 2004. Public deliberation in an age of direct citizen participation. American Review of Public Administration 34(4), 315-353.

[44] Korfmacher, K. S., 2001. The politics of participation in watershed modeling. Environmental Management 27(2), 161-176.

[45] Bergamaschi, S. Castano, S. and Vincini, M. 1999. Semantic integration of Semi-structured and Structured data Sources. Special Issue ACM SIGMOD Record Vol 28. No; 1 pp. 54-59.

[46] Bishr Y. 1998. Overcoming Semantic and other barriers to GIS interoperability. International Journal of Geographic Information Science. Vol. 12 No.4. pp. 299-314.

[47] Castillo, J. A., Silvescu, D. Caragea, J. Pathak, V. Honavar. 2003. Information Extraction and Integration from Heterogeneous, Distributed, Autonomous Information Sources: A federated Ontology-Driven Query-Centric Approach. IEEE International Conference on Information Reuse and Integration (IRI 2003) Las Vegas, NV.

[48] Garcia-Solaco, M., Saltor, F. and Castellanos, M. (1996). Semantic Heterogeneity in Multidatabase Systems, in Bukhres, O. and Elmagarmid, A (Eds) Object-Oriented Multidatabase Systems. Prentice Hall: Englewood Cliffs, NJ. pp 129-202.

[49] Kashyap, V. and Sheth, A. (1998). Semantic heterogeneity in global information systems: The role of metadata, context and ontologies. In Papazoglou, M. P. and Schlageter, G., editors, Cooperative Information Systems, pages 139--178. Academic Press, San Diego.

[50] McLeod, D. Fang, D. and Hammer, J. 1991. The Identification and Resolution of Semantic Heterogeneity. In Proceedings of First International Workshop on Interoperability in Multidatabase Systems. Kyoto, Japan, April 1991.

[51] Shahabi C. and Sheth, A. 1991. Semantic Issues in Multidatabase systems. SIGMOD special issue. Vol 20. No. 4.

[52] Welty, C. and N. Guarino 2001. Supporting Ontological Analysis of Taxonomic Relationships. Data and Knowledge Engineering, 39 (1), 51-74

[53] http://www.opengeospatial.org/standards/cat

[54] http://river.sdsc.edu/wiki/CSIRO%20Collaboration.ashx

[55] http://river.sdsc.edu/awdip/

[56] http://www.census.gov/geo/www/tiger/

[57] http://www.hydroseek.org

[58] http://water.sdsc.edu/tagger

[59] http://www.ogcii.org/bod

[60] Chesapeake Bay Environmental Observatory (http://ccmp.chesapeake.org/CBEO/resources.php)

[61] Observations and Measurements, 2006. Editor: Simon Cox, OGC Best Practices Document, OGC 05-087r4 (http://portal.opengeospatial.org/files/?artifact_id=17038).

[62] http://www.opengeospatial.org/standards/gml

[62] http://www.opengeospatial.org/standards/sensorml

[63] http://www.opengeospatial.org/standards/tml

[64] http://portal.opengeospatial.org/files/?artifact_id=12846

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

The Great Rivers Modeling Framework and Decision Support System

B. Eckman [a], **C. Barford** [b] **and P. West** [c]

[a] *IBM Big Green Innovations, 1475 Phoenixville Pike, West Chester, PA 19380, baeckman@us.ibm.com*

[b] *University of Wisconsin, Madison, 1710 University Avenue, Madison, WI 53726, clbarford@wisc.edu*

[b] *The Nature Conservancy, 633 W. Main St., Madison, WI 53703 pwest@tnc.org*

The health of the world's freshwater ecosystems is fundamental to the health of people, plants and animals around the world. The sustainable use of the world's freshwater resources is recognized as one of the most urgent challenges facing society today. An estimated 1.3 billion people currently lack access to safe drinking water, an issue the United Nations specifically includes in its recently published Millennium Development Goals (http://www.un.org/millenniumgoals/). To address the scale and urgency of this challenge, IBM is collaborating with The Nature Conservancy and the Center for Sustainability and the Environment (SAGE) at the University of Wisconsin, Madison to develop an innovative decision support tool and modeling framework for improved management of water resources worldwide.

The Great Rivers Modeling Framework and Decision Support System (DSS) is designed to help policy makers and a variety of stakeholders (farmers, fish & wildlife managers, hydropower operators, et al.) to assess, come to consensus, and act on land use decisions representing effective compromises between human use and ecosystem preservation/restoration. Initially focused on Brazil's Paraguay-Paraná, China's Yangtze, and the Mississippi Basin in the US, the DSS will integrate data and models from a wide variety of sectors, including water balance, water quality, carbon balance, crop production, hydropower, biodiversity, and economic valuation of ecosystem services. Our aim in this project is to build a community framework that supports the sharing and coupling of component models by modelers and other stakeholders worldwide.

The challenges of coupling component models in a framework are many and various. Common data and metadata representations are needed to represent interfaces between component models. Rules must be developed that express constraints on how and if component models may be coupled into a simulation, and which components are swappable. Once defined, simulations must be calibrated and validated; since autocalibration is arguably not yet well developed enough for unsupervised use, this step typically involves much manual intervention. The uncertainty of simulations built out of coupled models must be estimated and conveyed to the user. Data sets in diverse formats with disparate semantics and a variety of associated software must be integrated. Data must be interrogated at different temporal and spatial resolutions.

A key question in designing a community modeling framework is: on which level should the component models be defined? Two choices that readily come to mind are: 1) on the model level, resulting in components that are large multi-function programs like SWAT (Soil and Water Assessment Tool, Texas A&M University: www.brc.tamus.edu/swat/avswat.html); and 2) on the sector level, with components such as

water balance, water quality, crop production, carbon storage. But critical requirements for the framework are: defining interfaces between component models; expressing model semantics and assumptions in metadata; and developing rules that enforce constraints on how component models may be composed together into a simulation. These requirements are much more readily met if the model components are atomic, i.e., if each one has one and only one function. This is not true of large multi-function models, nor even of sector-level components like water balance. Following this reasoning, we are investigating defining model components at a lower level. For example, model components that deal with water balance include: Net radiation, Precipitation, Canopy loss, Transpiration, Soil evaporation/sublimation, Soil moisture, Runoff, Percolation, and Ground Water Recharge.

Our community modeling framework supports six basic tasks: 1) register and calibrate new component models and datasets; 2) define simulation workflows as compositions or couplings of model components; 3) run the simulations; 4) view the results of the simulations; 5) run searches over the database of simulation inputs, metadata, and results; and 6) tag, annotate, and interactively chat about models and simulations. Three levels of user are defined. Expert users may perform all 6 tasks; non-expert users perform tasks 4-6 only; and administrative users handle security, access control and system maintenance and operations. Sample searches over simulations include:

- "Show me all the grid cells from any simulation's output where fertilizer < x and crop yield > y and water quality > z", where x, y and z are user-defined threshold values.
- "In all streams that are order >4, within ½ mile buffer, what percentage of land is in crop and fertilized?"

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

A solute and water flux library for catchment models

Philipp Kraft, Kellie B. Vaché, Lutz Breuer, Hans-Georg Frede

Justus-Liebig Universität Gießen, philipp.kraft@umwelt.uni-giessen.de, kellie.vache@umwelt.uni-giessen.de, lutz.breuer@umwelt.uni-giessen.de, hans.georg.frede@umwelt.uni-giessen.de

Keywords: Hydrology; Catchment model; Model coupling; Object orientation

MOTIVATION

Simple coupling of models can be inhibited by overlapping process descriptions and differing spatial and temporal scales. However the process of bringing together multiple models can be enhanced if at least one of the models provides capabilities to adjust itself to the more specific requirements of other models.

Key design considerations in the development of such models include successive layers of abstraction in the process description, and efforts to preserve generality of process description in deep layers of specialization. Object orientated design of model code, using generalized interfaces, is one way to create such functionality. The advantages of an object oriented code structure for hydrological models, even without the utilization of inheritance, were outlined by Wang et al. [2005].

This paper focuses on the development of a coupled model system designed to evaluate landscape level nitrogen fluxes across multi-use watersheds. The goal is to outline the general steps and decision-making processes that are required in the development of coupled systems, with particular attention on the development of model code to facilitate coupling. This case study focuses on the coupling a hydrologic catchment modeling framework, CMF (see Vaché and McDonell [2006]), with a plot scale biogeochemical nitrogen flux model MOBILE, and extension of the DNDC model (Li et al. 1992).

The MOBILE (MOdular BIosphere SimuLation Environment) model is a plot scale model framework describing biomass production, decomposition, mineralization, nitrification and denitrification. CMF is a DHSVM (Wigmosta et al. [1994]) style catchment model, primarily designed for the calculation of catchment runoff generation, mean residence times, and of travel paths of conservative tracer.

THE MODEL FRAMEWORK

As a fully distributed simulation model, CMF is the natural framework for the coupled model, where the coupling invokes a number of instances of the lumped MOBILE model, each allocated to represent a different location within the CMF landscape. Despite the simplicity of the concept, a wide variety of details are involved, beginning with the fact that although the original CMF software was designed to operate in different modes of process description depth, its internal structure was not explicitly designed to facilitate the modular use of the different components. The initial step was therefore a redesign of the CMF code base to transform the solitary CMF model, including a GUI, into a solute and water flux library for catchment models, considering different kinds of application modes. The previous CMF version was implemented in C++ using the Microsoft Foundation Classes, and therefore only able to run on Windows PC. The library is implemented in strict ANSI C++ which allows compilation on any platform providing a standard conform C++ compiler.

SPATIAL RESOLUTION

The simulations are spatially-distributed with the catchment domain divided into model cells defined by location, center elevation, area, and adjacent neighbors. Under this definition the model is not restricted to square raster approaches, allowing to run the model in finer resolution in areas of greater spatial heterogeneity. A similar approach was taken by Band et al. [2001] in an application of the RHESSys model, where a fine resolution was used to capture the highly variable riparian dynamics, and was coarsened in less variable upslope regions. This type of flexibility is designed to balance the need to capture spatial dynamics with the pragmatic need to maintain acceptable computational times. Nonetheless, within the CMF framework, the use of a regular squared raster is still supported.

The cell class, called "UpslopeCell", is an abstract class. It implements the neighborhood relation to other cells, though, as an abstract class, the actual transport of water and solutes is not included. Interfaces are provided as pure virtual functions to access information about the cell state, for example depth of the saturated zone, amount of ponded water, or average saturation. A complete subbasin is described as a collection of interconnected UpslopeCells. Inheritance is used to support different implementations of the UpslopeCell class. The current version of the framework currently supports two implementation of UpslopeCell: the vertical homogenous cell and the layered cell. **Figure 1** outlines the this approach

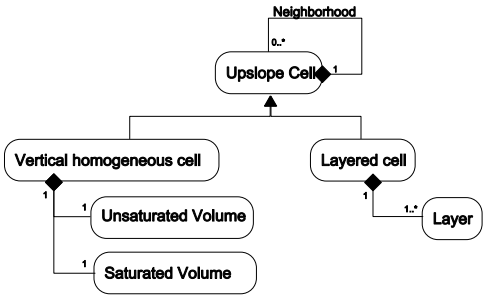


Figure 1: UML 2.0 diagram of the UpslopeCell class.

VERTICAL RESOLUTION

The key assumption of the vertical homogenous cell is that the shallow soil can be described using two layers, each with constant soil properties. The upper layer corresponds to the unsaturated zone, and the lower layer to the saturated zone. Vertical flow is assumed in the unsaturated zone, horizontal flow in the saturated zone. The size of the storages is variable according to the saturated depth. The mass balance equations of the two storages are based on DHSVM and described in detail in Vaché and McDonnel [2006]. The layered cell is tailored for application with biogeochemical models, like MOBILE, that require a detailed description of soil layer heterogeneity and concentration profiles. The vertical resolution of the soil layers is constant, and each layer is either unsaturated (with vertical flow only) or partially or fully saturated (with lateral flow). Each water store derives from the abstract class "StateVariable", and implements a function "Derive". This function calculates the storage change rate and includes a get/set function pair to interact with the current state of the storage variable. **Figure 2** shows the inheritance graph of "StateVariable".

The class "Concentration" also inherits from StateVariable and represents the concentration of an arbitrary conservative tracer. Derivates of "Concentration" can implement concentrations of species reacting with the environment, via simple time dependent decay functions, or via more complex transformation functions. Each water storage maintains a collection of concentrations.

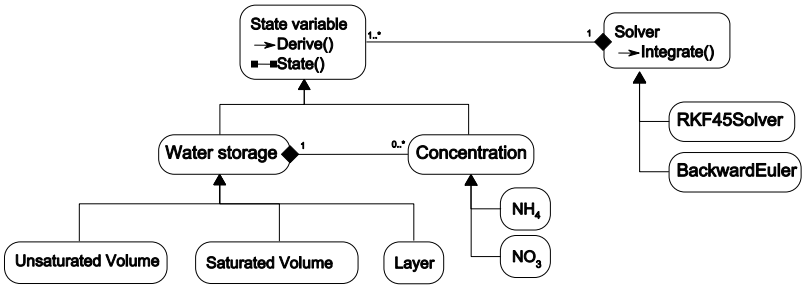


Figure 2: Inheritance diagram of "StateVariable"

The existence of a "StateVariable" abstract class allows the model to be solved by a central ODE-solver, which can be selected from a list of available solvers. Hillyer et al. [2003] built a COM-based general purpose modelling system, with a comparable structure.

MODEL COUPLING

To give the user the full power to combine these different features, a wrapper for scripting languages is provided, using a Simplified Wrapper and Interface Generator (SWIG, www.swig.org). Application is made more generic in that the user is not required to provide input files in a specific format, but rather writes a configuration script in Python or another SWIG supported language, and merges the environments. Any data source accessible through the scripting language, like (spatial) databases, plain text files, or XML files and web services can be incorporated. Many script programs are written with the purpose to build input files required by environmental models, with unit conversions and insertion of default values, where the direct use of a script language for model setup can spare the step of writing the specific model input files.

ACKNOWLEDGEMENTS

This work was funded by the NitroEurope IP (www.nitroeurope.eu)

REFERENCES

Band, L. E., C. Tague, P. Groffman, and K. Belt, Forest ecosystem processes at the watershed scale: hydrological and ecological controls of nitrogen export. *Hydrological Processes* 15, 2013–2028, 2001

Hillyer, C., J. Bolte, F. van Evert and A. Lamaker, The ModCom modular simulation system, *Europ. J. Agronomy*, 18, 333-343, 2003

Li, C., S. Frolking and T. A. Frolking, A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. *Journal of Geophysical Research*, 97, 9759-9776, 1992

Wang, J., T. E. Endreny and J. M. Hasset, A flexible modeling package for topographically based watershed hydrology, *Journal of Hydrology*, 314, 78-91, 2005

Wigmosta, M., L. Vail, and D. Lettenmaier, Distributed hydrology-vegetation model for complex terrain. *Water Resources Research* 30(6):1665–1679, 1994

Vaché, K. B. and J. J. McDonnell, A process-based rejectionist framework for evaluating catchment runoff model structure, *Water Resources Research*, 42(W02409), doi:10.1029/2005WR004247, 2006

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Promoting access to and reuse of modelling projects - A picture book approach

Ralf Seppelt
Helmholtz Centre for Environemtal Research, Department for Computational Landscape Ecology, Permoserstraße 15, 04318 Leipzig, Germany, ralf.seppelt@ufz.de,

Abstract: Use and misuse of simulation models, model documentation, model comparisons and model integration are topics that never vanished from discussion on modelling and simulation. Especially under the frame of collaborative modelling these issues again gain importance. This contribution to the discussion does not offer another technological approach for this task. Here, the focus is put on documenting and assessing models and by this supporting the part of (re-) application of models. Using the metaphor of a collapsible picture book with which modellers, practitioners, stakeholders and other can access regions, models and modelling questions. The approach is suitable for scientific groups, projects or even journals for storing and maintaining their knowledge accumulated through simulation models.

***Keywords**:* Model documentation, Model Meta-Data, Collapsible Picture Book

1. Introduction

In volume 100 of Ecological Modelling Müller (1997) stated that “we produced an enormous redundancy” by continuously developing simulation models. Actually this didn't change till today. Thus there is a need to bring modelling and simulation to a next level of applicability by the development of collaborative modelling platforms. This point has been made very strong in the contribution by Voinov et al. (2008), which clearly shows the pros and cons of collaborative modelling using a very technically point of view. In this short response I argue, that this line of thought can be extended and put into application. Within this commentary I would like to add some thoughts on the application part of these developments.

2. A “picture book” concept of making models more useable and consistent

Practitioners as well as scientists approach modelling and simulation developments following three different tracks. It is either

(1) a certain region, in which a modelling study is to be conducted,
(2) a certain problem or question to be tackled, and
(3) a specific modelling approach to be applied or developed.

These three approaches interrelate to each other. E.g. the problem reduces the set of possible model applications or developments; the region of course is closely related to the problems but not exclusively. This makes documenting models, simulations studies and projects difficult. All these three way of approaching a simulation study are of equivalent importance.

The metaphor that appears is that this could be tackled as a collapsible picture book, well known for young children. I mean these picture books that show for instance different animals, the back, the centre and the front on different cards. E.g. you can put together tail and back of a lion with the centre of an elephant and the head of zebra. Actually this is very likely to model application you need to identify solution for high nutrient load (middle part) in a certain catchments (e.g. Patuxent river), first part with a certain model approach, see figure upper part, see figure 1.

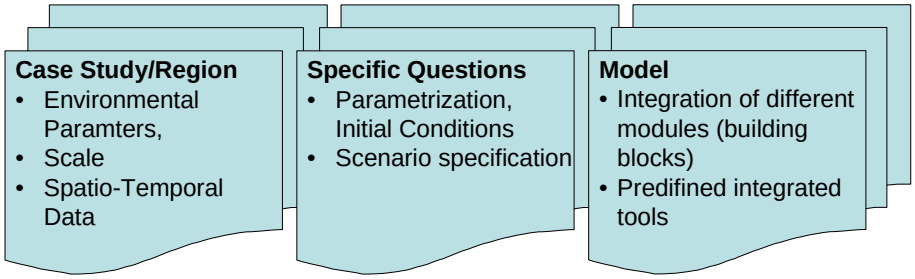


Figure 1. A "picture book" on model applications

Technically spoken, the implementation of such a collapsible picture book is a database with two main n:m-relations linking data sets on regions with scenarios and models. The major argument for such a generic database storing different modelling studies is that there are

- many models that can cope with the same problem fields (such as nutrient loads in catchments)
- many sites that have be analysed by different models (for instance the Elbe River basin)
- many model that have be applied to different regions/sites

The linkage is performed by scenarios.

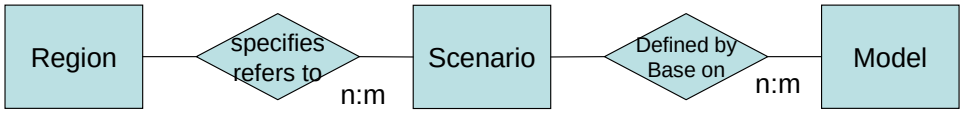


Figure 2. ER-Diagram of "picture book" on model applications.

3. Application and Discussion

The benefit of this approach can be seen for different applications First, research groups need to put effort in maintaining their suite of modelling developments. This task is of high importance, as team members are leaving (Ph.D. thesis finished, projects completed). This is crucial for not reinventing wheals and starting developments all over. Thus documentation and simplification of the model reuse is important. There are several developments available, but did not managed to be part of daily modelling work (Benz, 1997) of even shut down (Plentinger, 1996).

Second, the beginning phase of each new environmental research project is characterized by the question: Is there something out in the universe that we can use of our problem, task or scenario? Ususally meta documentation of models and databases (Bent, 1997) are useless unless results of the application show up. Thus all three parts of our picture book are of high importance.

Third, also journals such as Environmental Modelling and Software should aim at providing a database on their publications and products. The arguments above do support such an approach also for this field of application.

ACKNOWLEDGEMENTS

Many thanks go to Alexey Voinov, asking me to review his contribution to workshop 7 of the 2008 conference. By this I got the chance to put some vague ideas into words and this contribution to the discussion.

REFERENCES

Benz, J. and M. Knorrenschlid (1997). "Call for a common model documentation etiquette." Ecological Modelling 97: 141-143.

Plentinger, M.C. and Penning de Vries, F.W.T. (editors) (1996): CAMASE: a Concerted Action for the development and testing of quantitative Methods for research on Agricultural systems and the Environment, Wageningen University, The Netherlands

Müller, F. (1997). "State- of- the- art in ecosystem theory." *Ecological Modelling* 100: 135-161.

Voinov, A. (2008) “Community-based software tools to support participatory modeling: a vision”. This issue.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Actors and factors – bridging social science findings and urban land use change modeling

D. Haase[a], S. Kabisch[b], A. Haase[b], T. Filatova[c], A. van der Veen[d], T. Tötzer[e], W. Loibl[e], S. Scatasta[f], S. Schetke[g], A. Zuin[h], F. von Walter[i]

[a] UFZ – Centre for Environmental Research, Dept. of Computational Landscape Ecology, Permoserstr. 15, D-04318 Leipzig, Germany, (dagmar.haase@ufz.de)
[b] UFZ – Centre for Environmental Research, Dept. of Urban and Environmental Sociology, Permoserstr. 15, D-04318 Leipzig, Germany, (sigrun.kabisch@ufz.de)
[c] Department of Water Engineering and Management, University of Twente, The Netherlands, T.Filatova@ctw.utwente.nl
[d] International Institute for Geo-Information Science and Earth Observation, The Netherlands, Veen@itc.nl
[e] Austrian Research Centers – ARC GmbH, systems research, A-1220 Vienna, Austria (tanja.toetzer@arcs.ac.at, wolfgang.loibl@arcs.ac.at)
[f] Centre for European Economic Res., ZEW, L7,1 Mannheim, Germany (scatasta@zew.de)
[g] Institute of Geodesy and Geoinformation, Dept. of Urban Planning and Real Estate Management, University of Bonn, Germany, (schetke@uni-bonn.de)
[h] Edinburgh College of Art, OPENspace Research Centre,UK (azuin@eca.ac.uk)
[i] Georg-August-Universität Göttingen, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany, (fwalter2/jbarkma@gwdg.de)

Abstract: Recent uneven land use dynamics in urban areas resulting from demographic change, economic pressure and the cities' mutual competition in a globalising world challenge both scientists and practitioners, among them social scientists, modellers and spatial planners. Processes of growth and decline specifically affect the urban environment, the requirements of the residents on social and natural resources. Social and environmental research is interested in a better understanding and ways of explaining the interactions between society and landscape in urban areas. And it is also needed for making life in cities attractive, secure and affordable within or despite of uneven dynamics.

The position paper upon "Actors and factors – bridging social science findings and urban land use change modeling" presents approaches and ideas on how social science findings on the interaction of the social system (actors) and the land use (factors) are taken up and formalised using modelling and gaming techniques. It should be understood as a first sketch compiling major challenges and proposing exemplary solutions in the field of interest.

Keywords: *social science concepts; behavioural rules of agents; empirical knowledge; land use change models; demography; statistical functions; survey data; DPSIR; impact analysis; quality of life; ecosystem services.*

1. THE CHALLENGE

Recent uneven land use development patterns and demographic change in urban areas challenge modellers, social scientists and spatial planners. Processes of growth and decline specifically affect the urban environment as well as the demand on natural resources. Social and environmental research is interested in a better understanding and ways of explaining the interactions between society and landscape in urban areas. Planners and policy makers are in charge of making life in cities attractive, secure and affordable. Moreover, sustainability and green infrastructure gather increasing importance in cities' profiles.

The discussion paper upon "Actors and factors – bridging social science findings and urban land use change modeling" presents some basic ideas on how novel social science findings on the interaction of the social system (actors) and the landscape (factors) are taken up and formalised using modelling techniques. It should be understood as a first sketch compiling major challenges and proposing exemplary solutions in the field of interest.

The discussion paper is structured as follows. After this introduction some major and novel social science and urban land use findings are presented that represent operational and empirical challenges for urban (land use change) modelling. In a third section we select a number of different social science concepts in terms of their core ideas and novelties. Proposals are made how to "translate" these social science concepts and findings into a model and thus to formalise it, make it transferable. In addition, each example is used to ask important questions that have been raised during this modelling exercise. These questions should particularly encourage the discussion at the workshop. Participants and interested colleagues are invited to contribute to both the discussion paper and the workshop in Barcelona.

2. SOCIAL SCIENCE ISSUES TO BE ADDRESSED

In terms of social science concepts and findings on urban systems, particularly demographic and land use dynamics questions such as

- How does demography modify human demands on natural resources, real estate, housing space and transportation networks?
- What are the spatial, land use and environmental effects caused by demographic change such as an increase of new household types like single households, cohabitation, single-parent families or flat-sharers?
- What are the socio-environmental consequences of ongoing urban sprawl and of the opposite process of urban shrinkage characterized by large residential and commercial vacancies?

are high on the political agenda of many cities worldwide. Major factors of uncertainty can be found in the complex structure of the social systems, demographic and household pattern as well as of the respective housing markets.

Often, land use transitions are closely connected with humans acting on the real estate and land market. Land use changes thus can be understood as the consequences of actions. Since they are a heterogeneous group the behaviour of the "urban actors" is characterized by highly heterogeneous and controversy decisions that directly impact on and shape of land use change (in terms of the intensification of land cultivation, the amount of land take as well as forms of land abandonment).

This is the field of social sciences (sociology, political science, psychology, spatial planning) which carry out field research on effects of demographic change, urban sprawl and shrinkage. Their work delivers innovative empirical results in form of questionnaire survey data, series of interviews, perception data, agent profiles, behaviour settings through document analysis and observation.

Land use change model approaches have their strengths in setting up causal relationships between variables and to quantify them. Models generally aim at depicting multiple relationships and complex systems in a formalised and less complex way. They use major

variables to realise important stocks, flows and communication processes to reduce the real complexity to a point where we can see and analyse the influence of single parameters of the system. In urban land use change modelling there have been developed different modelling approaches such as system dynamics, cellular automata, and agent-based modelling.

Incorporating social science concepts and knowledge into modelling – bridging actors and factors – together we shall discuss the following questions in accordance with the above mentioned ones:

- How do we use social science concepts and findings to define the purpose of respective models?
- How to translate them into a) a formal way and b) into a model?
- What kind of empirical evidence we can use to a) feed and b) to verify the plausibility of our models? How can we formalise qualitative outputs from social science studies?
- How "reliable" and transferable case study based social science questionnaire data are in terms of applying them to other cases / regions?
- In terms of agent based models: How do we measure actors' behaviour and translate it into formal agents' rules accurately? Since persons are unique in their decision making, could we define a "mean person" as representative agent for an entire agent class?
- How do we depict trees, priorities and ranking of choices related to a given situation or scenario?
- Can we use a respondent stated choice as realistic? Some studies show that the intention to do some action stated in the survey does not always imply that people will actually do that. Is this fact worth considering in transferring survey data into ABM rules? If so, then how should we account for it?
- And finally, how can we test the plausibility and quality of social science (concepts) based model formalisations?

Bringing together the strengths of both social science and modelling the workshop integrates factors and actors and thus intends to endeavour ways of bridging social science and quantitative as well as qualitative modelling to find answer on the above posed questions and others related to them.

3. ISSUES TO BE DISCUSSED AT THE WORKSHOP

3. 1 Example I: Household diversification by demographic change

At present, large parts of Europe undergo considerable demographic changes (Cloet, 2003; Lutz, 2001). Related processes summarised as a Second Demographic Transition (SDT) such as an individualisation and institutionalised individualism of the (industrialised) society (Beck and Beck-Gernsheim, 2002), a worldwide decline in total fertility (up to <1.1 children per woman; Council of Europe, 2005) and a rising life expectancy, which results altogether in ageing of the total population. Furthermore demographers report the postponement of childbearing or the renouncement of the family phase in the life cycle (Buzar et al., 2005, 2007), respective changes in age group spectra and size, structure and the stability of household patterns (van de Kaa, 1987; Ogden & Hall, 2000; Haase et al., 2005; Bösch-Supan et al., 2005), which altogether results in a diversification of lifestyles as well as a considerable labour, (replacement) and retirement migration within Europe and worldwide followed by an increasing share of nationalities at national labour and residential markets (Bell et al., 2007).

These socio-demographic alterations enormously impact on lifestyles, expectations on quality of life as well as affiliations and motivations of residential mobility. Those – again –

challenge current land use patterns, housing markets and the development of urban space (Deutsch et al., 2003; Ekins et al., 2003).

Changing household structures in line with this second demographic transition (van de Kaa, 2004; Lestheaghe & Surkyn, 2002) brought about by societal processes, economic developments (rise of the service sector) and new fertility patterns (declining TFR[1]) are assumed to be preconditions of the current household formation: Households have become smaller in size and less stable since individuals shift from one living arrangement to another several times during their life course. In an EU FP5 project households had been classified according to their type based on empirical findings for different European cities[2]. Non-traditional or non-familiar household types such as one-person, cohabiting couples, single parents and young adults sharing a flat account for a good portion of the pluralised landscape of household types since their number has significantly increased. While the pluralisation of living arrangements is to be observed Europe-wide, the concrete forms are dependent upon cultural and societal specifics (Council of Europe, 2005).

Contemporary households act very situation-sensitive, i.e. their decision holds the capacity (and necessity) to adapt to changing internal and external conditions. Households are defined more subject-oriented, and living arrangements are adapted to individual life scripts (Buzar et al., 2005; van de Kaa, 2004; Ogden & Hall, 2000).

In terms of the rural-to-urban gradient, particularly the inner city (here understood as the part of city closed the city centre) is highly adaptable to the preferences and needs of SDT-sensitive households. It exhibits a range of appropriate characteristics – e.g. closeness to the city centre and to places of work, qualification and leisure, easily accessible by public transport. Their housing structures allow for flexible adaptations to changing personal circumstances and often the dwellings there are for rent, even in societies which are predominantly owner-occupied. Especially among non-traditional/non-familiar households, urban life with its central notions of density and diversity of both people and opportunities possesses a high value. In both cultural and symbolic terms, the inner city is transformed by these selected socio-demographic groups, their habits, interests and behaviour. But also for families, inner-city housing seems to be more attractive and desirable than hitherto assumptions of urban research have presumed (cf. a.o. Gober, 1990).

The questions:

For our workshop challenging questions emerge how this demographic transition and the related changes of lifestyles and household types detected by social scientists can be transferred into socio-demographic and spatial (urban) land use modeling? How far we can rely on these novel social science findings, are they replicable or transferable in a more mechanistic way of understanding? Could we “scale” them, aggregate or disaggregate them, too? Answers to such questions will be proposed in this discussion paper as well as touched during the workshop in Barcelona.

3.2 Example II: Newly evolving pattern of urban development

Among urban scholars the hypothesis raised whether current demographic decline, de-economisation and structural crisis are bringing about a new evolving type of urban development. Cities are at the same time sprawling and shrinking, here from arise spatially uneven patterns of development within the borders of an urban region. In doing so, we present a novel social science concept of urban shrinkage that goes beyond the focussed population-related view of shrinkage (Rink et al., 2008). It rather comprises the features of non-growth, on-going sprawl and upcoming reurbanisation alike, or if reurbanisation represents, on the other hand, merely a short interplay in a long-term period of desurbanisation under the condition of demographic change and decrease in population: the core city is gradually shrinking because of demographic decline.

[1] Total Fertility Rate
[2] FR5 EU Project Re Urban Mobil (www.re-urban.com)

Nevertheless, in the course of reurbanisation, the core city will be able to remain, or become again, compared to the outskirts and surroundings, as we can already see in cities that follow a consequent policy of strengthening the inner city. The less abrupt demography change the better the possibilities of cities to force reurbanisation and to support the development of their central areas are. Admittedly, reurbanisation takes place in a selective way and isles of reurbanisation arise within a context of emptiness and shrinking (Haase et al., submitted).

The questions:

There exists a multitude of urban land use change models varying from cellular automata (Silva and Clarke, 2002; Clarke et al., 1997; White et al., 1997; Wu and Webster, 1998; Wu, 1998) over stochastic models (Landis & Zhang, 1998a, b; Landis, 1994), rule- and agent-based approaches (Miller et al., 2004; Waddell, 2002; Ben-Akiva and Lerman, 1985; Horowitz et al., 1986; Klosterman, 1999) displaying urban growth and sprawl phenomena. According to mainly morphological or historical criteria cities expand into their hinterland. Protection of open or arable land is implemented using planning and protection restrictions. Such models assume that most of the urban land use change occurs at the periphery and the urban core remains more or less stable. We know from social and policy science that this is only half of the urban truth: how do we incorporate the concept of urban shrinkage into our land use change models? Is it a process that relates to built-up structures, age of urban structures, distances and accessibility or could it be better explained looking at households' residential or investor's decision-making? What variables do we have to include when we focus on urban shrinkage compared to growth? And how we model when suburban growth, reurbanisation and shrinkage, interact and occur simultaneously?

Using a system dynamics approach we apply this social science concept in a simulation model which uses variables emerging from this concept as descriptors for typical shrinkage related population, household, residential and other land use processes and pattern (McIntosh et al., in press). For the model implementation we had to "translate" non-numerical and only partially quantifiable social science knowledge into model variables such as urban stocks, flows and trends.

In our presentation we show the simulation of different urban development scenarios assuming a range of population and household formation changes, a further increase of single and single-parent households, effects of ageing and related altering demands and preferences on residential space and places to live. Accordingly, we simulate pattern of residential vacancy and potential sites of demolition (Haase & Seppelt, submitted). Based on these simulations we can show under which demographic conditions urban growth and shrinkage simultaneously occur or even dominate the development of a city/urban region.

3. 3 Example III: Understanding human decision-making

Identifying decision criteria and getting closer to human behaviour is fundamental for the understanding of decision-making processes and actions that shape our landscape and its environmental resources (Bharwani, 2006). Often we are confronted with multiple reasons for land use decisions of multiple actors that have a stake on one resource (e.g. land, water resources, soil fertility etc.).

Participatory methods such as Knowledge Elicitation Tools (KnETs) represent a new and reproducible way to formalise this knowledge using computational techniques and, what is more, to implement scenario techniques within the interviews. KnETs can be understood as an amplified methodology of classical social science empirical tools such as interviews and questionnaire surveys. It produces input data for the logic of agent-based models (ABM), decision trees or decision ranking. KnETs link qualitative and quantitative representations of stakeholder knowledge. The fusion of both is realised in form of an iterative process that incorporates a set-up phase of the game conducting interviews, a formalisation phase of the game, an empirical data collection phase of "gaming" with respective respondent groups of interest and a resulting decision tree creation and interpretation phase. Finally, another game round is conducted with a non-involved respondents' group to verify the results of the models and to assess our gained knowledge on decision making (Bharwani, 2006).

In the workshop two applications of the KnETs games methodology will be presented and discussed for firstly, getting insides in criteria and thresholds of decision making by municipal representatives pro or contra soft mitigation paths in flood risk management in the Ukrainian Tisza valley and, secondly, decision making of households to 'reurbanise' and to settle at recycled land within the inner city (success test of the realisation of a so-called 'program of urban villas'). The resulting trees shed light on what knowledge is used for decision making and how different criteria are weighted in these choices.

3. 4 Example IV: Mapping out the social impacts of land use: Soft-GIS and Choice-Based Conjoint Analysis as integrative methods

People's experiences of place are affected by planning strategies as much as community-based behaviours and emotions can affect planning and development. Residents' experiences of their environment are geographically located, since people can attach their experiences, memories, and feelings to specific locations. By allowing participants to locate these experiences, researchers can not only identify their specific affordances but also analyse and visualize the information for research and planning practices. This example aims to:

- Examine how soft-GIS and choice-based conjoint methods can help accurately measure the perceived importance of quality of life indicators and incorporate them into the impact assessment process by allowing researchers to map out the quality of the living environment and forecast land-use scenarios;
- Explore ways of analysing "soft knowledge" together with other geo-demographic data to understand land use in peri-urban areas. For example, to examine how people's preferences and perceived affordances relating to land use is affected by membership to different groups and segments in the population, such as older people and groups of immigrants;
- Discuss how these participatory methods can help build a bridge for more effective collaboration between researchers, stakeholders and planners and to provide a more holistic picture of the quality of the environment.

3.5 Example V: Mapping Survey Data into Agents' Behavioural Rules for ABMs: Motivation and Challenges

Modeling land use change inevitably involves modeling of an individual behaviour of land users in addition to modeling of spatial environment. The processes in the latter usually follow some physical laws. However, it is less straightforward for a modeller how to describe the process of human decision making (Berger and Schreinemachers 2006; Brown and Robinson 2006; Stites 2006). As it is observed by ABM-modellers, it is relatively easy to model the mechanical part of an ABM such as spatial environment, because their dynamics is described by a set of straightforward deterministic rules (with some uncertainty intervals sometimes).

In contrast, for human-beings it is not possible to say exactly how they (i.e., we) make decisions. Theoretically, land use behaviour is well formalized in economics. Farmers' (von Thünen 1826 (reprinted in 1966)), households' (Alonso 1964; Strazsheim 1987) and firms' (Fujita and Thisse 2002) decision making with respect to land is fully based on the assumption of a rational maximization, equilibrium, and representative behaviour.

In reality, people are boundedly rational, their behaviour is often unrepresentative, they choose different strategies in the same situation, their decisions are biased by previous experiences and emotions, and people sometimes make irrational decisions. All these observed characteristics of human behaviour make it difficult to use stylized theories of human decision making at the micro level.

Thus, how people make decisions (e.g. about land use) remains a black box for a modeller. The only way to open it a little bit is to analyze real world micro level *data*. These data could probably be obtained either by observing a land-user decision-making in the controlled environment (for example in the setting of a role-playing game (Barreteau et al.

2001; Bousquet et al. 2005), from interviews with stakeholders and during participatory workshops with them, or by gathering data in the form of surveys (Brown and Robinson 2006; Fernandez et al. 2005).

At the workshop we would like to discuss challenges and open questions with respect to using survey data for feeding ABMs. No doubts that surveys provide very valuable data about micro level decision making. However, the following issues might arise here:

1. It looks more straightforward to parameterize the initial agent population in the ABM with empirical data than to generate agent behavioural rules. If one uses statistical data for the latter then he can actually run into a problem of transferring aggregated data into agent rules, which are supposed to be heterogeneous (e.g., ask an agent to reproduce if she reaches a certain age) (Axtell et al. 2002). What is the best way to interpret survey data in term of agents' rules? Should we better categorize survey subjects into several groups and then clone them in the ABM? Or assign each agent a specific rule adopted from the survey with some probability (which might correlate with the percent of the sample, which has chosen it)?
2. Very often surveys produce qualitative measures. How can those measures be translated into formal agents' rules accurately? For example, how can we accurately transfer a likert scale into agents' behavioural rules? May the middle point choice of one agent on a likert scale mean the same as the low point choice for another agent? Can we use a respondent's stated choice as realistic? Some studies show that the intention to do some action stated in the survey does not always imply that people will actually do that. Is this fact worth considering in transferring survey data into ABM rules? If so, then how should we account for it?
3. With respect to climate change influencing human demands for housing: we are particularly interested in how perception of risk of flooding affects households' location choices in and outside flood-prone areas. How is it incorporated in the individual choice for housing and consequently in the aggregate demand for housing in the urban zone? Intensified climate change not only implies more risks and actually reduces supply of land available for development. Increasing storm and flood frequency, for example, may make people more aware of these risks and may change their location preferences.
4. Surveys produce answers on a typical situation (under investigation) in the field of study and a set of socio-economic characteristics of each respondent. Would it mean that the whole population can be divided into groups on the basis of their socio-economic characteristics (e.g. age, income) and each of the groups will be associated with a particular type of behavior?
5. Micro calibration and macro validation: comparing results of simulations runs (when micro behavior is calibrated with survey data) with statistical macro measures.

We are interested in discussing these questions in general and in application to our case study in the Netherlands in particular. We are working on an ABM simulating urban dynamics in a coastal city (Filatova et al. 2007; Filatova and van der Veen 2007). In parallel we are performing a survey on risk of flood perception and location choices of households in the province of Zeeland in the Netherlands. We hope that this workshop will provide possibilities for sharing our experiences and a ground for a brainstorming session.

3.6 Example VI: Actors as factors for land use change: Effects of demographic change on land use across Europe

Intuitively, it seems to be evident that there is a close relationship between human society and natural environment. People change their environment through intensive utilisation: for agriculture, for living, for transport, for leisure activities, for resources, etc. Thus, it appears clear that there must be a close link between demographic and land use development. However, quantifying and modelling this relationship based on statistically significant correlations is a challenging task. Even if we focus the research question on one single land use type, namely settlement area, where a very close and direct connection to population as driving force is assumed, it is not easy to explain the relationship statistically by a formula

fitting for different regions across Europe. In some European regions the development of settlement patterns and population numbers even diverges as on the one hand population stagnates or shrinks and on the other hand settlement areas grow. This phenomenon has socio-economic reasons. Empirical findings from Austrian and European studies proved that nowadays the growth of settlement area is not solely caused by growing population numbers but particularly in prosperous and urbanised regions by increasing demand for settlement area per person due to higher living standard and income (Loibl and Tötzer 2003, Tötzer et al. 2007). Thus, for modelling human actors as driving force for land use change, demand-related factors like settlement area/consumption per head, settlement density, household size, etc. have to be considered beside mere population numbers.

In this paper we will present first findings from our research within PLUREL, an Integrated Project funded within the 6th framework program of the EC. PLUREL deals with the relationships between urban, peri-urban and rural land use and aims to develop strategies for a sustainable development of these interlinked rural-urban regions (RURs). A key product will be the sustainability impact assessment tool for urban-rural linkages. PLUREL applies the DPSIR-concept (Driver-Pressure-State-Impact-Response) on different European scales: from NUTS0 to LAU2 in six European case study regions.

The research presented in this paper explores population as driving force for land use change focussing on settlement patterns and dynamics. We work on two different scales: on European scale and on micro-scale in six case study regions (among them e.g. Warsaw, Leipzig, Manchester). In this project we will examine if the findings from preceding research projects as well as from the PLUREL case studies could be confirmed on European scale, so that European-wide rules can be derived for certain region types (e.g. urban-monocentric, dispersed peri-urban, rural, urban-polycentric) across Europe. We will apply statistical analysis methods and carry out analyses on European NUTS3- level as well as on local scale for the case study regions.

Research on micro-scale allows in-depth analyses based on detailed data (e.g. LAU2) and gives insights into region-specific interdependencies between society, economy and environment. However, considering the macro-scale as well is particularly important for stakeholders and politicians in the EU. Thus, up-scaling the case study specific rules to the European scale will be an essential task within the project. With the help of pre-defined region types, results which correspond on local and on European scale will be proved and up scaled. Due to data constraints proxy-data will have to be used on European level.

The research will result in statistically based functions reflecting relationships between population and settlement areas in different region-types across Europe. These functions will be considered in many following work packages within the PLUREL project and will help to model the Sustainable Impact Assessment Tool for Rural-urban Relations – one of the key products of the PLUREL project.

3.7 Example VII: Land use policy impacts in the rural-urban region: a modelling framework

Because land is a fundamental input of anthropogenic (production, residence, recreation etc.) as well as non-anthropogenic (ecosystem functioning, refuge, habitat etc.) activities, modelling such impacts require a conceptual framework multi-sectoral and multi-zonal in scope. This framework should account for flows of people, commodities and services from one sector to the other and from one zone to the other. Because of the very nature of the problem at hand, an urban and rural context needs also be given to the conceptual framework. In this paper we propose an extension to the PACE model inspired by the modelling approach of metropolitan input-output models (see Jun, 1999 and 2005).

PACE is comparative-static multi-region, multi-sector CGE model. Primary factors of a region include labor, capital, and fossil-fuel resources. While in PACE regions are usually countries and cities in metropolitan input-output models, in our conceptual framework we refer to the following three regions: urban, peri-urban and rural. Each region can be divided in sub-region depending on the spatial disaggregation of available data. Capital is divided into two categories: manufactured capital (including infrastructure) and non-manufactured

capital (including forests, mineral deposits, fisheries, biodiversity etc.). We also add land as a further resource representing land that is not used for capital. In PACE the energy sector is modelled explicitly to better account for impacts of changes in climate policy. In our framework we add the real estate sector and show how the decision to develop land can be modelled using a real option approach.

Nested constant elasticity of substitution (CES) cost functions are employed to specify the substitution possibilities in domestic production between capital, labour, land, energy and non-energy, intermediate inputs, i.e. material. Final demand in each region is determined by a representative agent, who maximizes utility subject to a budget constraint.

In PACE, furthermore, labour and capital are intersectorally mobile within a region but cannot move between regions. A discussion on data needs to make this conceptual framework operative follows.

3.8 Example VIII: Quantification of land use changes, modelling of processes and impacts of shrinkage and the question of scale

Following the purpose of understanding and analysis of interactions between society and landscape in urban areas this paper tries to give a methodological and practical insight into the assessment of socio-environmental impacts under conditions of uneven spatial development. In that context the paper draws special attention to the quantification and modelling of land use changes and shrinkage processes, in particular, based on empirical ascertainable and communal data and empirical research techniques. A second focus highlights a conceptual approach for the assessment of functionality of urban green spaces against the concepts of quality of Life (QoL) and associated Ecosystem Services (ESS) in urban areas dealing with both quantitative modelling and empirical qualitative research techniques. In a third focus the question of choosing the right scale for assessment and modelling of land-use changes will be dealt with as spatial dynamics and their socio-environmental effects seem to be very variable on a smaller scale and extreme within shrinking cities.

Interactions of society and landscape

Friction is the prerequisite for dynamism and interaction. Heterogeneous and uneven spatial development are that friction between landscape and society in shrinking cities. Regarding a surplus of open spaces resulting from enormous vacancies and demolition on the one side and a decreasing number of inhabitants on the other both domains appear to be under pressure. How do these citizens deal with their changed – sometimes blurred and unstructured- living surroundings, what kind of socio-environmental impacts derive from that dynamism and come to pass in post-shrinkage landscapes? In that context I will introduce into a recent study focussing on the quantification and assessment of socio-environmental impacts according to various demolition scenarios in the city of Leipzig, Eastern-Germany (Schetke & Haase 2007, in press). Main outcome of this study is a set of indicators focussing on both structural-ecological assessment (e.g. LSM) as well as on modified social spaces deriving from modified accessibilities of social infrastructure (green and technical ones). The emphasis of this set of indicators is the quantification and socio-environmental assessment of land use changes and shrinkage processes using empirical ascertainable (e.g. communal) quantitative and qualitative data.

Attractiveness of urban life. Concepts for assessment

After the introduction into a model of the assessment of socio-environmental impacts I want to expand these findings by going a bit deeper into the analysis of the attractiveness and quality of urban live by presenting a conceptual approach (see Schetke, Haase & Breuste 2008, in prep.) focussing on the functionality of green spaces accoording to the concepts of quality of life (see Burgess et al. 1988; Givoni 1991; Kawka & Sturm 2006) and associated ecosystem services (Bolund & Hunhammar 1999; Constanza 1997 et al.; de Groot et al. 2002).

Following the presentation of an indicator set quantifying and assessing socio-environmental impacts in shrinking this conceptual approach widens its view by focussing

more clearly on the interactions between society and landscape in urban areas. To achieve this not only the area of research has been expanded by adjusting the model towards the analysis of the functionality of urban green under conditions of uneven – that means shrinkage and growth side by side- spatial development. But also the compilation of data implemented into the model has to be enlarged using not only administrative quantifiable land-use and social data but qualitative empirical data highlighting the perceptive side of land- use changes by clearly focussing on social-spatial interdependencies, as well.

Beside the scientific claim in analysing the socio-spatial interactions in urban landscapes the focus on specifically the functionality of urban green is founded on the fact that urban green and open spaces are among the driving factors of quality of life and of that attractiveness of urban life planners are in charge to provide.

Keeping in mind that this affects both planners and scientists the gap between the two domains can be bridged by the conscious implementation of valuable communal and empirical ascertainable data into such a model. Also planners are more and more in charge to estimate spatial and social dynamics in order to remain able to steer a sustainable spatial development and to promote the attractiveness of urban life. In that context a conceptual model not only dedicated to quantify land-use changes but also able to analyse and to depict socio-spatial interactions contribute to both side's gain of knowledge and progress.

Spatial and land use effects caused by demographic change. Choosing the right scale

In shrinking cites socio-environmental dynamics can be very heterogenous and diverse amongst different neighbourhoods and urban structure types. Both ecological and social affairs - each type of a single urban ecosystem, each social group of citizens- are affected in a different way with sometimes severe and sometime almost unnoticeable impact. The more, the process of shrinkage and the often accompanying demolition occurs in different spatial and temporal ways which make a general large-scale assessment and modeling (e.g. on city or regional scale) of land use changes and shrinkage processes difficult and redundant in meaning. The question of scale is long discussed in terms of land use assessment. Especially in terms of quantifying and modelling of both presented socio-environmental impacts of shrinkage and a more in-depth analysis of socio-spatial interactions and the functionality (see QoL and ESS) of urban green it is still a sensitive matter to discuss and to keep in mind.

3.9 Example IX: Integrated Modelling of Smallholder Land-use Decisions on the Availability of Rattan

The margins of national parks belong officially to the protected area of the parks. However, they are used for the livelihood of smallholders in the park's vicinity. Conflicts and continued poverty around protected areas suggest that existing approaches to conservation lack understanding of links between maintenance of diversity and human well-being [Bawa et al., 2004]. Regularly, the livelihood of the rural poor includes agricultural land-use. Large areas are converted by smallholder farmers, despite the fact that tropical rain forests play a major role in the provision of ecosystem services [Achard et al., 2002]. The interaction between human and natural systems at the rain forest margin play a key role in preserving and stabilising forested ecosystems [Maertens et al., 2006].

Ecosystem stability is a multi-dimensional concept requiring multi-disciplinary analysis because it concerns economic, ecological and social issues. This way we hope to identify more sustainable development options, such as mitigation of climate change effects and biodiversity loss, poverty alleviation and economic development for rural areas [Balmford et al., 2002].

One tool to address the livelihoods of the smallholders as well as impacts on ecosystem functioning are integrated assessment models. Mallawaarachchi and Quiggin [2001] e.g., present a model for the integrated assessment of sugar cane production in Queensland/ Australia. We will extend such an approach by a spatially explicit agent-based-modelling component. Recently the application of spatially explicit agent-based models have become a focus in the research on land-use and cover change [Bousquet and Le Page, 2004]. This modelling concept – originally from the field of problem solving in (distributed) artificial intelligence - provides an excellent frame to simulate the problem at hand. With the SIEHL-

framework (Spatially and Institutionally Explit Household decision-making and Land use model) we apply this rationale to the margins of tropical rainforests, first to Indonesia (Lore Lindu National Park); *In-SIEHL* and later to Ecuador (Podocarpus National Park; *Ec-SIEHL*).

The main component of this framework is an Agent-Based Model, which allows us to differentiate between the perception of individual agents and system response modelled with 'full' information. The land-use decision-making options will be the result of an economical optimized result of the production opportunities. We first apply this model to rattan collectors in Lore Lindu National Park in Central Sulawesi.

Multi-agent-model. The core element of the IN-SIEHL-framework is a Multi-Agent System (*M*AS). According to the Multi-Agent-Paradigm, agents are software representations ". . . that are situated in some environment and capable of autonomous actions in this environment in order to meet their design objectives" [Wooldridge, 2001]. MAS allow us to simulate the land-use decisions of smallholders on an individual level. Particularly, we can differentiate between how an individual agent perceives its environment ('*Merkwelt'*) and the 'full-information' system response ('*Wirkwelt'*). A common framework for the agent's rationale is the *Belief-Desire-Intention*-Architecture.

In our case, the framework can be explained as follows: Each agent has a certain *Desire* (*profitmaximisation*), which he tries to pursue. An ('Wirkwelt') stimulus triggers the revision of the *Belief* ('Merkwelt'). According to its beliefs, a set of options is generated, from which the agent choses one option which will – according to his belief – fit best to fulfill its desire (*Intention*).

For rattan collection this can be exemplified by the problem to actually design an optimal rattan harvesting strategy that maximises long-term rattan benefits for a poor household. The smallholders know about putative rattan sites (*Merkwelt*) and also roughly to which extent harvesting has taken place there. However, actual rattan harvesting decisions are not made from an omniscient position according to analyst knowledge on potentially optimal and sustainable harvesting strategies as obtained from an analysis of all *Wirkwelt* data available in the model.

Agents. In the Central Sulawesi In-SIEHL implementation. the agents in the model are aggregated entities, because they do not represent individual humans, but smallholder households as a decision-maker. Data of the availability of the capital were collected in social science surveys and fed into a database (Data from the STORMA-Project and Schwarze [2004]). These sources supply data on capital access and restrictions to economic activity of different classes of smallholder households. Off-farm work will be modelled as rattan extraction. Particularly for remote villages such as Au and Moa this is highly realistic. This decision-process will be optimised as a result of an analysis (Lingo) of single household production opportunities). Every agent has a plot of land he uses exclusively. The main activity of an agent is to decide on and implement the land use of their plot.

Bio-physical Environment. The original landscape data from the STORMA-project will be used. Based on LandSat-images on a 30 £ 30 m grid, we use the data on land-cover and altitude. Agents are located in the villages and rattan will be distributed according to a height gradient. Furthermore previous harvesting activities will be taken into account. With an GIS-Analysis and expert knowledge, the most probable entry points for rattan-harvesting will be located and supposed trails (e.g. a gradient with the smallest slope) used for modelling purposes. The abundance data of *Calamus zollingeri*, is used for initialisation Siebert [2004]. Since only two initial factors are known, which determine the abundance of rattan (distance to town and altitude), those will be used to simulate the distribution of rattan. After initialisation, the agent will decide – for example every "year" – on its land-use, which will be "implemented" subsequently. This will be the yield that corresponds to the harvesting rates that can be estimated using the Siebert [2004] data.

Model structure. The technical Framework structure is presented in figure 2. From a central Graphical-Users Interface (*JAVA-Application*), different background scenarios (price, climate variation) and evaluation components can be selected, and the simulation of the MAS is triggered. The MAS is modelled in Repast (http://repast.sourceforge.net), a

toolkit with various built-in methods for data analysis and extension possibilities such as database connectivity or GIS-extension.

The economic and ecological data needed for simulation are loaded from an Entity-Relationshipdatabase. In this database, various data from different sources are stored, i.e. Land-Sat images on a 30£ 30m basis, information on vegetation cover and detailed household survey data including geo-referenced resource access. Depending on the research question, different 'impact-modules' can be included into the framework.

The ecological module 'implements' the land-use decision in '*Wirkwelt*',. . . its ecological consequences; e.g. on the hydrological cycle or on rattan stocks. The economic impact is a consequence of the implemented land-use decision, too. In this case exogenous factors (climate, . . .) influence crop yield, which might be very different from what the agent anticipated ('*Merkwelt*'). The profits are added to the financial capital. The results of the impacts are again fed back into the database and can be used for further analysis.

Calibration and validation. Data from work of Siebert [2000] and Siebert [2004] will be taken to estimate the spatial distribution of the rattan. However, both publications only offer rattan data in an aggregated form. In order to get a more realistic spatial distribution we apply the following steps. First, we calculate a weighted path raster. These paths are those, which are travelled most likely by the rattan collectors. We assign the weights, which stand for the movement costs. The values of the weights are chosen in accordance with expert knowledge. This means that travelling along rivers has the lowest movement cost and travelling within the forest has the highest costs. The cells of this raster contain the accumulated cost values. In the next step we combine the information of the rattan on the height gradient of Siebert [2000] with this raster.

REFERENCES (cited and background)

Achard, F., Hugh D.E., Stibig, H.J., Mayaux, P., Gallego, J., Richards, T., Malingreau, J.-P. (2002) Determination of Deforestation Rates of the World's Humid Tropical Forests, *Science, 297* (5583), 999–1002.

Alonso, W. (1964), *Location and Land Use* (Cambridge, MA: Harvard University Press.).

Axtell, Robert, et al. (2002), 'Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley', *Proceedings of the National Academy of Sciences (PNAS)*, 99 (3), 7275–79.

Balmford, A., Bruner, A., Cooper, P., Costanza, R., Farber, S., Green, R.E., Jenkins, M., Jefferiss, P., Jessamy, V., Madden, J., Munro, K., Myers, N., Naeem, S., Paavola, J., Rayment, M., Rosendo, S., Roughgarden, J., Trumper, K., Turner, R.K. (2002) Economic Reasons for Conserving Wild Nature, *Science 297* (5583), 950–953.

Barreteau, O., Bousquet, F., and Attonaty, J.-M. (2001), 'Role playing game for opening the black box of multi-agent systems: method and lessons of its application to Senegal River Valley irrigated systems.' *Journal of Artificial Societies and Social Simulation,* 4 (2), 12.

Beck, U., and E. Beck-Gernsheim. 2002. *Individualisation. Institutionalized Individualism and its Social and Political Consequences*. SAGE.

Bell, S., Alves, S., Oliveira, E.S., 2007. Response functions for social issues. Typescript, PLUREL-GA 2007, Leipzig.

Ben-Akiva, M.E., Lerman, S.R. (1985) Discrete choise analysis: Theory and application to travel demand. MIT press, Cambridge. MA.

Berger, T., Schreinemachers, P. (2006), Creating Agents and Landscapes for Multiagent Systems from Random Samples, *Ecology and Society*, 11 (2), art. 19.

Bharwani, S., Understanding complex behaviour and decision making using ethnographic knowledge elicitation games (KnETs). *Social Science Computer Review* 24(1), 78-105, 2006.

Bolund, P., Hunhammar, S. 1999: Ecosystem Services in urban areas. In: Ecological Economics 29, 293-301.

Bösch-Supan, A., Brugiavini, A., Jürgs, H., Mackenbach, J., Siegrist, J., Weber, G., 2005. *Health, Ageing and Retirement in Europe*. MEA Mannheim.

Bousquet, F., Le Page, C. (2004) Multi-agent simulations and ecosystem management: a review, *Ecological Modelling 176* (3-4), 313–332.

Bousquet, F., Trebuil, G., and Hardy, B. (2005), *Companion Modeling and Multi-Agent Systems for Integrated Natural Resource Management in Asia* (Los Banos, Philippines: International Rice Research INstitute).

Brown, D.G., Robinson, D.T. (2006), 'Effects of Heterogeneity in Residential Preferences on an Agent-Based Model of Urban Sprawl', *Ecology and Society*, 11 (1), art. 46.

Burgess, J. et al. 1988: People, Parks and Urban Green: A Study of Popular Meanings and Values for Open Spaces in the City. In: Urban Studies 25, 455-473.

Buzar, S., Odgen, P., Hall, R., 2005. Households matter: the quiet demography of urban transformation. Progress of Human Geography 29(4), 413-436.

Buzar, Stephan, Ogden, Philip E., Hall, Ray, Haase, Annegret, Kabisch, Sigrun, and Annett Steinführer. 2007. "Splintering Urban Populations: Emergent Landscapes of Reurbanisation in four European Cities." Urban Studies 44(4): 651-677.

Clarke, K.C., Hoppen, S., Gaydos, L. (1997) A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area, *Environment and Planning B: Planning and Design* 24, pp. 247-261.

Cloet, R., 2003. Population Changes 1950-2050 in Europe and North America. Population Statstics.doc 3-03, 1-11.

Constanza, R. et al. 1997: The value of the world's ecosystem services and natural capital. Nature, 387, 253-260.

Council of Europe, 2005. Recent demographic developments in Europe, Council of Europe publishing, Strasbourg.

De Groot, R. et al. 2002: A typology for the classification , description and valuation of ecosystem functions, goods and services. In: Ecological Economics 41, 393-408.

Deutsch, L., Folke, C., Skånberg, K., 2003. The Critical Natural Capital of Ecosystem Performance as Insurance for Human Well-Being. Ecological Economics 44, 205-217.

Ekins, P., Folke, C., De Groot, R., 2003. Identifying Critical Natural Capital. Ecological Economics 44, 159-163.

Fernandez, Luis E , et al. (2005), 'Characterizing location preferences in an exurban population: implications for agent-based modeling', *Environment and Planning B: Planning and Design*, 32, 799-820.

Filatova, Tatiana and van der Veen, Anne (2007), 'Scales in coastal land use: policy and individual decision-making (an economic perspective)', in Caroline van Bers, Daniel Petry, and Claudia Pahl-Wostl (eds.), *Issues in Global Water System Research. Global Assessments: Bridging Scales and Linking to Policy* (#2; Bonn: The Global Water System Project), 61-68.

Filatova, Tatiana, Parker, Dawn C., and van der Veen, Anne (2007), 'Agent-Based Land Markets: Heterogeneous Agents, Land Prices and Urban Land Use Change', in Frederic Amblard (ed.), *Proceedings of the 4th Conference of the European Social Simulation Association (ESSA'07)* (Toulouse, France), 263-77.

Fujita, M. and Thisse, J-F (2002), *Economics of agglomeration. Cities, industrial location and regional growth* (Cambridge University Press).

Givoni, B. 1991: Impact of planted areas on urban environmental quality: a review. In: Atmospheric Environment Vol. 25B No.3, pp. 289-299.

Gober, P. 1990. The Urban Demographic Landscape: A Geographic Perspective. In Housing Demography. Linking Demographic Structure and Housing Markets, ed. D. Myers, 232-248. Madison & London: University of Wisconsin Press.

Haase, A., Kabisch, S., Steinführer, A., 2005. Reurbanisation of Inner-City Areas in European Cities. In: Sagan, I, Smith, D M (Eds): Society, economy, environment – towards the sustainable city, Gdańsk, Poznań, 75-91.

Haase, A., Nuissl, H., Rink, D., submitted, New pathways of urban sprawl and reurbanisation. Impacts of demographic change and population decrease on current urban change. *Urban Studies*.

Haase, D., Seppelt, R., submitted. Applying social science concepts: modelling and simulating scenarios for residential mobility. *Environmental Modelling and Software*.

Horowitz, J.L., Koppelman, F.S., Lerman, S.R. (1986) Self-instructing course in disaggregate mode choise modeling. Technology Sharing Program, US Department of Transportation, Washington DC.

Kaa, D. van de, 1987. Europe's Second Demographic Transition. Population Bulletin 42, 1-57.

Kaa, D. van de, 2004. Is the Second Demographic Transition a useful research concept. Vienna Yearbook of Population Research 2004, pp. 4-10.

Kabisch, S., Haase, A., Haase, D., 2006. Beyond growth – urban development in shrinking cities as a challenge for modeling approaches. In: Voinov, A., Jakeman, A., Rizzoli, A. (eds). *Proceedings of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software".* International Environmental Modelling and Software Society, Burlington, USA, July 2006.

Kawka, R., Sturm, I. 2006: Objektive regionale Lebensqualität und subjektives Wohlbefinden- Was macht Bürgerinnen und Bürger zufrieden? In: Informationen zur Raumentwicklung Heft 6/7 2006.

Klosterman, R.E. (1999) The What if? Collaborative Planning Support System. *Environment and Planning, B: Planning and Design* 26, pp. 393-408.

Landis J., Zhang, M. (1998a) The second generation of the California urban futures model. Part 2: Specification and calibration results of the land-use change submodel. *Environment and Planning B: Planning and Design* 25, pp. 795-824.

Landis J., Zhang, M. (1998b) The second generation of the California urban futures model. Part 1: Model logic and theory, *Environment and Planning A*, 30, pp. 657-666.

Landis, J. (1994) The California Urban Futures Model: a new generation of metropolitan simulation models, *Environment and Planning B: Planning and Design*, 21, pp. 399-420.

Lesthaeghe, R., and J.R. Surkyn. 2002. "New Forms of Household Formation in Central and Eastern Europe. Are they related to Newly Emerging Value Orientations?" UNECE Economic Survey of Europe, 2002-1, Chapter 6, 197-216. Geneva: United Nations Commission for Europe.

Loibl, W., and T. Tötzer, Modeling Growth and Densification Processes in Sub-urban Regions – Simulation of Landscape Transition with Spatial Agents. *Environmental Modelling and Software. Volume 18, Issue 6*, pp. 485-593, 2003.

Lutz, W., 2001. The end of World Population Growth. Nature 412, 543-545.

Maertens, M., Zeller, M., Birner, R. (2006) Sustainable agricultural intensification in forest frontier areas, *Agricultural Economics 34*(2), 197–206.

Mallawaarachchi, T., Quiggin, J. (2001) Modelling socially optimal land allocations for sugar cane growing in North Queensland: a linked mathematical programming and choice modelling study, *The Australian Journal of Agricultural and Resource Economics 45* (3), 383–409.

McIntosh, B.S., C. Giupponi, A. Voinov, C. Smith, K.B. Matthews, M. Monticino, M.J. Kolkman, N. Crossman, M. van Ittersum, D. Haase, A. Haase, J. Mysiak, J.C.J. Groot, S. Sieber, P. Verweij, N. Quinn, P. Waeger, N. Gaber, D. Hepting, H. Scholten, A. Sulis, H. van Delden, E. Gaddis, H. Assaf, 2008. Bridging the gaps between design and use: developing tools to support environmental management and policy. In: Jakeman, A., Voinov, A., Rizzoli, A., Chen (eds.) *State of the art and Futures in Environmental Modelling and Software*. Elsevier, in press.

Miller, E.J., Hunt, J.D., Abraham, J.E., Salvini, P.A. (2004) Microsimulating urban systems, Computers, *Environment and Urban Systems*, 28, pp. 9-44.

Ogden, P. E., Hall, R., 2000. Households, Reurbanisation and the Rise of Living Alone in the Principal French Cities, 1975–90, Urban Studies 37(2), 367-390.

Rink, D., Haase, A., Nuissl, H. (2008) New Pathways for urban sprawl and Reurbanisation. Presentation at the IALUC conference: Impact Assessment of Land Use Changes, Berlin, April 6-9, 2008.

Schetke, S., Haase, D. 2007: Multicriteria assessment of socio-environmental aspects in shrinking cities. Experiences from eastern Germany. In: Environmental Impact Assessment Reviews (in press).

Schetke, S., Haase, D., Breuste, J. 2008: Recreational land use and green space functionality in uneven spatial urban development (in prep.).

Siebert, S. F. (2000) Survival and growth of rattan intercropped with coffee and cacao in the agroforests of Indonesia, *Agroforestry Systems 50*, 95–102.

Siebert, S. (2004) Demographic Effects of Collecting Rattan Cane and Their Implications for Sustainable Harvesting, *Conservation Biology 18*(2), 424–431.

Silva, E.A., Clarke, K.C. (2002) Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. *Computers, Environment and Urban Systems*, 26, pp. 525-552.

Stites, Janet (2006), 'Agent-based modeling. The future of financial markets', *Santa Fe Bulletin*, Winter 2006, 31-35.

Strazsheim, M. (1987), 'The Theory of Urban Residential Location', in E.S. Mills (ed.), *Handbook of Regional and Urban Economics* (Volume II: Elsevier Science Publishers B.V), 717-57.

Tötzer, T, M. Köstl, K. Steinnocher, Scenarios of land use change in Europe based on socio-economic and demographic driving factors. *Proceedings of the 12th International Conference on Urban Planning & Regional Development in the Information Society*, 2007.

von Thünen, Johann Heinrich (1826 (reprinted in 1966)), *Isolated State: An English Edition of Der Isolierte Staat* (Pergamon Press).

Waddell, P. (2002) Urbansim: Modeling Urban Development for Land Use, Transportation and Environmental Planning, *Journal of the American Planning Association*, 68(3), pp. 297-314.

White, R., Engelen, G., Uljee, I. (1997) The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics, *Environment and Planning B: Planning and Design*, 24, pp. 323-343.

Wu, F. (1998) Simulating urban encroachment on rural land with fuzzy-logic-controlled cellular automata in a geographical information system, *Journal of Environmental Management*, 53, pp. 293-308.

Wu, F., Webster, C.J. (1998) Simulation of land development through the integration of cellular automata and multicriteria evaluation, *Environment and Planning B: Planning and Design*, 25, pp. 103-126.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Simulation and scenarios of a social science concept on urban shrinkage

A.Haase[a], D. Haase[b], S. Kabisch[a]

[a]UFZ – Centre for Environmental Research, Dept. of Urban and Environmental Sociology, Permoserstr. 15, D-04318 Leipzig, Germany, (annegret.haase@ufz.de)
[b]UFZ – Centre for Environmental Research, Dept. of Computational Landscape Ecology, Permoserstr. 15, D-04318 Leipzig, Germany

Keywords: *social science concepts; urban shrinkage; causalities; simulation model.*

In our paper the hypothesis is tested whether current demographic decline, de-economisation and structural crisis are bringing about a new evolving type of urban development. In doing so, we present a novel social science concept of urban shrinkage that goes beyond the focussed population-related view of shrinkage (Kabisch et al., 2006). It rather comprises the features of non-growth, on-going sprawl and upcoming reurbanisation alike, or if reurbanisation represents, on the other hand, merely a short interplay in a long-term period of desurbanisation under the condition of demographic change and decrease in population: the core city is gradually shrinking because of demographic decline.

Nevertheless, in the course of reurbanisation, the core city will be able to remain, or become again, compared to the outskirts and surroundings, as we can already see in cities that follow a consequent policy of strengthening the inner city. The less abrupt demography change the better the possibilities of cities to force reurbanisation and to support the development of their central areas are. Admittedly, reurbanisation takes place in a selective way and isles of reurbanisation arise within a context of emptiness and shrinking (Haase et al., submitted).

Using a system dynamics approach we apply this social science concept in a simulation model which uses variables emerging from this concept as descriptors for typical shrinkage related population, household, residential and other land use processes and pattern (McIntosh et al., in press). For the model implementation we had to “translate” non-numerical and only partially quantifiable social science knowledge into model variables such as urban stocks, flows and trends.

In our presentation we show the simulation of different urban development scenarios assuming a range of population and household formation changes, a further increase of single and single-parent households, effects of ageing and related altering demands and preferences on residential space and places to live. Accordingly, we simulate pattern of residential vacancy and potential sites of demolition (Haase & Seppelt, submitted). Based on these simulations we can show under which demographic conditions urban

growth and shrinkage simultaneously occur or even dominate the development of a city/urban region.

REFERENCES

Kabisch, S., Haase, A., Haase, D., 2006. Beyond growth – urban development in shrinking cities as a challenge for modeling approaches. In: Voinov, A., Jakeman, A., Rizzoli, A. (eds). *Proceedings of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software"*. International Environmental Modelling and Software Society, Burlington, USA, July 2006.

Haase, A., Nuissl, H., Rink, D., submitted, New pathways of urban sprawl and reurbanisation. Impacts of demographic change and population decrease on current urban change. *Urban Studies*.

Haase, D., Seppelt, R., submitted. Applying social science concepts: modelling and simulating scenarios for residential mobility. *Environmental Modelling and Software*.

McIntosh, B.S., C. Giupponi, A. Voinov, C. Smith, K.B. Matthews, M. Monticino, M.J. Kolkman, N. Crossman, M. van Ittersum, D. Haase, A. Haase, J. Mysiak, J.C.J. Groot, S. Sieber, P. Verweij, N. Quinn, P. Waeger, N. Gaber, D. Hepting, H. Scholten, A. Sulis, H. van Delden, E. Gaddis, H. Assaf, 2008. Bridging the gaps between design and use: developing tools to support environmental management and policy. In: Jakeman, A., Voinov, A., Rizzoli, A., Chen (eds.) *State of the art and Futures in Environmental Modelling and Software*. Elsevier, in press.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Mapping out the social impacts of land use: Soft-GIS and Choice-Based Conjoint Analysis as integrative methods

A.Zuin[a]

Edinburgh College of Art, OPENspace,UK (azuin@eca.ac.uk)

Keywords: *social impacts, Soft-GIS, choice-based Conjoint Analysis.*

People's experiences of place are affected by planning strategies as much as community-based behaviours and emotions can affect planning and development. Residents' experiences of their environment are geographically located, since people can attach their experiences, memories, and feelings to specific locations. By allowing participants to locate these experiences, researchers can not only identify their specific affordances but also analyse and visualize the information for research and planning practices. This example aims to:

- Examine how soft-GIS and choice-based conjoint methods can help accurately measure the perceived importance of quality of life indicators and incorporate them into the impact assessment process by allowing researchers to map out the quality of the living environment and forecast land-use scenarios;
- Explore ways of analysing "soft knowledge" together with other geo-demographic data to understand land use in peri-urban areas. For example, to examine how people's preferences and perceived affordances relating to land use is affected by membership to different groups and segments in the population, such as older people and groups of immigrants;
- Discuss how these participatory methods can help build a bridge for more effective collaboration between researchers, stakeholders and planners and to provide a more holistic picture of the quality of the environment.

Mapping Survey Data into Agents' Behavioral Rules for ABMs: Motivation and Challenges

Tatiana Filatova[a], Anne van der Veen[ab]
a - Department of Water Engineering and Management,
University of Twente, The Netherlands, T.Filatova@ctw.utwente.nl
b - International Institute for Geo-Information Science
and Earth Observation, The Netherlands, Veen@itc.nl

Keywords: behavioral rules of agents; survey data

Modeling land use change inevitably involves modeling of an individual behavior of land users in addition to modeling of spatial environment. The processes in the latter usually follow some physical laws. However, it is less straightforward for a modeler how to describe the process of human decision making (Berger and Schreinemachers 2006; Brown and Robinson 2006; Stites 2006). As it is observed by ABM-modelers, it is relatively easy to model the mechanical part of an ABM such as spatial environment, because their dynamics is described by a set of straightforward deterministic rules (with some uncertainty intervals sometimes). In contrast, for human-beings it is not possible to say exactly how they (i.e., we) make decisions. Theoretically, land use behavior is well formalized in economics. Farmers' (von Thünen 1826 (reprinted in 1966)), households' (Alonso 1964; Strazsheim 1987) and firms' (Fujita and Thisse 2002) decision making with respect to land is fully based on the assumption of a rational maximization, equilibrium, and representative behavior. In reality people are boundedly rational, their behavior is often unrepresentative, they choose different strategies in the same situation, their decisions are biased by previous experiences and emotions, and people sometimes make irrational decisions. All these observed characteristics of human behavior make it difficult to use stylized theories of human decision making at the micro level. Thus, how people make decisions (e.g. about land use) remains a black box for a modeler. The only way to open it a little bit is to analyze real world micro level *data*. These data could probably be obtained either by observing a land-user decision-making in the controlled environment (for example in the setting of a role-playing game (Barreteau et al. 2001; Bousquet et al. 2005)), from interviews with stakeholders and during participatory workshops with them, or by gathering data in the form of surveys (Brown and Robinson 2006; Fernandez et al. 2005). During this workshop we would like to discuss challenges and open questions with respect to using survey data for feeding ABMs. No doubts that surveys provide very valuable data about micro level decision making. However, the following issues might arise here:

1. It looks more straightforward to parameterize the initial agent population in the ABM with empirical data than to generate agent behavioral rules. If one uses statistical data for the latter then he can actually run into a problem of transferring aggregated data into agent rules, which are supposed to be heterogeneous (e.g., ask an agent to reproduce if she reaches a certain age) (Axtell et al. 2002). What is the best way to interpret survey data in term of agents' rules? Should we better categorize survey subjects into several groups and then clone them in the ABM? Or assign each agent a

specific rule adopted from the survey with some probability (which might correlate with the percent of the sample, which has chosen it)?

2. Very often surveys produce qualitative measures. How can those measures be translated into formal agents' rules accurately? For example, how can we accurately transfer a likert scale into agents' behavioral rules? May the middle point choice of one agent on a likert scale mean the same as the low point choice for another agent? Can we use a respondent's stated choice as realistic? Some studies show that the intention to do some action stated in the survey does not always imply that people will actually do that. Is this fact worth considering in transferring survey data into ABM rules? If so, then how should we account for it?

3. With respect to climate change influencing human demands for housing: we are particularly interested in how perception of risk of flooding affects households' location choices in and outside flood-prone areas. How is it incorporated in the individual choice for housing and consequently in the aggregate demand for housing in the urban zone? Intensified climate change not only implies more risks and actually reduces supply of land available for development. Increasing storm and flood frequency, for example, may make people more aware of these risks and may change their location preferences.

4. Surveys produce answers on a typical situation (under investigation) in the field of study and a set of socio-economic characteristics of each respondent. Would it mean that the whole population can be divided into groups on the basis of their socio-economic characteristics (e.g. age, income) and each of the groups will be associated with a particular type of behavior?

5. Micro calibration and macro validation: comparing results of simulations runs (when micro behavior is calibrated with survey data) with statistical macro measures.

We are interested in discussing these questions in general and in application to our case study in the Netherlands in particular. We are working on an ABM simulating urban dynamics in a coastal city (Filatova et al. 2007; Filatova and van der Veen 2007). In parallel we are performing a survey on risk of flood perception and location choices of households in the province of Zeeland in the Netherlands. We hope that this workshop will provide possibilities for sharing our experiences and a ground for a brainstorming session.

Alonso, W. (1964), *Location and Land Use* (Cambridge, MA: Harvard University Press.).

Axtell, Robert, et al. (2002), 'Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley', *Proceedings of the National Academy of Sciences (PNAS)*, 99 (3), 7275–79.

Barreteau, O., Bousquet, F., and Attonaty, J.-M. (2001), 'Role playing game for opening the black box of multi-agent systems: method and lessons of its application to Senegal River Valley irrigated systems.' *Journal of Artificial Societies and Social Simulation*, 4 (2), 12.

Berger, Thomas and Schreinemachers, Pepijn (2006), 'Creating Agents and Landscapes for Multiagent Systems from Random Samples', *Ecology and Society*, 11 (2), art. 19.

Bousquet, F., Trebuil, G., and Hardy, B. (2005), *Companion Modeling and Multi-Agent Systems for Integrated Natural Resource Management in Asia* (Los Banos, Philippines: International Rice Research INstitute).

Brown, Daniel G. and Robinson, Derek T. (2006), 'Effects of Heterogeneity in Residential Preferences on an Agent-Based Model of Urban Sprawl', *Ecology and Society*, 11 (1), art. 46.

Fernandez, Luis E , et al. (2005), 'Characterizing location preferences in an exurban population: implications for agent-based modeling', *Environment and Planning B: Planning and Design*, 32, 799-820.

Filatova, Tatiana and van der Veen, Anne (2007), 'Scales in coastal land use: policy and individual decision-making (an economic perspective)', in Caroline van Bers, Daniel Petry, and Claudia Pahl-Wostl (eds.), *Issues in Global Water System Research. Global Assessments: Bridging Scales and Linking to Policy* (#2; Bonn: The Global Water System Project), 61-68.

Filatova, Tatiana, Parker, Dawn C., and van der Veen, Anne (2007), 'Agent-Based Land Markets: Heterogeneous Agents, Land Prices and Urban Land Use Change', in Frederic Amblard (ed.), *Proceedings of the 4th Conference of the European Social Simulation Association (ESSA'07)* (Toulouse, France), 263-77.

Fujita, M. and Thisse, J-F (2002), *Economics of agglomeration. Cities, industrial location and regional growth* (Cambridge University Press).

Stites, Janet (2006), 'Agent-based modeling. The future of financial markets', *Santa Fe Bulletin*, Winter 2006, 31-35.

Strazsheim, M. (1987), 'The Theory of Urban Residential Location', in E.S. Mills (ed.), *Handbook of Regional and Urban Economics* (Volume II: Elsevier Science Publishers B.V), 717-57.

von Thünen, Johann Heinrich (1826 (reprinted in 1966)), *Isolated State: An English Edition of Der Isolierte Staat* (Pergamon Press).

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Actors as factors for land use change: Effects of demographic change on land use across Europe

T. Tötzer[a], **W. Loibl**[a]
[a] *Austrian Research Centers – ARC GmbH, systems research, A-1220 Vienna, Austria (tanja.toetzer@arcs.ac.at, wolfgang.loibl@arcs.ac.at)*

***Keywords**:* Demography, Land use change, Statistical functions, PLUREL, DPSIR-concept.

Intuitively, it seems to be evident that there is a close relationship between human society and natural environment. People change their environment through intensive utilisation: for agriculture, for living, for transport, for leisure activities, for resources, etc. Thus, it appears clear that there must be a close link between demographic and land use development. However, quantifying and modelling this relationship based on statistically significant correlations is a challenging task. Even if we focus the research question on one single land use type, namely settlement area, where a very close and direct connection to population as driving force is assumed, it is not easy to explain the relationship statistically by a formula fitting for different regions across Europe. In some European regions the development of settlement patterns and population numbers even diverges as on the one hand population stagnates or shrinks and on the other hand settlement areas grow. This phenomenon has socio-economic reasons. Empirical findings from Austrian and European studies proved that nowadays the growth of settlement area is not solely caused by growing population numbers but particularly in prosperous and urbanised regions by increasing demand for settlement area per person due to higher living standard and income (Loibl and Tötzer 2003, Tötzer et al. 2007). Thus, for modelling human actors as driving force for land use change, demand-related factors like settlement area/consumption per head, settlement density, household size, etc. have to be considered beside mere population numbers.

In this paper we will present first findings from our research within PLUREL, an Integrated Project funded within the 6th framework program of the EC. PLUREL deals with the relationships between urban, peri-urban and rural land use and aims to develop strategies for a sustainable development of these interlinked rural-urban regions (RURs). A key product will be the sustainability impact assessment tool for urban-rural linkages. PLUREL applies the DPSIR-concept (Driver-Pressure-State-Impact-Response) on different European scales: from NUTS0 to LAU2 in six European case study regions.

The research presented in this paper explores population as driving force for land use change focussing on settlement patterns and dynamics. We work on two different scales: on European scale and on micro-scale in six case study regions (among them e.g. Warsaw, Leipzig, Manchester). In this project we will examine if the findings from preceding research projects as well as from the PLUREL case studies could be confirmed on European scale, so that European-wide rules can be derived for certain region types (e.g. urban-monocentric, dispersed peri-urban, rural, urban-polycentric) across Europe. We will apply statistical analysis methods and carry out analyses on European NUTS3- level as well as on local scale for the case study regions.

Research on micro-scale allows in-depth analyses based on detailed data (e.g. LAU2) and gives insights into region-specific interdependencies between society, economy and environment. However, considering the macro-scale as well is particularly important for stakeholders and politicians in the EU. Thus, up-scaling the case study specific rules to the European scale will be an essential task within the project. With the help of pre-defined region types, results which correspond on local and on European scale will be proved and up scaled. Due to data constraints proxy-data will have to be used on European level.

The research will result in statistically based functions reflecting relationships between population and settlement areas in different region-types across Europe. These functions will be considered in many following work packages within the PLUREL project and will help to model the Sustainable Impact Assessment Tool for Rural-urban Relations – one of the key products of the PLUREL project.

REFERENCES

Loibl, W., and T. Tötzer, Modeling Growth and Densification Processes in Sub-urban Regions – Simulation of Landscape Transition with Spatial Agents. *Environmental Modelling and Software. Volume 18, Issue 6*, pp. 485-593, 2003.

Tötzer, T, M. Köstl, K. Steinnocher, Scenarios of land use change in Europe based on socio-economic and demographic driving factors. *Proceedings of the 12th International Conference on Urban Planning & Regional Development in the Information Society*, 2007.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Land use policy impacts in the rural-urban region: a modelling framework

S. Scatasta[a]

[a] *Centre for European Economic Res., ZEW, L7,1 Mannheim, Germany (scatasta@zew.de)*

Keywords: *land use policy, impact assessment, urban region.*

Because land is a fundamental input of anthropogenic (production, residence, recreation etc.) as well as non-anthropogenic (ecosystem functioning, refuge, habitat etc.) activities, modelling such impacts require a conceptual framework multi-sectoral and multi-zonal in scope. This framework should account for flows of people, commodities and services from one sector to the other and from one zone to the other. Because of the very nature of the problem at hand, an urban and rural context needs also be given to the conceptual framework. In this paper we propose an extension to the PACE model inspired by the modelling approach of metropolitan input-output models (see Jun, 1999 and 2005).

PACE is comparative-static multi-region, multi-sector CGE model. Primary factors of a region include labour, capital, and fossil-fuel resources. While in PACE regions are usually countries and cities in metropolitan input-output models, in our conceptual framework we refer to the following three regions: urban, peri-urban and rural. Each region can be divided in sub-region depending on the spatial disaggregation of available data. Capital is divided into two categories: manufactured capital (including infrastructure) and non-manufactured capital (including forests, mineral deposits, fisheries, biodiversity etc.). We also add land as a further resource representing land that is not used for capital. In PACE the energy sector is modelled explicitly to better account for impacts of changes in climate policy. In our framework we add the real estate sector and show how the decision to develop land can be modelled using a real option approach.

Nested constant elasticity of substitution (CES) cost functions are employed to specify the substitution possibilities in domestic production between capital, labour, land, energy and non-energy, intermediate inputs, i.e. material. Final demand in each region is determined by a representative agent, who maximizes utility subject to a budget constraint.

In PACE, furthermore, labour and capital are intersectorally mobile within a region but cannot move between regions. A discussion on data needs to make this conceptual framework operative follows.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Quantification of land use changes, modelling of processes and impacts of shrinkage and the question of scale

S.Schetke[a]

[a] *Institute of Geodesy and Geoinformation, Dept. of Urban Planning and Real Estate Management, University of Bonn, Germany, (schetke@uni-bonn.de)*

Keywords: *land use changes, impact analysis, quality of life, ecosystem services, scale.*

Introduction

Following the purpose of understanding and analysis of interactions between society and landscape in urban areas this paper tries to give a methodological and practical insight into the assessment of socio-environmental impacts under conditions of uneven spatial development. In that context the paper draws special attention to the quantification and modelling of land use changes and shrinkage processes, in particular, based on empirical ascertainable and communal data and empirical research techniques. A second focus highlights a conceptual approach for the assessment of functionality of urban green spaces against the concepts of quality of Life (QoL) and associated Ecosystem Services (ESS) in urban areas dealing with both quantitative modelling and emprirical qualitative research techniques. In a third focus the question of choosing the right scale for assessment and modelling of land-use changes will be dealt with as spatial dynamics and their socio-environmental effects seem to be very variable on a smaller scale and extreme within shrinking cities.

Interactions of society and landscape

Friction is the prerequisite for dynamism and interaction. Heterogenous and uneven spatial development are that friction between landscape and society in shrinking cities. Regarding a surplus of open spaces resulting from enormous vacancies and demolition on the one side and a decreasing number of inhabitants on the other both domains appear to be under pressure. How do these citizens deal with their changed – sometimes blurred and unstructured- living surroundings, what kind of socio-environmental impacts derive from that dynamism and come to pass in post-shrinkage landscapes? In that context I will introduce into a recent study focussing on the quantification and assessment of socio-environmental impacts according to various demolition scenarios in the city of Leipzig, Eastern-Germany (Schetke & Haase 2007, in press). Main outcome of this study is a set of indicators focussing on both structural-ecological assessment (e.g. LSM) as well as on modified social spaces deriving from modified accessibilities of social infrastructure (green and technical ones). The emphasis of this set of indicators is the quantification and socio-environmental assessment of land use changes and shrinkage processes using empirical ascertainable (e.g. communal) quantitative and qualitative data.

Attractiveness of urban life. Concepts for assessment

After the introduction into a model of the assessment of socio-environmental impacts I want to expand these findings by going a bit deeper into the analysis of the attractiveness and quality of urban live by presenting a conceptual approach (see Schetke, Haase & Breuste 2008, in prep.) focussing on the functionality of green spaces accoording to the concepts of quality of life (see Burgess et al. 1988; Givoni 1991; Kawka & Sturm 2006) and associated ecosystem services (Bolund & Hunhammar 1999; Constanza 1997 et al.; de Groot et al. 2002).

Following the presentation of an indicator set quantifying and assessing socio-environmental impacts in shrinking this conceptual approach widens its view by focussing more clearly on the interactions between society and landscape in urban areas. To achieve this not only the area of research has been expanded by adjusting the model towards the analysis of the functionality of urban green under conditions of uneven – that means shrinkage and growth side by side- spatial development. But also the compilation of data implemented into the model has to be enlarged using not only administrative quantifiable land-use and social data but qualitative empirical data highlighting the perceptive side of land- use changes by clearly focussing on social-spatial interdependencies, as well.

Beside the scientific claim in analysing the socio-spatial interactions in urban landscapes the focus on specificly the functionality of urban green is founded on the fact that urban green and open spaces are among the driving factors of quality of life and of that attractiveness of urban life planners are in charge to provide.

Keeping in mind that this affects both planners and scientists the gap between the two domains can be bridged by the concious implementation of valuable communal and empricial ascertainable data into such a model. Also planners are more and more in charge to estimate spatial and social dynamics in order to remain able to steer a sustainable spatial development and to promote the attractiveness of urban life. In that concext a conceptual model not only dedicated to quantify land-use changes but also able to analyse and to depict socio-spatial interactions contributes to both side´s gain of knowledge and progress.

Spatial and land use effects caused by demographic change. Chosing the right scale

In shrinking cites socio-environmental dynamics can be very heterogenous and diverse amongst different neighbourhoods and urban structure types. Both ecological and social affairs - each type of a single urban ecosystem, each social group of citizens- are affected in adifferentw ay with sometimes severe and sometime almost unnoticeable impact. The more, the process of shrinkage and the often accompanying demolition occurs in different spatial and temporal ways which make a general large-scale assessment and modeling (e.g. on city or regional scale) of land use changes and shrinkage processes difficult and redundant in meaning. The question of scale is long discussed in terms of land use assessment. Especially in terms of quantifying and modelling of both presented socio-environmental impacts of shrinkage and a more indepth analysis ofsocio-spatial interactions and the functionality (see QoL and ESS) of urban green it is still a sensitive matter to discuss and to keep in mind.

REFERENCES

Bolund, P., Hunhammar, S. 1999: Ecosystem Services in urban areas. In: Ecological Economics 29, 293-301.

Burgess, J. et al. 1988: People, Parks and Urban Green: A Study of Popular Meanings and Values for Open Spaces in the City. In: Urban Studies 25, 455-473.

Constanza, R. et al. 1997: The value of the world's ecosystem services and natural capital. Nature, 387, 253-260.

De Groot, R. et al. 2002: A typology for the classification , description and valuation of ecosystem functions, goods and services. In: Ecological Economics 41, 393-408.

Givoni, B. 1991: Impact of planted areas on urban environmental quality: a review. In: Atmospheric Environment Vol. 25B No.3, pp. 289-299.

Kawka, R., Sturm, I. 2006: Objektive regionale Lebensqualität und subjektives Wohlbefinden- Was macht Bürgerinnen und Bürger zufrieden? In: Informationen zur Raumentwicklung Heft 6/7 2006.

Schetke, S., Haase, D. 2007: Multicriteria assessment of socio-environmental aspects in shrinking cities. Experiences from eastern Germany. In: Environmental Impact Assessment Reviews (in press).

Schetke, S., Haase, D., Breuste, J. 2008: Recreational land use and green space functionality in uneven spatial urban development (in prep.).

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Integrated Modelling of Smallholder Land-use Decisions on the Availability of Rattan

Frank v. Walter[a], **Jan Barkmann**[a]

[a]*Georg-August-Universität Göttingen, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany (fwalter2/jbarkma@gwdg.de)*

Abstract: Although being officially protected areas, the margins of national parks are often cultivated by local agricultural smallholders. This is the case for example in the Lore Lindu Park in Central Sulawesi/ Indonesia or the Podocarpus National Park in Southern Ecuador.
To assess the impacts of the decisions of smallholders in these margins, we propose an integrated modelling framework, with which we are able to analyse the economic decision-making process of smallholders and the ecological and economic consequences.
The main component of this framework is an Agent-Based Model, which allows us to differentiate between the perception of individual agents and system response modelled with 'full' information. The land-use decision-making options will be the result of an economical optimized result of the production opportunities.
We first apply this model to rattan collectors in Lore Lindu National Park in Central Sulawesi.

Keywords: Integrated Modelling; Multi-Agent-Systems; Sulawesi

1 INTRODUCTION

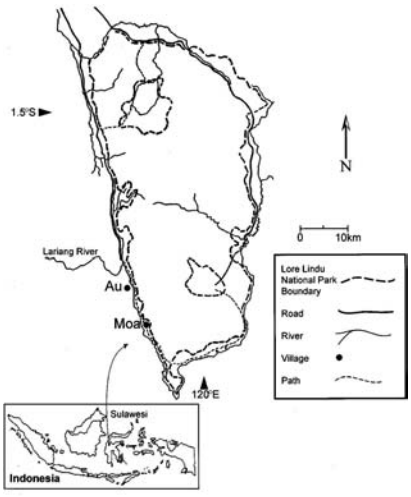
The margins of national parks belong officially to the protected area of the parks. However, they are used for the livelihood of smallholders in the park's vicinity. Conflicts and continued poverty around protected areas suggest that existing approaches to conservation lack understanding of links between maintenance of diversity and human well-being [Bawa et al., 2004]. Regularly, the livelihood of the rural poor includes agricultural land-use. Large areas are converted by smallholder farmers, despite the fact that tropical rain forests play a major role in the provision of ecosystem services [Achard et al., 2002]. The interaction between human and natural systems at the rain forest margin play a key role in preserving and stabilising forested ecosystems [Maertens et al., 2006]. Ecosystem stability is a multi-dimensional concept requiring multi-disciplinary analysis because it concerns economic, ecological and social issues. This way we hope to identify more sustainable development options, such as mitigation of climate change effects and biodiversity loss, poverty alleviation and economic development for rural areas [Balmford et al., 2002].
One tool to address the livelihoods of the smallholders as well as impacts on ecosystem functioning are integrated assessment models. Mallawaarachchi and Quiggin [2001] e.g., present a model for the integrated assessment of sugar cane production in Queensland/ Australia. We will extend such an approach by a spatially explicit agent-based-modelling component. Recently the application of spatially explicit agent-based models have become a focus in the research on land-use and cover change [Bousquet and Le Page, 2004]. This modelling concept - originally from the field of problem solving in (distributed) artificial intelligence - provides an excellent frame to simulate the problem at hand. With the SIEHL-framework (**S**patially and **I**nstitutionally **E**xplit **H**ousehold decision-making and **L**and use model) we apply this rationale to the margins of tropical rainforests, first to Indonesia (Lore Lindu National Park); *In-SIEHL* and later to Ecuador (Podocarpus National Park; *Ec-SIEHL* see section 4).

2 RESEARCH SITE

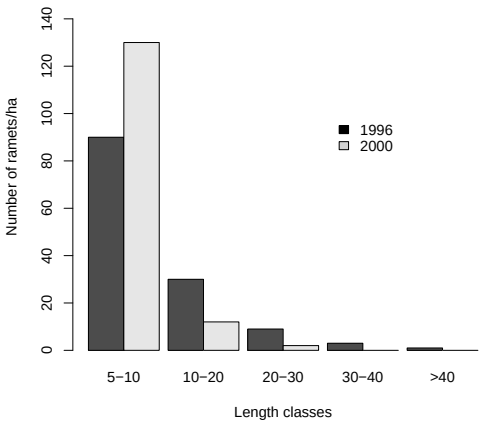
Lore Lindu National Park (LLNP) is situated in Central Sulawesi, Indonesia $(1°80' - 1°30'S;$ $119°90' - 120°16'E)$. The park is a core area of the Wallacea biodiversity hotspot [Myers et al., 2000], harbouring many endemic plant and animal species. LLNP covers an area of 217,000 ha, which consists mainly of lower montane forest. Annual rainfall ranges from 2000 to 4000 mm. LLNP was declared as a *"Man and the Biosphere Reserve"* (MAB) by UNESCO in 1977 [Pangau-Adam et al., 2006]. Especially the margins of the national park are encroached by smallhold farmers and the natural land cover is changed. This encroachment includes logging, planting of cash-crops (coffee, cacoa, rice) and extraction of wood and non-timber forest products (NTFP) such as rattan. In addition to forest conversion, cacoa agroforestry is intesified, a process that further increases pressure on biological diversity Steffan-Dewenter et al. [2007]. Thus the pressure on the forest within the national park increases. In some regions of mountainous Sulawesi, this already as led to a virtually complete conversion from forest to cultivated land [e.f. Barber and Matthews, 2002].
In villages of Au and Moa in the vicinity of LLNP rattan is a major source of income. We chose this area to carry out our first case-study. Moa and Au are located in the south-west of the LLNP (see figure 1(a)). There the abundance and distribution of rattan (*Calamus zollingeri*) was measured along 5 transects in the years 1996 and 2000 and a strong decline in abundance was observed (see figure 1(b)). The transects were located along a height-gradient with those of Moa located within the National Park and those of Au just outside the park's borders.

2.1 Importance of rattan

Rattans (*Calamus*), spiny climbing or trailing palms with some 600 species, can only be found in the Old World [Sastry, 2001]. It is an important non-timber forest product (NTFP), which is mainly collected from natural forests. However, harvest of rattan canes is far from being sustainable, as more canes are extracted than regrow [Dransfield, 2000; Siebert, 2004]. Rattan plays an important role in smallholder economy: it has the largest share of non-timber forest products, which are sold in the LLNP area [Schwarze, 2004]. However, rattan extraction is extremely hard work. Therefore only the poorest groups of residents are involved in rattan extraction. Because of its local importance as an NTFP we chose rattan as an example resource for the first implementation of the SIEHL-framework.



(a) Villages of Au and Moa. Taken from [Siebert, 2000]



(b) Distribution of *Calamus zollingeri* canes (ramets) in 1996 and 2000 (from Siebert [2004]).

Figure 1: Project Area and distribution of rattan *Calamus zollingeri*

3 SIEHL-FRAMEWORK

In the SIEHL-framework, we use integrated assessment methods to combine knowledge of socio-economic and ecological impacts of land-use changes in a spatially explicit way on the landscape scale. The SIEHL-framework consists of five parts: a Central Java-Application, a Multi-Agent-System (Repast - MAS), a database the implemented (and optimised) land-use-decisions, ecological and economic impact modules and an evaluation component (see figure 2).

3.1 Multi-Agent-Model

The core element of the IN-SIEHL-framework is a Multi-Agent System (*M*AS). According to the Multi-Agent-Paradigm, agents are software representations

> "... that are situated in some environment and capable of autonomous actions in this environment in order to meet their design objectives" [Wooldridge, 2001].

MAS allow us to simulate the land-use decisions of smallholders on an individual level. Particularly, we can differentiate between how an individual agent perceives its environment (*'Merkwelt'*) and the 'full-information' system response (*'Wirkwelt'*). A common framework for the agent's rationale is the *Belief-Desire-Intention*-Architecture (see fig. 3(b)).
In our case, the framework can be explained as follows: Each agent has a certain *Desire* (*profit-maximisation*), which he tries to pursue. An ('Wirkwelt') stimulus triggers the revision of the *Belief* ('Merkwelt'). According to its beliefs, a set of options is generated, from which the agent choses one option which will – according to his belief – fit best to fulfill its desire (*Intention*).
For rattan collection This can be exemplified by the problem to actually design an optimal rattan harvesting strategy that maximises long-term rattan benefits for a poor household. The smallholders know about putative rattan sites (Merkwelt) and also roughly to which extent harvesting has taken place there. However, actual rattan harvesting decisions are not made from an omniscient position according to analyst knowledge on potentially optimal and sustainable harvesting strategies as obtained from an analysis of all Wirkwelt data available in the model.

Agents. In the Central Sulawesi In-SIEHL implementation. the agents in the model are aggregated entities, because they do not represent individual humans, but smallholder households as a decision-maker. Data of the availability of the capital were collected in social science surveys and fed into a database (Data from the STORMA-Project and Schwarze [2004]). These sources supply data on capital access and restrictions to economic activity of different classes of smallholder households. Off-farm work will be modelled as rattan extraction. Particularly for remote villages such as Au and Moa this is highly realistic. This decision-process will be optimised as a result of an analysis (Lingo) of single household production opportunities (see figure 3(a)) Every agent has a plot of land he uses exclusively. The main activities of an agent is to decide on and implement the land use of their plot. .

Bio-physical Environment. The original landscape data from the STORMA-project will be used. Based on LandSat-images on a 30×30 m grid, we use the data on land-cover and altitude. Agents are located in the villages and rattan will be distributed according to a height gradient. Furthermore previous harvesting activities will be taken into account. With an GIS-Analysis and expert knowledge, the most probable entry points for rattan-harvesting will be located and supposed trails (e.g. a gradient with the smallest slope) used for modelling purposes.
The abundance data of *Calamus zollingeri*, is used for initialisation Siebert [2004]. Since only two initial factors are known, which determine the abundance of rattan (distance to town and altitude), those will be used to simulate the distribution of rattan.
After initialisation, the agent will decide – for example every "year" – on its land-use, which will be "implemented" subsequently. This will be the the yield that corresponds to the harvesting rates, that can be estimated using the Siebert [2004] data.

3.2 Technical structure

The technical Framework structure is presented in figure 2. From a central Graphical-Users Interface (*JAVA-Application*), different background scenarios (price, climate variation) and evaluation components can be selected, and the simulation of the MAS is triggered. The MAS is modelled in Repast (`http://repast.sourceforge.net`), a toolkit with various built-in methods for data analysis and extension possibilities such as database connectivity or GIS-extension. The economic and ecological data needed for simulation are loaded from an Entity-Relationship-database. In this database, various data from different sources are stored, i.e. Land-Sat images on a 30×30 m basis, information on vegetation cover and detailed household survey data including geo-referenced resource access.
Depending on the research question, different 'impact-modules' can be included into the framework. The ecological module 'implements' the land-use decision in 'Wirkwelt',... its ecological consequences; e.g. on the hydrological cycle or on rattan stocks.
The economic impact is a consequence of the implemented land-use decision, too. In this case exogenous factors (climate, ...) influence crop yield, which might be very different from what the agent anticipated ('Merkwelt'). The profits are added to the financial capital. The results of the impacts are again fed back into the database and can be used for further analysis.

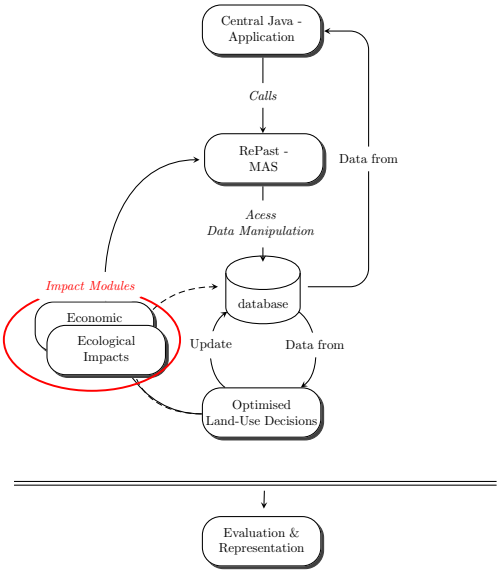
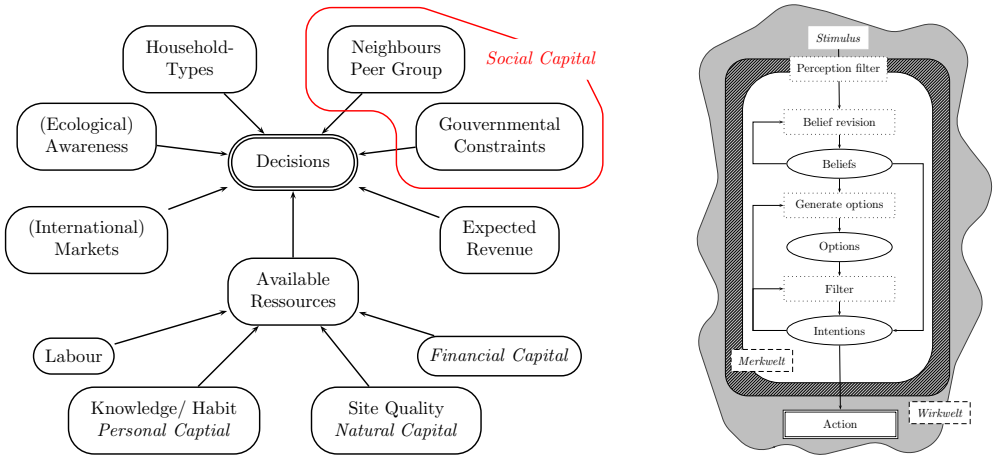


Figure 2: Program-flow

3.3 Calibration and Validation

Data from work of Siebert [2000] and Siebert [2004] will be taken to estimate the spatial distribution of the rattan. However, both publications only offer rattan data in an agreggated form. In order to get a more realistic spatial distribution we apply the following steps. First, we calculate a weighted path raster. These paths are those, which are travelled most likely by the rattan collectors. We assign the weights, which stand for the movement costs. The values of the weights are chosen in accordance with expert knowledge. This means that travelling along rivers has the lowest movement cost and travelling within the forest has the highest costs. The cells of this raster contain the accumulated cost values. In the next step we combine the information of the rattan on the height gradient of Siebert [2000] with this raster. As rattan stocks of 1996 are already subject to harvesting we $a_i i^b b$



(a) Influencing factors for decision-making

(b) Belief-Desire-Intention Architecture

Figure 3: Schemes for the decision-making process and -factors

4 Future Work

In the course of the next three years the framework will also be applied to *Podocarpus National Park* (PNP) in Southern Ecuador. The park is located between the cities of Loja and Zamora. Situated between the Andean and Amazonian climates it is one of the hottest "hotspots" in biodiversity [Myers et al., 2000; Brehm et al., 2008]. The park covers an area over 146,300 hectares, with elevations ranging from around 1000 m to 3690 m. The main research area of the biodiversity within the park and the adjacent pastures is the "Estación Científica San Francisco" (ECSF $3°58'S$; $79°4'W$). However, protected areas, as for example the region of Corazón de Oro (a fragmented forest of 533 km^2 north of ECSF, is inhabited and cultivated by smallholders, even though they do not have land titles. The main threat for the stability is therefore (illegal) conversion of forest to pastures, as livestock provides far higher revenues than any other land-use types [Wunder, 2000], as in Indonesia, the profit-maximizing smallholder will sooner or later switch to this cash "crop".

REFERENCES

Achard, Frederic, Hugh D. Eva, Hans-Jurgen Stibig, Philippe Mayaux, Javier Gallego, Timothy Richards, and Jean-Paul Malingreau, "Determination of Deforestation Rates of the World's Humid Tropical Forests," *Science*, August 2002, *297* (5583), 999–1002.

Balmford, Andrew, Aaron Bruner, Philip Cooper, Robert Costanza, Stephen Farber, Rhys E. Green, Martin Jenkins, Paul Jefferiss, Valma Jessamy, Joah Madden, Kat Munro, Norman Myers, Shahid Naeem, Jouni Paavola, Matthew Rayment, Sergio Rosendo, Joan Roughgarden, Kate Trumper, and R. Kerry Turner, "Economic Reasons for Conserving Wild Nature," *Science*, August 2002, *297* (5583), 950–953.

Barber, Charles Victor and Emily Matthews, *The state of the forest: Indonesia.*, FWI/GWF, 2002.

Bawa, Kamaljit S., W. John Kress, Nalini M. Nadkarni, Sharachchandra Lele, Peter H. Raven, Daniel H. Janzen, Ariel E. Lugo, Peter S. Ashton, and Thomas E. Lovejoy, "Tropical Ecosystems into the 21st Century," *Science*, October 2004, *306* (5694), 227 – 228.

Bousquet, F. and C. Le Page, "Multi-agent simulations and ecosystem management: a review," *Ecological Modelling*, September 2004, *176* (3-4), 313–332.

Brehm, G., J. Homeier, K. Fiedler, I. Kottke, J. Illig, N.M. Nöske, F. Werner, and S.W. Breckle, "Mountain Rain Forests in Southern Ecuador as a Hotspot of Biodiversity – Limited Knowledge and Diverging Patterns," in Erwin Beck, Jörg Bendix, Ingrid Kottke, Franz Makeschin, and Reinhard Mosandl, eds., *Gradients in a Tropical Mountain Ecosystem of Ecuador*, number 198. In 'Ecological Studies.', Springer, 2008, chapter 1, pp. 1–13.

Dransfield, John, "Species Profiles Rattans," in Florentino O. Dransfield, John ; Tesoro and N. Manokaran, eds., *Rattan, Current research issues and prospects for conservation and sustainable development*, Vol. 14 of *Non-Wood Forest Products* FAO INBAR December 2000, pp. 23–34.

Maertens, Miet, Manfred Zeller, and Regina Birner, "Sustainable agricultural intensification in forest frontier areas," *Agricultural Economics*, 2006, *34* (2), 197–206.

Mallawaarachchi, Thilak and John Quiggin, "Modelling socially optimal land allocations for sugar cane growing in North Queensland: a linked mathematical programming and choice modelling study," *The Australian Journal of Agricultural and Resource Economics*, September 2001, *45* (3), 383–409.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent, "Biodiversity hotspots for conservation priorities," *Nature*, February 2000, *403* (6772), 853–858.

Pangau-Adam, Margaretha, Matthias Waltert, and Michael MÃijhlenberg, "Nest Predation Risk on Ground and Shrub Nests in Forest Margin Areas of Sulawesi, Indonesia," *Biodiversity and Conservation*, December 2006, *15* (13), 4143–4158.

Sastry, C.B., "Rattan in the twenty-first century - an overview," *Unasylva*, 2001, *52* (205), 3–10.

Schwarze, Stefan, "Determinants of Income Generating Activities of Rural Households A Quantititive Study in the Vicinity of the Lore-Lindu National Park in Central Sulawesi/ Indoensia." Dissertation, Department of Agricuture 2004.

Siebert, S. F., "Survival and growth of rattan intercropped with coffee and cacao in the agroforests of Indonesia," *Agroforestry Systems*, 2000, *50*, 95–102.

Siebert, Stephen, "Demographic Effects of Collecting Rattan Cane and Their Implications for Sustainable Harvesting," *Conservation Biology*, 2004, *18* (2), 424–431.

Steffan-Dewenter, Ingolf, Michael Kessler, Jan Barkmann, Merijn M. Bos, Damayanti Buchori, Stefan Erasmi, Heiko Faust, Gerhard Gerold, Klaus Glenk, S. Robbert Gradstein, Edi Guhardja, Marieke Harteveld, Dietrich Hertel, Patrick Hohn, Martin Kappas, Stefan Kohler, Christoph Leuschner, Miet Maertens, Rainer Marggraf, Sonja Migge-Kleian, Johanis Mogea, Ramadhaniel Pitopang, Matthias Schaefer, Stefan Schwarze, Simone G. Sporn, Andrea Steingrebe, Sri S. Tjitrosoedirdjo, Soekisman Tjitrosoemito, Andre Twele, Robert Weber, Lars Woltmann, Manfred Zeller, and Teja Tscharntke, "Tradeoffs between income, biodiversity, and ecosystem functioning during tropical rainforest conversion and agroforestry intensification," *PNAS*, 2007, *104* (12), 4973–4978.

Wooldridge, Michael, *An Introduction to Multi-Agent-Systems*, John Wiley & Sons, 2001.

Wunder, Sven, *The Economics of Deforestation - The Example of Ecuador*, Macmillan Press, 2000.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Using a Bayesian network to assess the population-level risks of endocrine disruptors to brown trout

Mark Borsuk[1], Sara Schubert[2], Armin Peter[2], Patricia Burkhardt-Holm[3]

[1] Thayer School of Engineering, Dartmouth College,
Hanover, NH 03755 USA (mark.borsuk@dartmouth.edu)

[2] Swiss Federal Institute of Aquatic Science and Technology (Eawag),
6047 Kastanienbaum, Switzerland

[3]Department of Environmental Sciences, University of Basel,
4051 Basel, Switzerland

Abstract: Much has been learned recently about the biochemical, morphological, and reproductive effects of endocrine disruptors on wildlife, both in the laboratory and in the field. However, relatively little is known about how individual-level effects will impact population persistence. It is possible that the parameter that is the most sensitive indicator of endocrine exposure may not be the most ecologically relevant one. Additionally, to date, much research has focused on directly reproductive attributes such as fecundity or fertility. However, it is possible that more subtle changes induced by endocrine disruptors (e.g., modifications of growth rate or time of maturation) may also have important population-level consequences.

The practical difficulties of performing controlled experiments on entire populations over multiple generations means that the population-level effects of endocrine active substances are unlikely to be revealed through experimentation alone. Further, the dependence of population dynamics on environmental variables other than chemical exposure, such as climate, habitat, and prey availability, implies that any empirical studies will be highly site-specific. Dynamic simulation models represent an alternative means of population assessment. These mathematical representations of a species' lifecycle can be used to translate effects on the individual to measures of population persistence such as rate of change, extinction risk, or long-term abundance. Models can then be parameterized for different locations using appropriate site-specific habitat and stressor data to assess the susceptibility of a population living under one set of conditions compared with another.

Unfortunately, most natural populations are highly variable, and nearly all predictive models are uncertain. Therefore, it is an open question as to whether any population changes that may be caused by exposure to endocrine disruptors will be distinguishable against the significant noise in model assessments. We addressed this question using a detailed population model developed for brown trout (*Salmo trutta*) in Switzerland and a set of recent research findings concerning the individual-level effects of endocrine disruptors on this species. Being in the form of a Bayesian network, the model includes rigorous probabilistic representations of natural variability and model uncertainty and is also transparent with regards to causal structure. Elasticity analysis was used to assess the sensitivity of the population to various endocrine disruptor-induced impacts, which were then compared against model prediction intervals.

Keywords: belief networks, causal models, uncertainty, model evaluation, environmental management and decision-making

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Assessment of agri-environmental measures on water quality using a Bayesian Network

M. Carpani [a], **C. Giupponi** [b]

[a] *Department of Crop Production, University of Milan, Milan, Italy (marta.carpani@unimi.it)*
[b] *Center for Environmental Economics and Management, Department of Economic Science, Ca' Foscari University of Venice, Venice, Italy (cgiupponi@unive.it)*

Keywords: Bayesian Network; water quality; nitrogen; policy evaluation.

The present abstract refers to a work in progress for developing a model to assess water quality in the sensitive area of Northern Italy, the Venice Lagoon Watershed (VLW).
Considering that in the past the VLW have been responsible of severe eutrophication events of the lagoon and that in the past decade the municipal and industrial loads have been reduced, through significant investments in treatment plants, agriculture activities are now on top of the agenda regarding the pollution sources to be further controlled and abated.
The Master Plan of the Regional Government has thus planned actions to reduce pollutants loads in the lagoon and among them is a package of agri-environmental measures (D.G..R. 2116, 02/08/2002) promoting:

- reduction of applied water volumes,
- creation of buffer strips and set-aside,
- introduction of a low input agriculture,
- structural interventions for breeding.

The issue of assessing the effectiveness of those investments has raised relevance for ex-post evaluation of the policy measures and for planning new strategies for the future within the local implementation of the Water Framework Directive (CE/200/60). In order to do that, scenario analysis and modelling tools are needed for the integrated assessment of both the water cycle at the watershed scale and the agriculture activities. Simulation models for diffuse pollution of agricultural origin provide a great potential within this context, but the complexity of the problems requires also that model outcomes are integrated with other sources of information, such as statistics, qualitative data, and expert knowledge.

The conceptual model proposed by the NetSyMod approach [Giupponi et al., 2008] was used in a research project financed by the Regional Environmental Protection Agency as the methodological framework for participatory modelling and policy making support. As a hydrologic model engine, SWAT (Soil and Water Assessment Tool) [Arnold, 1998] was used for the simulation of nutrient loads, transport and discharge [Salvetti et al., 2007], within two alternative scenarios: with and without measures implementation. Expert knowledge was instead elicitated by means of a participatory modelling workshop. Parallel activities were conducted to acquire the other information needed, for instance from census data.

The research referred herein focused on the analysis of the potentials of the Bayesian Network (BN) approach to provide an intermediate and comprehensive modelling structure, allowing to make operational the – upper level and synthetic – conceptual model developed with participatory modelling within the DPSIR framework (Driving force, Pressure, State,

Impact and Response) [EEA, 1999], in integration with the – detailed mechanistic – SWAT model and the other – lower level and detailed – sources of information.

The BN designed for the contribution to lagoon nitrate pollution is reported in Figure 1.

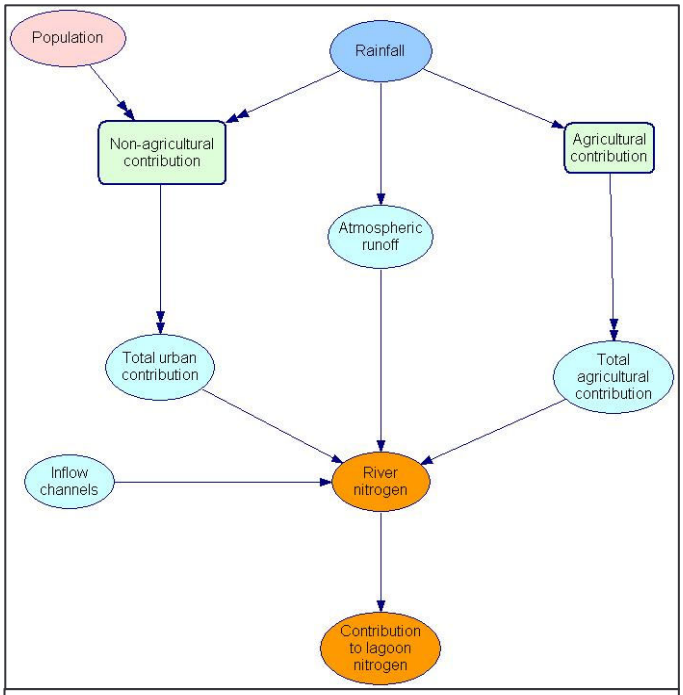


Figure 1. Overall Bayesian Network for the Venice Lagoon Watershed.

In order to focus specifically on the agri-environmental measures an operational BN was developed (Figure 2). The network represents the most important nodes and their causal connections, elaborated upon the outcomes of participatory modelling and consolidated through the analytical approach provided by the SWAT model. The BN considers first financial aspects related to the amount of resources made available by the Regional Government for the various voluntary agri-environmental measures, and the degree of uptake by farmers. Nodes follow to represent the physical aspects of cause-effect relationships between agricultural activities, the water cycle and the quality of discharged water. As a final node the quality of rivers flowing into the lagoon is reported.

The software utilised in this study to create and deal with Bayesian Networks is GeNIe (http://genie.sis.pitt.edu/) that manage just discrete variables, therefore it was necessary to discretise all the continuous variables, data and model outputs (i.e. SWAT outputs) in particular.

Currently, the BN is being presented to three different group of experts with diverse competences: administrative, agricultural and water experts. Each expert validates the agricultural network and, by means of an ad hoc questionnaire s/he fills the conditional probability tables for a group of nodes related to his/her expertise.

Once these steps will be concluded the model will be fully functional and then updating with monitoring data and model outputs are planned to represent the two scenarios to be analysed. Such knowledge update will lead to the assessment of agri-environmental measures effectiveness, but also, in the future may represent an ex-ante assessment tool, for resource allocation to alternative measures. This potential of the network could be extremely useful in a framework of policy evaluation considering as well, as pointed out by Usitalo [2007], the possibility of studying consequences of different management decisions from two perspective: expected values and risks of highly undesirable outcomes.

So far, the BN approach has shown great potential for bridging hard-science models with participatory techniques and modelling, thanks to facilitated communication through user-friendly graphical interface, making explicit cause-effect relationships among systems variables. In particular, BN approaches may represent an effective and functional intermediary between qualitative cognitive mapping and simplified conceptual frameworks, such as the DPSIR, and dynamic modelling allowing the exploration of the overall structure and main relationships of complex socio-ecosystems. Several drawbacks and limitations should also be mentioned, such as the need to discretise continuous functions for operational reasons, the limited details manageable with the networks in order to avoid exploding complexity of conditional probability functions, or the lack of means for representing feedback loops.

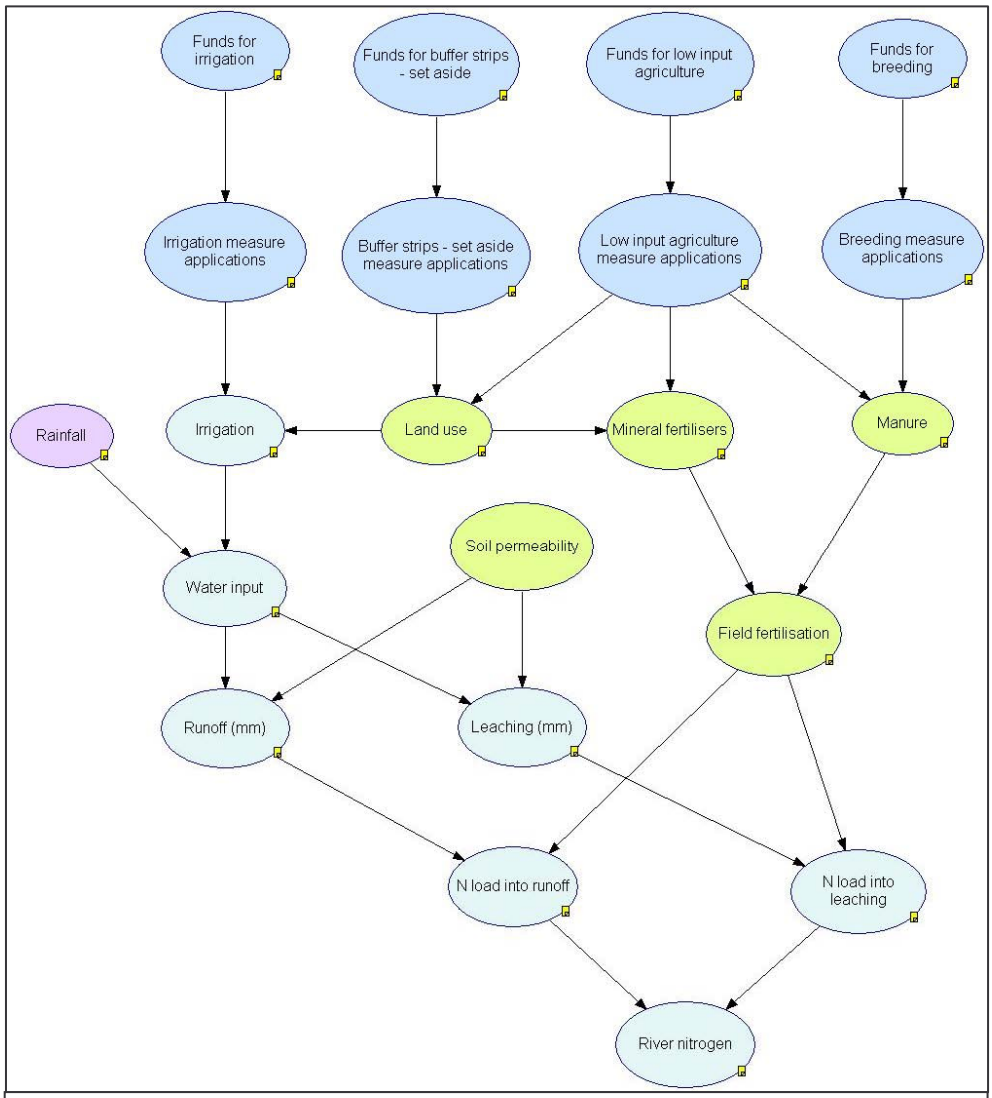


Figure 2. Agricultural Bayesian Network for the Venice Lagoon Watershed.

REFERENCES

Arnold, J.G., R. Srinivasan, R.S. Muttiah and J.R. Williams Large area hydrologic modeling and assessment part I: model development. *Journal of the American Water Resources Association*, 34 (1), 73-89, 1998.

EEA (European Environmental Agency), Environmental indicators: typology and overview, *Technical Report No 25*, Copenhagen, 1999.

Giupponi, C., A. Sgobbi, J. Mysiak, R. Camera, and A. Fassio, NetSyMoD: an integrated approach for water resources management. In: P. Maire, M. Coenen, C. Lombardo, M. Robba, and R. Sacile (eds.), *Integrated water management: practical experiences and case studies,* Springer, 69-93, 2008.

Salvetti, R., A. Azzellino, D. Gardoni, R. Vismara, M. Carpani, C. Giupponi, M. Acutis, M. Vale, P. Parati, Application of SWAT Model on Three Watersheds within the Venice Lagoon Watershed (Italy): Source Apportionment and Scenario Analysis, presented at 4th International SWAT Conference at UNESCO-IHE, Delft, The Netherlands, July 2-6, 2007.

Uusitalo, L., Advantages and challenges on Bayesian networks in environmental modelling, *Ecological Modelling,* 203, 312-318, 2007.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Integrated methodology for the management of uncertainty and complexity in climate change adaptation policies: a Bayesian Decision Network approach

M. Catenacci [a,b], C. Giupponi[a,c]

[a] *Fondazione Eni Enrico Mattei, Corso Magenta 63, Milan; Castello 5252, Venice (michela.catenacci@feem.it; carlo.giupponi@feem.it)*
[b] *International School for Advanced Studies, Island of San Servolo, Venice*
[c] *Ca'Foscari University of Venice, Department of Economics, Venice*

Keywords: Adaptation to climate change; Participatory approach; Expert judgment elicitation; Bayesian decision network.

The main aim of the research presented herein is to define and implement an integrated methodological framework to identify and assess potential adaptation strategies to the impacts of sea level rise on coastal areas, and in particular on the Grado-Marano lagoon, in north-eastern Italy.

The approach analyses the potential synergies of combining elements belonging to different research fields, with the aim of contributing to the management of the complexity and uncertainty characterizing climate change issues.

The proposed methodology is built upon the sequential use of specific tools, which are based on the integrated application of conceptual modelling and expert judgment elicitation techniques, decision analysis and probabilistic modelling, within the broader framework of the NetSyMoD methodology (Giupponi et al. 2008; www.netsymod.eu/), previously developed, facilitating the involvement of stakeholders or experts in policy- or decision-making processes.

The assessment of the impacts of sea level rise on the lagoon of Grado and Marano is characterized by the involvement of a high number of disciplines, which apply different measurement and analysis methods, and produce a huge amount of heterogeneous data that need to be analysed jointly to constitute an appropriate support for the definition of effective adaptation policies. Uncertainty is pervasive in the definition of future projections of sea-level change and of technological and socio-economic development. Several attempts have been made to assess the complexity and incorporate the uncertainty in climate change policy analyses and, despite their different perspectives, they all agree on the importance to identify and implement an integrated approach to formally structure environmental and socio-economic issues into a single assessment framework.

The methodology developed for the Grado-Marano case study, proposes a suite of tools, which starts from the definition of a shared conceptual model for the local socio-ecosystem analysis, by means of participatory modelling techniques allowing the integration of knowledge of experts belonging to distinct disciplines. In particular, the main vulnerability factors of the area were discussed during a workshop in which the system's components were identified and the relationships existing among them were collected and discussed using cognitive maps, and then formalized into the DPSIR causal framework.

Later on, a Bayesian Decision Network (BDN) model was built, where nodes represented the main variables of the system and arcs made each cause-effect link explicit. The

conditional probability tables underlying each node were populated using judgments elicited from the experts through interviews and specific questionnaires. Value nodes measuring the expected gain or loss of each economic sector were inserted as outputs of the BN. The propagation of probabilities through the connected nodes made it possible to analyse different scenarios of sea level rise and to assess the effectiveness of alternative adaptation strategies, according to their consequences on the factors of the systems and on the value nodes.

Drawing on the application of the methodology to the case study, the main advantages and the possible shortcomings arising from the development and application of the integrated approach can be preliminarily discussed. The use of cognitive maps enhances the interaction among the experts and the consequent integration of different domains of knowledge into a shared framework of analysis. Its formalization into a Bayesian Decision Network allows the incorporation of expert subjective knowledge with empirical data and information coming from models, improving the management of issues characterized by lack of data. Alternative management approaches can be tested through the development of BNs, and different future scenarios can be studied by entering new information into the network, since this change is propagated throughout the BN, to the endpoints. The updating process of the BN and the use of conditional probabilities within the model implicitly incorporate uncertainty in the outcomes. The Bayesian decision network developed for the case study combines probabilistic reasoning with value functions, allowing to make decisions that maximize the expected value, which could also be expressed in monetary values. Together with the advantages of the tool, a number of issues should also be considered, such as BNs' difficulties in dealing with situations where the complexity needs to be represented in great detail or where feedback loops must be incorporated into the analysis, or the limited potential of the tool to incorporate continuous variables, and the risks linked to the use of experts judgments. Alternative solutions are then proposed to overcome the above limits.

ACKNOWLEDGEMENTS

The authors wish to thank ENEA for the financial support of the VECTOR project. The authors also want to acknowledge to Prof. Ken Reckhow (Duke University) for guiding the preparatory process of the Bayesian network model.

REFERENCES

Giupponi, C., A. Sgobbi, J. Mysiak, R. Camera, and A. Fassio, 2008. *NetSyMoD: an integrated approach for water resources management*. In: P. Maire, M. Coenen, C. Lombardo, M. Robba, and R. Sacile (eds.), Integrated water management: practical experiences and case studies. Springer. 69-93.

Korn K.B., Nicholson, A.E., 2004, Bayesian Artificial Intelligence, Chapman & Hall/CRC

Morgan, M.G., 2007, Best Practice Approaches for Characterizing, Communicating and Incorporating Scientific Uncertainty in Climate Decision Making, CCSP Product, forthcoming.

Morgan, M.G., Henrion, M., 1990, Uncertainty. In: A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, ISBN 0-521-36542-2.

Reckhow, K.H., 2002, Bayesian approaches in ecological analysis and modeling

Uusitalo, L., 2006, Advantages and challenges on Bayesian networks in environmental modelling, Ecological Modelling 203 (2007) 312-318.

Varis, O., Kuikka, S., 1997, BeNe-EIA: a Bayesian approach to expert judgment elicitation with case studies on climate change impacts on surface water, Climatic Change 37, 539-563.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Actor Modeling and Actor-based Modeling: New Methods to Analyze the Human Component of Human-Environment Systems

Petra Döll, Christiane Döll and Alexandra Titz

Institute of Physical Geography, University of Frankfurt am Main, P.O. Box 111932, 60054 Frankfurt am Main, Germany (p.doell@em.uni-frankfurt.de, c.doell@em.uni-frankfurt.de, titz@em.uni-frankfurt.de)

Keywords: Actor; Agent-Based Modeling; Integrated Assessment; Modeling; Perception Graph; Scenarios.

1. Motivation

The future of natural resources depends on the actions of people. Thus, whenever assessing options for a sustainable management of natural resources, it may not good enough to provide decision makers with a decision support system that integrates the technical and natural system components only. To include the human system component, it is state-of-the-art to — in addition to demographic modeling and economic optimization — consider human actions in storylines of qualitative scenarios, which are then translated ad-hoc into quantitative scenarios of input variables of models of the technical and natural systems components. We propose modeling of societal actors and actor-based modeling of physical variables (AM) as science-based and transparent methods for analyzing the human system component and for, in addition, quantifying input to environmental systems models. AM is particularly useful for problems fields where little knowledge is available and where a large number of actors with different values and goals have to cooperate to obtain a sustainable outcome. Actor-based modeling can be regarded as a type of agent-based modeling where the actions of the actors are determined from complex actor-specific perceptions graphs which are the core of actor modeling.

2. Modeling of societal actors

Actor modeling is a software-supported formalized actor analysis, which focuses on modeling the different problem perceptions of institutional actors (e.g. chemical industry or drinking water suppliers) and includes semi-quantitative computations of the utilities of actions, allowing thus the identification of optimal strategies (set of actions). We apply the software DANA (Dynamic Actor Network Analysis) [Bots et al., 1999; Bots, 2007, http://dana.actoranalysis.com]. The problem perception of an actor is modeled by perception graphs (PGs) where the most important goals of the actor, possible actions by the actor itself and by other actors, expectations and relevant factors (boxes) are connected by causal links (influence arrows) among each other in a consistent and transparent form. A PG is a type of causal map, and a directed acyclic graph. All boxes and arrows are given symbols which are translated to quantitative values on a scale of 7, which in the case of the boxes refer to 7 change levels (strong, medium, slight decrease, no change, slight, medium strong increase). Both causal links and expectations can be expressed probabilistically, but this possibility has not been used by us. The premise of actor modeling is that the behavior of actors is led by their own subjective perception of the problem field. Model input are interviews with representatives of relevant actors, where first PGs may be constructed by the representatives.

3. Actor-based modeling

Actor-based modeling refers to modeling of future development of factors (e.g. production of chemicals) based on all actors' PGs and the analyst's PG (which reflects the system view of the analyst and encompasses most of the actions and factors of the actors' PGs). In the analyst's PG, the development of factors due to the change level of all actions is computed, the change level of each action being determined as the (change level of) action with the highest utility in the PG of the actor that performs the action. In each decision round, the actors act in a specified order (Fig. 1). The determination of the optimal action of an actor takes into account the actions taken by the actors before him. Different scenarios are computed by representing the scenario axes as additional actors. For example, if two scenarios differ with the respect of regulation by the EU, a different perception graph for the EU is added such that the optimal EU action is either a more or a less strict regulation.

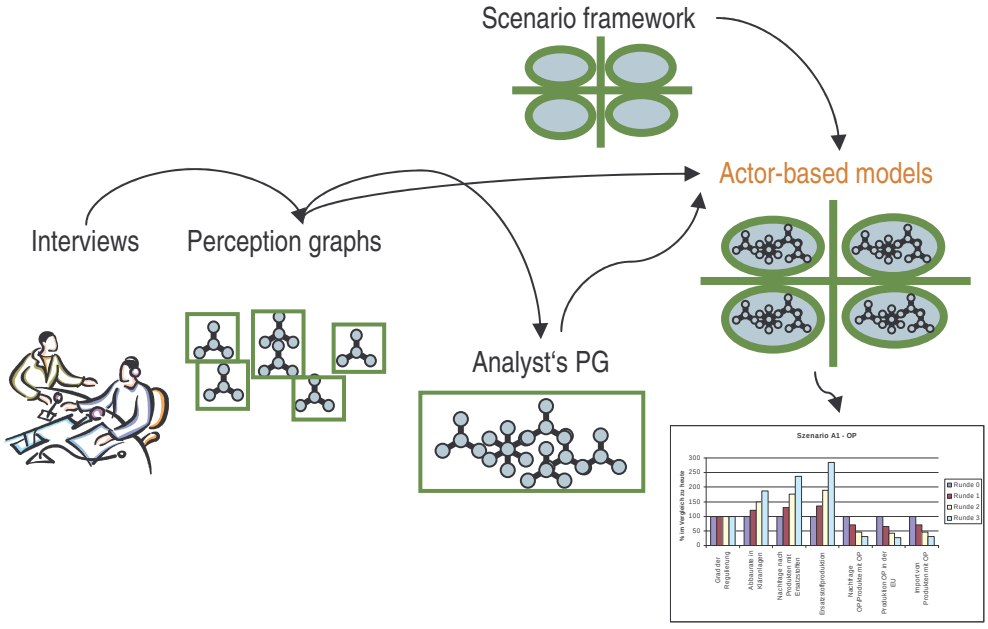


Fig. 1 Overview of deriving scenarios of the development of factors (e.g. production volume of a chemical substance) by actor-based modeling.

4. Applications

We applied actor modeling and actor-based modeling to support the identification of strategies for sustainable development in two problem fields where knowledge is rather poor and the cooperation of a number of societal actors with differing problem perceptions, values and goals is required:

- mobile organic xenobiotics in surface waters (www.intafere.de) [Döll and Döll, 2006, 2008]
- pharmaceuticals in drinking water (www.start-project.de)

Both projects included stakeholder participation over 2-3 years, i.e. stakeholders were not only interviewed but also participated in workshops. While in "intafere", 9 structured interviews with representatives of 9 actors were done, in "start", 1-3 representatives of 7 actors developed 7 perception graphs of their respective actor (on paper). The PGs as constructed with DANA were verified by the representatives, partly in a second meeting.

In "intafere", explorative scenarios of future production of specific organic xenobiotics (e.g. the flame retardant TCPP) in the EU, their import into the EU, the production of chemical substitutes and their average removal rate in waste water treatment plants was

modeled. The scenarios varied with respect to the degree of 1) regulation of chemical substances by the EU and 2) the sustainability-oriented behavior of the consumers. The actors whose perception graphs influenced the scenarios were the producers of the chemical of interest, the German environmental protection agency, the regional water authorities, water suppliers and environmental NGOs. Within the integrated assessment done in "intafere", the computed future developments of the four factors were then used as input to models that computed the emissions to a small drainage basin south of Frankfurt as well as the resulting concentrations and ecological impacts in the local surface water system.

In "start", following the precautionary principle applicable under circumstances of unknown risk, an integrated strategy for decreasing emissions of pharmaceuticals into the water cycle were to be developed. Here, the actor modeling served to identify which measures can be combined successfully, e.g. with the support of the actors including pharmaceutical industry, waste water treatment operators, drinking water suppliers, physicians, pharmacists, health insurers and the Ministry of Health and subordinated authorities. A selected result was that most actors support the development of the environmentally-friendly pharmaceuticals. The pharmaceutical industry, however, does not think that its goals, which include that release of pharmaceuticals into the environment should not increase, can be reached by the explicit development of environmentally-friendly pharmaceuticals but will be influenced positively by the development of more highly active pharmaceutical agents.

5. Conclusions

Actor modeling makes the perceptions of the actors in a problem field more transparent to the actors themselves and to the other actors, and it increases the system knowledge of actors and scientists. Actor-based modeling can be used to estimate how actors might act in certain scenarios of the future, and to compute, in a semi-quantitative way, the relative changes of factors that result from these actions. Both methods are best implemented in participative processes which aim at the development of sustainable development strategies by stakeholders. In the framework of integrated assessment and modeling, these changes of factors can then serve as input to physical models. Limitations of these new methods include the assumption that actors are homogeneous and that actor perceptions do not change in the future. It is not possible to obtain statistical representativeness, and it is difficult to balance between complexity and triviality in the PGs. Besides, the semi-quantitative computations might lead to spurious results. Model validation should certainly be performed together with the actors. In the future, we plan to perform sensitivity analyses, and we want to explore how learning of actors can be included.

ACKNOWLEDGEMENTS

The authors wish to thank Pieter Bots, TU Delft, The Netherlands, for modifying the software DANA according to our scientific needs.

REFERENCE

Bots, P. W. G., van Twist, M. J. W., van Duin, J. H. R., Designing a power tool for policy analysts: Dynamic Actor Network Analysis. In Nunamaker, F. M., Sprague, R. H. (eds.): *Proceedings HICSS 33*, IEEE Press, 1999.

Bots, P. W. G., Analysis of Multi-Actor Policy Contexts Using Perception Graphs, *IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'07),* 160-167, 2007.

Döll, C., Döll, P., Akteursmodellierung im Rahmen des Projektes "Integrierte Analyse von mobilen, organischen Fremdstoffen in Fließgewässern". In Wittmann, J., Müller, M. (eds.): *Simulation in Umwelt- und Geowissenschaften*. Workshop Leipzig 2006. Shaker Verlag, Aachen (ISBN 3-8322-5132-4), 109-118, 2006.

Döll, C., Döll, P., Modellierung der Problemwahrnehmungen und Handlungen von Akteuren im Problemfeld "Mobile organische Fremdstoffe in Gewässern". In: Koch, A., Mandel, P. (eds.): *Modellierung und Simulation komplexer geographischer Systeme*. Salzburger Geographische Mitteilungen, Salzburg, 2008 (in press).

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Knowledge elicitation games on stakeholder decision-making: do we get closer to agents' behaviour?

D.Haase[a]

[a] *UFZ – Centre for Environmental Research, Dept. of Computational Landscape Ecology, Permoserstr. 15, D-04318 Leipzig, Germany, (dagmar.haase@ufz.de)*

Keywords: *knowledge elicitation games; decision-making; flood adaptation; housing choices; decision tree interpretation.*

Identifying decision criteria and getting closer to human behaviour is fundamental for the understanding of decision-making processes and actions that shape our landscape and its environmental resources (Bharwani, 2006). Often we are confronted with multiple reasons for land use decisions of multiple actors that have a stake on one resource (e.g. land, water resources, soil fertility etc.).

Participatory methods such as Knowledge Elicitation Tools (KnETs) represent a new and reproducible way to formalise this knowledge using computational techniques and, what is more, to implement scenario techniques within the interviews. KnETs can be understood as an amplified methodology of classical social science empirical tools such as interviews and questionnaire surveys. It produces input data for the logic of agent-based models (ABM), decision trees or decision ranking. KnETs link qualitative and quantitative representations of stakeholder knowledge. The fusion of both is realised in form of an iterative process that incorporates a set-up phase of the game conducting interviews, a formalisation phase of the game, an empirical data collection phase of "gaming" with respective respondent groups of interest and a resulting decision tree creation and interpretation phase. Finally, another game round is conducted with a non-involved respondents' group to verify the results of the models and to assess our gained knowledge on decision making (Bharwani, 2006).

In this paper I present two applications of the KnETs games methodology for firstly, getting insides in criteria and thresholds of decision making by municipal representatives pro or contra soft mitigation paths in flood risk management in the Ukrainian Tisza valley and, secondly, decision making of households to 'reurbanise' and to settle at recycled land within the inner city (success test of the realisation of a so-called 'program of urban villas'). The resulting trees shed light on what knowledge is used for decision making and how different criteria are weighted in these choices.

ACKNOWLEDGEMENTS

I would like to thank Sukaina Bharwani for introducing the KnETs methodology to our UFZ team, our Newater colleagues Svetlana Kuptsova and Yulia Demenko from the water direction in Ushgorod (Ukraine) for conducting the games upon flood risk adaptation in the Zacarpathian region as well as Anna Kunath for collaborating in the urban games sessions.

REFERENCES

Bharwani, S., Understanding complex behaviour and decision making using ethnographic knowledge elicitation games (KnETs). *Social Science Computer Review* 24(1), 78-105, 2006.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Waterlogging in Upper Tisza and the driving forces behind

J. Mysiak[a] **and Z. Nagy**[b]

[a] *Fondazione Eni Enrico Mattei, Castelo 5252, Venice 30122, Italy (jaroslav.mysiak@feem.it)*
[b] *Corvinus University of Budapest, Department of Soil Science and Water Management, Budapest, Hungary, (zsuzsanna.nagy14@uni-corvinus.hu)*

Keywords: Rural policies; conceptual modelling; waterlogging; policy analysis.

Agricultural and environmental policies can have positive, mutually reinforcing effects. The recent reform to the EU Common Agricultural Policy, for example, introduced compliance mechanisms and incentives from which both, environment and rural communities, can benefit. In some cases however, the policies respond to different, incompatible concerns, producing situations in which positive attainments are reciprocally obstructed.

This paper analyses the synergies and barriers encountered when tackling the waterlogging issue in Upper Tisza River Basin, North East Hungary. The frequent floods and ensuing waterlogging of large areas, former floodplains turned into crop fields, are serious issues and their importance increases with the gloom prospects of climate change. The floods in 2001, 2005 and 2006 affected several hundred thousands hectares and caused damages costing billions. In the past, the Tisza river had been streamlined to prevent floods and to gain new agriculture land in landscapes where wetlands and marshes dominated before. This has further amplified the vulnerability to flood and draughts, the latter being the consequence of low water retention to bridge the meagre summer periods.

We analysed existing legislation, recent policies and subsidy schemes with potential effects on farmers decisions, eventually affecting the damages from waterlogging and floods. Similar to other studies employing qualitative, conceptual models to examine the behaviour of the systems before getting hold of quantitative data, this paper analyses obligations and incentives encouraging farmers to employ management practices which ultimately reduce their vulnerability to flood and waterlogging damages.

The policies we look into include compulsory rules and financial or material incentives realized in the context of the European Agricultural and Rural Development Operational Programmes; national programmes for rural development, environmental protection and spatial development; diverse compensation and insurance schemes, and legislations referring to land use and water management in general. Particular attention is paid to single area payments (SAP) and the cross-compliance obligations the farmers have to meet, and the agri-environmental schemes reducing the vulnerability to floods and waterlogging (e.g. afforestation, planting of autochthonous plants more resistant to high water tables, flood plain restoration and establishments of environmental set-aside areas). We consider policies currently in practice, as well as measures realised in the past and newly planned interventions. When possible, the policies are analysed quantitatively in terms of financial resources allocated, number of actors who adhered to the policy schemes, and the tangible impacts. When such information is not accessible, the policies are delved into qualitatively, in terms of compliance criteria and revealed preferences of the intended beneficiaries. For this scope a series of semi-structured interviews had been conducted, involving farmers, representatives of agricultural water boars, water authority, nature protection agencies and others relevant actors.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Sustainability Evaluation for Shellfish Production in Gamak Bay based on the Systems Ecology

H.T. Oh[a], W.C. Lee[a], S.M. Lee[b], R.H. Jung, S.J. Hong[a]

[a] National Fisheries R&D Institute, Marine Environment Management Team
(ohhyuntaik@gmail.com)
[b] National Pukyung Univ. Ecological Engineering Department
(leesm@pknu.ac.kr)

Keywords: EMERGY Analysis, Gamak (Kamak) Bay, Shellfish Aquaculture, Marine Management Policy, System Ecology

This research outlines a method for sustainability evaluation for shellfish aquaculture production in Gamak Bay for based on Emergy evaluation of the environmental resources. For a better understanding of environmental factor influencing oyster production and the management of oyster stocks, it is important to understand and assess the real value of environmental sources such as solar energy, river, tide, wave, wind, etc. In this research, EMERGY(Odum, 1996; Ulgiati and Brown, 2007) flows from environment were 76% for shellfish aquaculture in Gamak Bay. EMERGY yield ratio, Environmental Loading Ratio, and Sustainability Index (Brown and Ulgiati, 1997; Tilley and Vito, 2006) are 4.2, 0.3, 13.8 respectively.

Using the Emergy evaluation data, the predicted maximum shellfish aquaculture production in Gamak Bay and FDA (Food and Drug Administration, U.S.) designated area in Gamak Bay were 10,845 ton/yr, 7,548 ton/yr, respectively. Actually predicted shellfish production were about 1.3 times more than produced shellfish production in 2005. Therefore, the carrying capacity of the Gamak Bay may be estimated to be 1.3 times as high as the present oyster production.

This research was performed to develop a method of strategic environmental assessment on the operation of sewage treatment plant and reuse of seeding areas by using the environmental accounting based on emergy evaluation (Kang and Park, 2002; Valyi and Ortega, 2004). The result was applied to marine environment policy in order to evaluate the real wealth of the regional environment and economy for present phase, and developed phase from scenario. Using the result of comparison of Emergy indices between present system and new system (Sewage Treatment Plant, Seeding), the new 1(STP), new 2(Seeding), and new 3(STP & Seeding) cost-benefit output show 1.88, 0.78, and 1.29, respectively.

In this study, an energy systems model was built to simulate the variation of sustainability of the oyster aquaculture. The results of simulation based on calibration data in 2005 show that oyster production yield slowly slightly increase and money and asset show increase to the steady state because of the law of supply and demand. The results of simulation based on the variation of inflation rate show that the more increase the oil price, the more decrease the oyster production and asset after using the excess environmental resources around 10 years.

ACKNOWLEDGEMENTS

The authors wish to thank the National Fisheries R&D Institute for the financial support of the project “Assessment of Carrying Capacity in Coastal Aquaculture Sites”.

REFERENCES

Brown M. T., S. Ulgiati, 1997, Emergy based indices and ratios to evaluate sustainabnility : monitoring economics and technology toward environmentally sound innovation, Ecological Engineering, 9, pp. 51 ~ 69.

Kang, Daeseok, and Seok Soon Park, 2002, Emergy Evaluation Perspectives of a Multipurpose Dam Proposal in Korea, Journal of Environmental Management, Vol 66, pp. 293~306

Odum, H.T., 1996, Environmental Accounting, Emergy and Environmental Decision Making. John Wiley & Sons, New York, pp. 294~303

Odum, H.T., and J.E. Arding, 1991. Emergy analysis of shrimp mariculture in Ecuador. Univ. Florida, pp. 37~44

Tilley, David and Vito Comar, 2006, Emergy-based Simulation to Assesses Brazil's long term carrying capacity : Environment, Electricity and Population, Population & Environment, Vol 27(3), pp. 307~326

Ulgiati, S., and M. T. Brown, 2007, Emergy and Ecosystem Complexity, Communications in Nonlinear Science and Numerical Simulation, DOI: 10. 1016

Valyi, R., and E. Ortega, 2004, Emergy simulator, an open source simulation platform dedicated to systems ecology and Emergy studies, Sourceforge website (http://emsim.sourceforge.net)

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Modelers and stakeholders: who is the boss?

A. Voinov[a,b] **and F. Bousquet**[c]

[a] *Chesapeake Research Consortium Community Modeling Program &*
[b] *Johns Hopkins University Dept. of Geography and Environmental Engineering*
645 Contees Wharf Road, P.O. Box 28, Edgewater, MD 21037, United States
(alexey.voinov@uvm.edu)
[b] *CIRAD GREEN Research Unit, Montpellier, France (francois.bousquet@cirad.fr)*

Computer modeling has been regarded as an efficient way of solving problems. Yet it has a history of failure when applied to decision making. Models have not necessarily resulted in better decisions. There is a clear disjoint between analytical quantitative analysis and decision making. How do we make models more appealing and convincing for the decision makers?

One possible approach is to open up the modeling process and engage the model users, the stakeholders in the modeling process. This is known as participatory or companion modeling, or shared vision planning, or mediated modeling. This increases the transparency of the model and brings more trust to the results. In most cases we would claim that the model is a product of the stakeholder process and is designed to suit the needs of the stakeholders and the goals of the project. However in reality much of the modeling process is predetermined by the skills and preferences of the modelers who come with their set of tools and drive the process.

We review a series of modeling projects that involve stakeholders and classify the various tools that they have used. We then try to evaluate to what extent the modeling approaches may have predetermined the outcomes. We also argue that the stakeholder process could benefit from a community effort that would make models and data more readily available and interoperable to make it easier to switch from one modeling paradigms to another and involve models of varying complextiy in the decision making process.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Some Thoughts on Success in Applying Models to Water Resource Systems

Giovanni M. Sechi [a] and Andrea Sulis [a]
[a] *Dept. of Land Engineering, University of Cagliari, 09123 Cagliari, Italy*
(asulis@unica.it)

***Keywords**:* Models, Complex Water Systems, Evaluation Process.

In the attempts to apply models to complex water systems we must consider the interests and opinion of numerous stakeholders. The use of these models does not eliminate the risk of making mistakes but can better inform all these different stakeholders (Loucks and van Beek, 2005). The success (or failure) of these attempts is often hard to judge. Many possible frameworks have been presented to judge success.

Following Goeller (1988), in this paper we suggest that the effectiveness of models in practice can be judged by three measures, each ones applies to a different time period (Figure 1). The *analysis success* reflects how the analysis was performed and presented to the stakeholders. The analysts must take care of client satisfaction but a success only based on this measure will be transitional. In our experience the client may be not satisfied because he is learning something he does not want to accept. The *application success* is concerned with how the model was used in the decision making process and by whom. Also, the extent to which the information from the model application influences the decision making process is a good indicator of success in model application. As the identification of problems is a key aspect of their solution, the application success can support the framing of problems to better identify problems worthy to be solved (Hermans, 2005). The *outcome success* gives information on how the use of analytic results from the model affects the system planning and management and if information resulting from the model improve the problem.

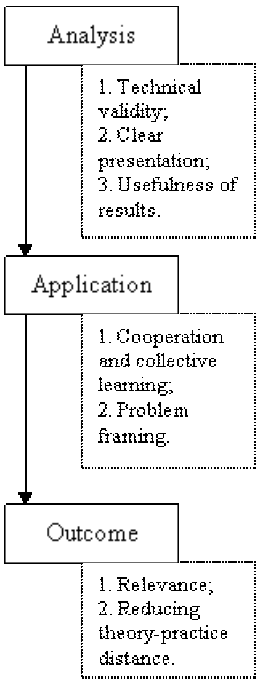


Figure 1. Success Evaluation Steps

All of us must judge the success in a "smart way". Problems and measures can change over time. Monitoring the performance of any decision influenced by the models and working closely with the client is a central aspect in any modelers' activities.

REFERENCES

Goeller, B.F., A framework for evaluating success in systems analysis, *Report P-7454*, The Rand Corp., Santa Monica, California, 1988.

Hermans, L.M., *Actor analysis for water resources management*, Eburon Publishers, Delft, 2005.

Loucks, D.P. and van Beek, E., *Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications*, UNESCO Press, Paris, 2005.

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

From INGIS to NAMOS – Operationalisation and contextualisation of sustainability at the local level

Developing and improving sustainability-related indicator systems as decision and support tools in urban theory and practice

D.Rink[a], A. Haase[a], S. Kabisch[a], D. Haase[b], G. Hartmuth[c]

[a]UFZ – Centre for Environmental Research, Dept. of Urban and Environmental Sociology, Permoserstr. 15, D-04318 Leipzig, Germany, (dieter.rink@ufz.de)
[b]UFZ – Centre for Environmental Research, Dept. of Computational Landscape Ecology, Permoserstr. 15, D-04318 Leipzig, Germany
[c]UFZ – Centre for Environmental Research, Dept. of Environmental Microbiology, Permoserstr. 15, D-04318 Leipzig, Germany

Keywords: *sustainability, monitoring; operationalisation; indicators; Leipzig.*

Due to its fuzziness the model of sustainable development has to be particularised and contextualised before any positioning with respect to this yardstick. The spread of the concept of sustainable development has not been thwarted by its extreme vagueness or its current lack of generally binding operationalisation. On the contrary: its conceptual ambiguity has in all likelihood significantly boosted its appeal by enabling stakeholders to focus on their favourite elements within a process of discourse dominated by the struggle for the power of definition.

Based on a newly put forward conceptualisation of sustainability the realisation of both postulates is shown within the development of a local indicator system for two German cities (IGNIS). It serves as a decision and support tool for urban planning or practice. It could be used as an information and communication medium on the local role of sustainability which is geared to the different needs of the addressees consulting it (e.g. local politicians, local authority personnel, the general public, Local Agenda 21 officers) and the aims in mind (e.g. reporting, transparency, education and training, comparison with other towns and cities). In addition, the information system could acquire a guidance function by supporting local politicians and authorities in the early identification and analysis of problem areas, the definition of aims and the selection of suitable measures

In the project, initially, local problem areas were identified in a bottom up approach by local authority personnel. These problem areas were contrasted with a set of so called sustainability rules which had been systematically derived from the basic sustainability norm, stating minimum requirements for sustainable development. At the interface of local problem areas and sustainability rules indicators were determined. These are supposed to

provide information on whether the city is over time becoming closer to or farther removed from the respective sustainability goals in its problem areas.

The project was set up in cooperation between the Helmholtz Centre for Environmental research and the cities of Leipzig and Halle (Germany) which was lead and guided by scientists of the first-mentioned institution in 2002-2006. In the result both cities have implemented the indicator system in their work. In both cities special groups with representatives from different local authorities were founded, which reports one time a year about the efforts towards sustainability.

Given this background, the paper presents in a first part the INGIS decision and support tool and discusses successes and challenges of the cooperation between scientists and practitioners while setting it up. In a second part, it expands on how the tool is used and developed further by the tow municipalities in urban practice and about its applicability in other European cities. Last but not least, the paper informs in a third part about a follow-up project (started in 2006; at the moment under work) that develops an indicator set and monitoring system for the sustainable development of shrinking cities (NAMOS), also starting with the local case of Leipzig and drawing on the base of the IGNIS system. NAMOS is a specification of IGNIS and tries to elaborate the sustainability concept to the specific problems of shrinking cities like the shrinkage of population, vacancies of flats and abandonment. It will develop a tool to steer the process of shrinkage.

REFERENCES

Rink, D.; Hartmuth, G.; Huber, K. (Eds.) (2005): Raum für Nachhaltigkeit. Zur Kontextualisierung des Leitbilds, Berlin.

Hartmuth, G.; Rink, D.; Huber, K (2007): Operationalisation and contextualisation of sustainablility at the local level: Stages in the development of a sustainability indicator system, in: Sustainable Development (submitted).

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

ICT Tools For Environmental And Human Risk Assessment

I. Holoubek[a], **J. Hřebíček**[b] and **L. Dušek**[b]

[a] *Research Centre for Environmental Chemistry and Ecotoxicology, Masaryk University, Kamenice 126/3, 625 00 Brno, Czech Republic (holoubek@recetox.muni.cz)*

[b] *Institute of Biostatistics and Analyses, Masaryk University, Kamenice 126/3, 625 00 Brno, Czech Republic (hrebicek&dusek@iba.muni.cz)*

Abstract: Environmental (Ecological) and Human Risk Assessment (*EcoHRA*) is a complex process of assessing the influence of human activities on ecosystem duality and vice versa. The endpoint of *EcoHRA* is a prospective or retrospective assessment of stress factors influence (chemical contaminants, anthropogenic interventions or natural disasters) on ecosystems and their parts. Monitoring and collecting of available information as the source data for *EcoHRA* is a key part of the process for the assessor. The availability and quality of input data determines the complexity and feasibility of the process. In the problem description phase it is possible to use the public available high-quality data from on-line web databases and web services, e.g. toxicological properties of chemicals, GIS characteristics and maps, etc. Fundamentally, the problem relates to the quality and completeness of the data inputs and to the capability of assessor to aggregate key information according to methodical requirements and given situation. Recent development of information and communication technologies provides very powerful platforms for effective processing of multiple data sources. Particularly, web searches, database systems and data mining tools oriented on key environmental components and their descriptors, regionally-specific data aggregation, mapping of exposure and segmentation of the region of interest using Geographic Information Systems technology, automated processing of laboratory tests, namely dose-response curves using eco-toxicologically relevant models, algorithms and statistical packages and probabilistic estimates of risk level and associated uncertainties and their reflection in standardization of used environmental information systems by EEA, EPA, OECD and UNEP.

Keywords: Environmental Data; Environmental Information; Ecological Risk Assessment, Environmental and Human Risk Assessment; Monitoring; ICT; Environmental Information Systems; SEIS.

1. INTRODUCTION

Environmental monitoring is the very important part of *EcoHRA* framework. Environment and environmental impacts are monitored and connected with appropriate data that are collected by responsible organizations like Eurostat, European Environment Agency (EEA), U.S. Environmental Protection Agency (EPA), Organization for Economic Cooperation and Development (OECD), and United Nations Environment Programme (UNEP) in different ways e.g. using sensors, satellites, monitoring facilities and devices, people, etc.

For example, the United Nations (UN) System *Earthwatch* {http://earthwatch.unep.net/) is a broad UN initiative to coordinate, harmonize and catalyze environmental observation activities among all UN agencies for integrated assessment purposes. Through *Earthwatch*,

UN agencies work together on global environmental issues, by exchanging and sharing environmental data and information.

The North American Node of UNEP *GRID*, located at the USGS EROS Data Center (http://na.unep.net/), is in the forefront of applying information and communication technologies (ICT) such as remote sensing, Geographic Information Systems (GIS) and web mapping to address the relationships between the environment and human populations.

The Global Monitoring for Environment and Security (GMES), (http://www.gmes.info/), represents a concerted effort to bring data and information providers together with users, so they can better understand each other and make environmental and security-related information available to the people who need it through enhanced or new services.

The International Register of Potentially Toxic Chemicals, or the UNEP Chemicals (IRPTC) helps the world community make better use of existing global resources and to give developing countries the information base to manage chemicals effectively. IRPTC collects information on hazardous chemicals through the Member countries Network. Based on the collected information, data profiles are prepared by contributing network partners, consultants and IRPTC staff.

2. ENVIRONMENTAL MONITORING

Environmental monitoring is shortly analysed in this section. It is the basis of environmental data collection, environmental assessment, reporting and also environmental research and the basis of understanding of environmental problems and trends. Environmental monitoring is therefore a powerful tool for supporting *EcoHRA*, decision-making, enforcing policy decisions, and for assessing compliance with policy regulations and objectives [Gilbert, 1987], [Holoubek, 2004].

Over the last decade, ICTs and new indicators have been developed in many countries to measure environmental impacts and performance, support *EcoHRA*, better integrate environmental and economic decision-making and better communicate environmental information to decision-makers and the public.

Monitoring institutions of many countries have received important autonomy to help ensure the independence and quality of their data, Hřebíček et al [2006], Olson [2003]. The integration of national information systems with international networks EIONET - European Environment Information and Observation Network (http://www.eionet.europa.eu/) and UNEP has been improved, including those co-ordinated by the governing bodies of international environmental conventions like UNEP, OECD or EEA, EPA, etc.

Here something could be added on the value of optimisation and co-ordination of monitoring networks. The Group on Earth Observations (or GEO) (http://earthobservations.org/) is coordinating international efforts to build a GEOSS: Global Earth Observation System of Systems. This emerging public infrastructure is interconnecting a diverse and growing array of instruments and systems for monitoring and forecasting changes in the global environment. GEO's current Members include 72 countries and the European Commission. Important role play NASA - National Aeronautics and Space Administration (http://www.nasa.gov/), ESA - European Space Agency (http://www.esa.int/), Russia and China with their satellite monitoring systems.

Despite this progress, a number of important weaknesses remain. Some current environmental monitoring systems in many countries do not meet priority demands. Some important environmental areas in *EcoHRA*, such as hazardous waste, heavy metals and other toxic substances, particulate matter, acidification, indoor air quality, drinking water, groundwater and wildlife outside protected areas, are not properly monitored in many countries. Inventories are lacking in several countries of waste of high potential hazard, which were (or continue to be) dumped on landfill sites, especially in rural areas. In a number of countries, biological monitoring systems are too cumbersome and expensive to

manage. Monitoring practices to monitor environmental effects of armed conflicts in those countries where these conflicts occurred are practically nonexistent.

There is, furthermore, a contrast between the large volume of data produced by ICTs and the difficulty in using these data in *EcoHRA* to support decision making. On-line data banks of time-series data are poorly developed in a number of European countries and therefore in the EEA. This handicaps, in particular, analysis of cause-effect relationships, Hejč et al [2006].

According to the Communication COM(2008) 46 final, the SEIS (Shared Environmental Information System) has been developed and will be able to provide the knowledge base required to design, implement and evaluate the environmental and other policies. The SEIS will enable up-to-date quality data and information covering all elements of the DPSIR (Driving forces, Pressures, State of the environment, Impacts and Responses) framework of EcoHRA.

The general objectives of the SEIS are:

- to secure a clear political agreement in EU around a set of principles on which the SEIS is to be based;
- to continue rationalising the *knowledge base* through the assessment and streamlining of existing reporting requirements within environmental legislation while implementing information and communication technology solutions for electronic reporting;
- to establish and implement data and information sharing agreements in addition to an efficient ICT infrastructure to facilitate the discovery, assessment, access and sharing of environment-related data and information;
- to reinforce and, where necessary, establish monitoring infrastructures and surveys for the collection and archiving of *fit-for-purpose* environment-related data that are cost effective and flexible but can be sustained over the long term.

Most countries have increased the amount of environmental data and information provided to the public to follow Aarhus convention and Directive 2003/4/EC. Most of them now produce annual state of the environment reports that contain data from main sources. An increasing number provide this information through the web data basis. However, not everywhere is free and easy access to environmental information ensured for the general public, Hejč et al [2007].

Environmental authorities lack both monitoring experts and technical equipment in many countries. Industrial facilities also lack financial resources and are therefore not able to equip the sources of pollution with appropriate monitoring devices. Both the equipment that is available and sampling methods have become obsolete. As a result, routine monitoring activities have been handicapped or even discontinued altogether. The reliability and accuracy of air and water quality information is often questionable in many respects. Thus, it is impossible to fully evaluate the current environmental situation in these countries. Furthermore, ongoing national monitoring systems could hardly be integrated in international (global or regional) programmes of UNEP (GEO).

Another aspect to be covered is the dissemination of knowledge: Across the various European Countries, monitoring networks (GMES, EIONET) and monitoring strategies are at a different level of development. Fortunately, the development of SEIS will improve this level. Its co-ordinated action is, amongst others, aimed at harmonising efforts across Europe and exchange of knowledge with regard to:

- Development of monitoring strategies, optimising the cost/benefit ratio and taking transboundary activities into account. In case of the European Water Framework Directive (WFD) the latter is for instance done on the basis of the pre-defined basin-approach.
- Integration of modelling, unifying GIS (INSPIRE directive), and monitoring.

- Truly representative and comparable (amongst others because of harmonisation of methodologies) data-sets.

3. ECOLOGICAL AND HUMAN RISK ASSESSMENT

The ecological and human risk assessment (is a rapidly evolving discipline that is quickly being incorporated into the daily routines of industry and government. The *EcoHRA* framework has recently been completed or is under development in numerous countries of Europe [EU, 2003], in USA [EPA, 2001], [EPA, 2003], and Canada. In present time, many specific approaches and models exist in numerous sub-disciplines of *EcoHRA*, [Dolislager et al, 2007], such as chemical-property and fate estimation, toxicity, biological uptake, and population effects with limited co-ordination across these sub-disciplines [Holoubek, 2004].

The basic components of the *EcoHRA* framework are summarised in the Figure1.

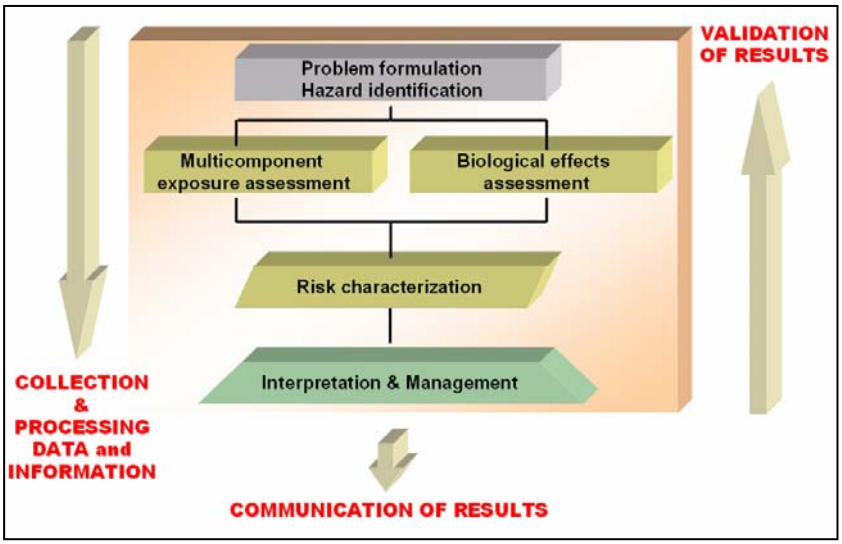


Figure 1: The key methodological step of the *EcoHRA* framework [Holoubek et al, 2004]

The *EcoHRA* framework shows the trends of global harmonization of *EcoHRA* and development of common methods, data, ICT and criteria and pooled resources to obtain more cost-effective, consistent, and flexible systems with reduced duplication of effort and incompatibilities, [Stewart et al, 2007].

The basic components of traditional ecological and human risk assessment are [Holoubek et al, 2001]:

- Stressors - released into the environment;
- Receptors - living in and using that environment;
- Response - of receptor to the stressors.

Measurements of exposure and effects represent the interactions between the components. Conventional *EcoHRA* depends on measurements of exposure and effects to calculate an estimate of risk. The regional concept of ecological and human risk assessment (*ReEcoHRA*) requires additional consideration of scale, complexity of the structure, and regional spatial components [Holoubek et al, 2001]: sources - release stressors; habitats - where the receptor reside; impacts - to the assessment endpoints. The *ReEcoHRA* uses multiple stressor assessment, where the number of possible interactions increases

combinatorially. The stressors are derived from diverse sources; receptors are often derived from diverse sources, Figure 2. Receptors are often associated with a variety of habitats. One impact can lead to additional impacts.

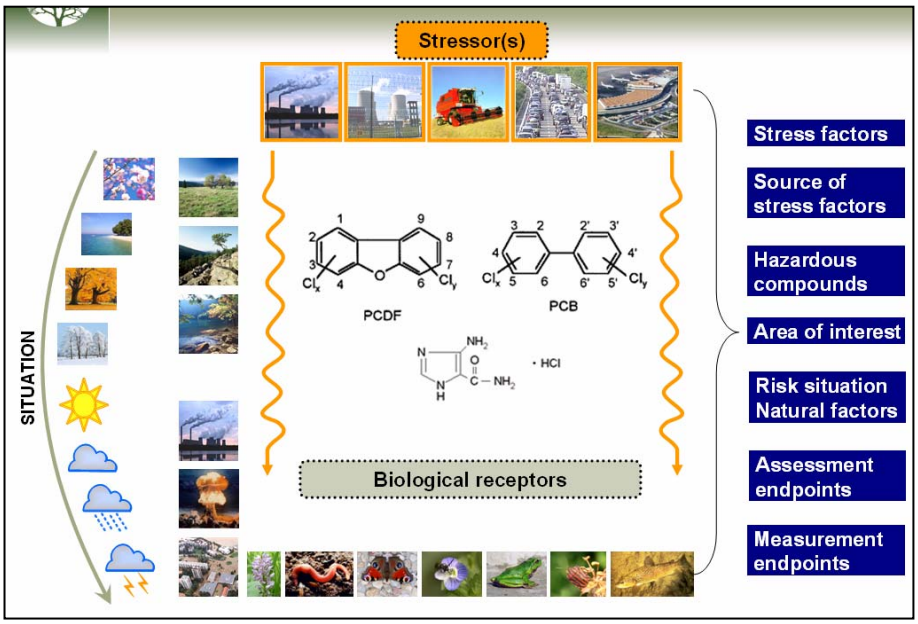


Figure 2: The basic components of the *ReEcoHRA* framework [Holoubek et al, 2004]

The basic approaches of *ReEcoRA* framework are [Holoubek et al, 2001]:

- To evaluate the risk components at different locations in the region; Risk characterization defines the likelihood that humans or wildlife will be exposed to hazardous concentrations. Thus, risk characterization describes the relationship between exposure and toxicity. Risk assessors identify species likely to be exposed, the probability of such exposure occurring, and effects that might be expected.
- Rank the importance of these locations;
- Combine this information to predict the relative risk among these areas;
- The number of possible combinations that can results from this approach depends on the number of categories that are identified for each risk component.
- Each identified combination establishes a possible pathway to a hazard.

Assessing and characterizing risk to ecological systems, including a myriad of non-target aquatic and terrestrial organisms as well as surface and ground water, is a much younger and more complex science than that of human health risk consideration. Ecological risk assessment considers a greater range of complex issues and covers more species than does human health risk assessment: fish, aquatic invertebrates, aquatic and terrestrial plants, non-target insects, birds, wild mammals, reptiles, and amphibians.

The first step of *EcoHRA* on the local or regional level - problem formulation, see Figure 3, has these basic steps:

- Identify local specific or regional characterisation of a study site or region;
- Qualitative evaluate of contaminant release, migration and fate of stressors (pollutants);
- Identify contaminants with potential ecological effects; receptors; exposure pathways; known effects;
- Select endpoints of concern;

- Specify objectives and scopes.

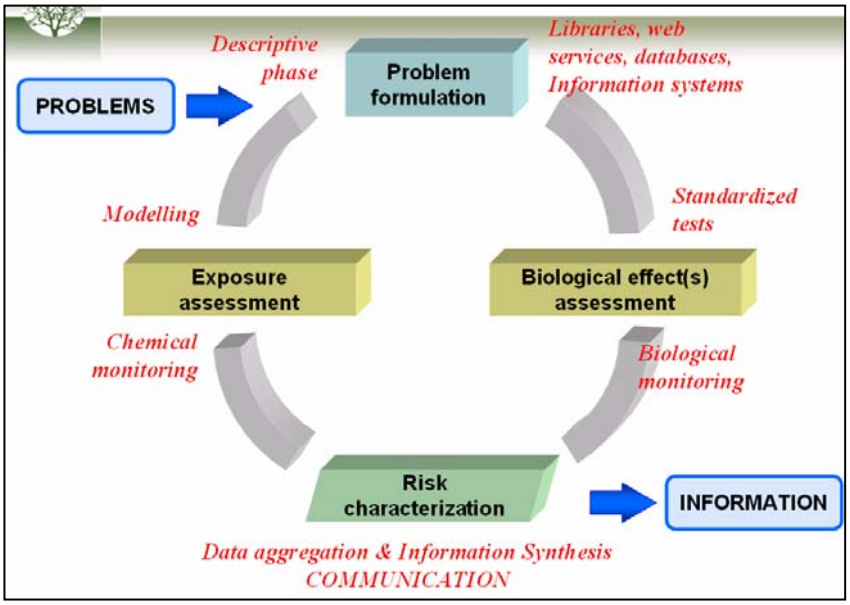


Figure 3: The *EcoHRA* framework and related methods and information sources.

The multicomponents exposure and human risk assessment mainly are mainly focused on quantifying of releases, migration and fate, characterisation of receptors and measurement or estimation of exposure point concentration.

Risk characterisation as a final step of whole procedure contents current adverse effects, future adverse effects, uncertainty analysis and ecological and human significance.

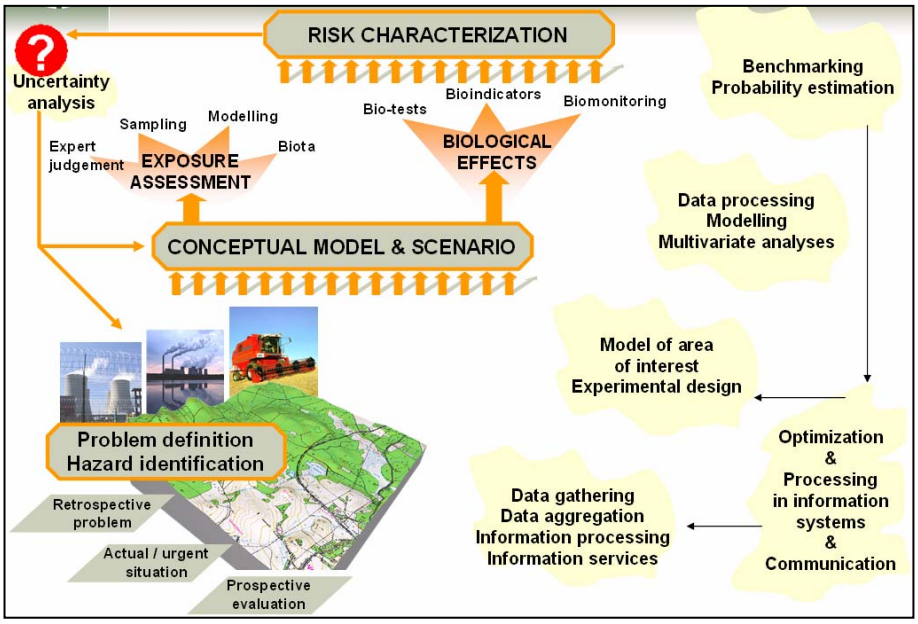


Figure 4: The *EcoHRA* risk characterisation and related used tools and information sources.

4. REQUESTS ON COMPONENTS OF ICT TOOLS

All ICT tools (databases, information systems, environmental models) have to be accessible through standard web browsers (MS Explorer, Mozzila, Firefox, Opera, etc) like for example are web information systems RAIS [Dolislager et al, 2007] or SADA [Stewart et al, 2007]. It is expected that they will have friendly Graphical User Interface (GUI). Their query formats will have to be the easy-to-use with GUI for selection of input or result data that could be consistent across the given ICT tool. The efficient help system with tutorial could be available for the user through the selection and information retrieval process.

Further, we shortly describe the basic components of ICT tools which could be available for *EcoHRA*.

4.1 Databases

This ICT tools could contain fully-searchable databases, for example:

- **Toxicity Values.** This database could be a fully-searchable database of chemical (organics and inorganics) and radionuclide human health toxicological values. This is needed to perform risk evaluations and assessments in accordance with the standard methods of the Risk Assessment e.g. [EPA, 2003]. It is possible to download these values from EPA's Integrated Risk Information System (IRIS), or incorporates data from another sources, e.g the Health Effects Assessment Summary Tables (HEAST), the Agency for Toxic Substances and Disease Registry (ATSDR) Minimal Risk Levels (MRLs), World Health Organization (WHO) values, and California Environmental Protection Agency (CalEPA), etc.
- **Toxicity Profiles.** EPA and EU legislation provide many excellent toxicity profiles that deal with the most important chemicals on a national and international scale [Patterson et al. 2002].
- **Chemical-specific factors.** During both the exposure and the dose-response (toxicity) assessments of a risk assessment, various chemical-specific parameters are needed to determine dose and risk. In support of this need, is appropriate for instance ICT tool, which provides a fully-searchable database of over 20,000 chemical records for 22 parameters of 1635 chemicals and radionuclides combined like the RAIS Chemical-specific Factors [Dolislager et al, 2007].
- **Established Regulatory Limits for Surface Water and Groundwater.** Various national agencies, EEA and EPA use toxicity and risk information to establish drinking water criteria for specific chemicals in environmental media. Verifying that current or future contaminant levels are below or above the criteria is an important task in risk assessment. This ICT tool could provide a fully-searchable database for surface water and groundwater criteria established by EU legislation and state agencies.

4.2 Ecological Benchmarks

There are few web-accessible information sites, which provide both human health and ecological toxicity information. In addition to its human health toxicity values, ICT tools could contain a appropriate standard ecological benchmarks for given chemicals. Screening ecological benchmarks are used to identify chemical concentrations in environmental media (surface water, sediment, and surface soil) that are at or below thresholds for effects to ecological receptors, specifically a range of aquatic organisms, soil invertebrates, and terrestrial plants.

This ICT tool can help risk assessors ensure that contaminants of potential concern that are to be dropped in the risk assessment on the basis of negligible contribution to human health risk are not in fact risk drivers for ecological receptors.

4.3 Human Health Risk Exposure Modelling

Very important could be ICT tool on modelling human health risk exposure. This tool will provide the ability to calculate web-based risk estimates in real time. This tool will allow users to upload small data sets for a first-pass risk assessment. Users should first familiarize themselves with the applicable guidance documents and special cases information provided as links on the page.

4.4 Links to Additional ICT Tools

We can recommend the official web-based EPA Soil Screening Guidance calculation tools for both chemicals (http://rais.ornl.gov/calc_start.shtml) and radioactive isotopes (http://rais.ornl.gov/rad_start.shtml) and the official EPA online calculation tools for radionuclide PRGs (http://epa-prgs.ornl.gov/radionuclides/) and EPA Radionuclide Dose Compliance Concentrations (http://epa-dccs.ornl.gov/).

5. CONCLUSIONS

The principal role of ICT tools in *EcoHRA* cannot be narrowed only to standardized monitoring, gathering and aggregation of primary environmental and related data. All phases of the *EcoHRA* process (hazard identification, exposure assessment, dose-response monitoring, etc) are intrinsically associated with some level of uncertainty and so the final conclusions are based on stochastic analytic methods. At this point we must accentuate key role of GIS technology, multivariate processing of ecological bioindicators and finally very important dose-response modeling and probabilistic characterization of risk. Each of these methods, and many others as well, represent unique field of computational science with its own methodology and progress and developed ITC tools. Notwithstanding the methodical variety, *EcoHRA* must assimilate only verified approaches with sufficiently robust algorithms, suitable to heterogeneous or incomplete data. Reliable computational methodology and appropriate ICT tools bridge the gap between environmental and experimental data and makes the whole *EcoHRA* process as effective as it is possible.

ACKNOWLEDGEMENTS

The authors wish to thank the Ministry of Education of the Czech Republic for the financial support of the project No MSM 0021622412 INCHEMBIOL and the Ministry of Environment of the Czech Republic for the financial support of the projects No SP/4i2/26/07 and SP/1b1/30/07.

REFERENCES

EC, Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information. Official Journal of the European Union. L 41/26. 2003.

EEA, European Environment Agency, <http://www.eea.europa.eu/>, 2008/04/30

EPA, U.S. Environmental Protection Agency, http://www.epa.gov/>, 2008/04/30

EPA, Risk Assessment Guidance for Superfund: Volume I - Human Health Evaluation Manual. (Part A) EPA/540/1-89/002. Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, D.C. 1989.

EPA, Risk Assessment Guidance for Superfund (RAGS): Part D. Volume I -Human Health Evaluation Manual (Part D, Standardized Planning, Reporting and Review of Superfund Risk Assessments). Final December 2001.

EPA, Human health toxicity values in Superfund risk assessments. OSWER Directive 9285.7-53. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, D.C. 2003.

EU, Updated Opinion of the Scientific Steering Committee on Harmonization of Risk Assessment Procedures. European Commission. Health and Consumer Protection Directorate General. 2003. <http://ec.europa.eu/food/fs/sc/ssc/out82_en.html>, 2008/04/30

Eurostat, <http://ec.europa.eu/eurostat>, 2008/04/30

Dolislager, F., Galloway, L., Stewart D., The Risk Assessment Information System (RAIS): Insights from the Evolution of an Online Region-Specific Tool to and International Environmental Risk Assessment Center, In: *6th International Symposium on Environmental Software Systems (ISESS 2007).* Guelph: IFIP WG 5.11, 2007.

Gilbert, R., Statistical Methods for Environmental Pollution Monitoring, John Wiley &Sons, New York, 1987.

Hejč, M., Hlaváček, M., and Hřebíček, J., eGovernment Services in Environment - Automate Data Quality Assessment - Czech Republic Approach, In EnviroInfo 2007. 21st International Conference on Informatics for Environmental Protection. Environmental Informatics and System Research. Volume 2. Workshop and application papers. Warsaw, Poland: Shaker Verlag, Aachen, 159-166, 2007.

Holoubek, I., Monitoring, modelling and information system for persistent organic pollutants. In: Prastacos, P., Cortés, U., de Leon, J. L. D., Murillo, M. (Eds.): e-Environment: Progress and Challenges. *Research on Computing Science*. Vol. 11, 117-134, 2004.

Holoubek, I., Dušek, L., Machala, M., Čupr, P., Bláha, K.: Project IDRIS - Ecological Risk Assessment - regional approaches. Assessment and Management of Environmental Risks: Methods and Applications in Eastern European and Developing Countries. Kluwer Academic Publishers, 283-298, 2001.

Hřebíček, J., Hejč, M., and Holoubek, I., Current Trends in Environmental Modelling with Uncertainty, Proceedings of the iEMSs Third Biennial Meeting "Summit on Environmental Modelling&Software". International Environmental Modelling and Software Society, Burlington, 342-347, 2006.

Jolma A., and J., Bortin, Methods of uncertainty treatment in environmental models, *Environmental Modelling & Software*, 20(8), 979-980, 2005.

OECD, Organization for Economic Cooperation and Development, Environment Directorate, <http://www.oecd.org/department/0,3355,en_2649_33713_1_1_1_1_1,00.html>, 2008/04/30

Olson, J. E., Data Quality: the accuracy dimension. Morgan Kaufmann Publishers, San Francisco. 2003.

Patterson, J., Hakkinen, P.J., Wullenweber, A.E., Human health risk assessment: selected internet and world wide web resources, *Toxicology*, 173, 123-143, 2002..

Pick, T., From Aarhus to INSPIRE: Putting Environmental Information on the Map. In EnviroInfo 2007. 21st International Conference on Informatics for Environmental Protection. Environmental Informatics and System Research. Volume 2. Workshop and application papers. Warsaw, Poland: Shaker Verlag, 239-246, 2007.

Pipino, L., Lee, Y., and Wang, R., Data Quality Assessment. *Communications of the ACM*, 45(4), 2002.

RAIS, The Risk Assessment Information System, <http://rais.ornl.gov/>, 2008/04/30

SEIS, Shared Environment Information System, <http://ec.europa.eu/environment/seis/>, 2008/04/30

Stewart, R. N., Powers, G., and Purucker, T., SADA: A Freeware Decision Support Tool Integrating GIS, Sample design, Spatial Modeling, and Risk Assessment. In: 6th International Symposium on Environmental Software Systems (ISESS 2007). Guelph: IFIP WG 5.11, 2007.

UNEP, United Nations Environment Programme, <http://www.unep.org/>, 2008/04/30

An Intelligent Decision Support System for Environmental Impact Assessment of Industrial/Agricultural Projects

Shady Aly[a], Ivan Vrana[b]
[a]*Assistant Professor, Helwan University - Faculty of Engineering, Cairo, Egypt*
[b]*Professor, Czech University of Life Sciences, Prague, Czech Republic*
vrana@pef.czu.cz, shadyref@yahoo.com

Abstract

An Environmental Impact Assessment (EIA) is an assessment of the likely positive and/or negative influences an industrial project may have on the human health or environment. The assessment can have several dimensions like Air pollution impact, Noise/vibration impact, Water pollution impact, Ecological impact (Geological hazard impact & Endangered species impact), etc. Actually, there is a wide prevalence of situations of environmental impact and human risk assessment. This paper attempts to present a non-conventional solution for treating the complex EIA decision problem. It is typically a complex ill-structured decision making problem that involves: huge quantities of data to manipulate, low quality of data (uncertainty, measurement errors, missing data), different spatial and temporal scales (from seconds to years, from local to global), dynamic and stochastic behavior, and being at the crossroad among many disciplines/domains, and so many qualitative or subjective factors. In addition, the mentioned environmental dimensions often include non-homogenous subsets of input factors or variables. Consequently, in these situations, the decision-making problem becomes so complicated when we need to assess the overall environmental impact of a newly proposed or existing industrial project, specifically, we need to decide whether to adopt or proceed with the project. Such a binary, YES/NO, decision problem requires assessing the influences of all non-homogeneous inputs factors within the given relevant environmental dimensions or views. One way to obtain a reliable solution to such a complex, critical decision problem is to rely on the knowledge, heuristics, and experience of different expertise's corresponding to each environmental dimension. In this paper, we propose a modular intelligent environmental decision support system that makes use of fuzzy logics and appropriate human expertise's in rigorously assessing the impacts of the proposed or existing project on the surrounding environment or human health. The proposed system includes multiple fuzzy expert systems (FESs), each of which contains the homogenous and aggregated knowledge and expertise's relevant to a one relevant environmental dimension. An objective numerical scale is established to assess the decision outputs of each individual system, expressing the degree of bias to "Positive" or "Negative" environmental impacts. Consequently, for a project to be adopted or accepted it must attain a combined or aggregate value in favor of the "Positive" impact outcome; otherwise the project will be rejected.

Integrating multiple intelligent decision support systems is considered particularly useful in obtaining, high quality, more comprehensible, and reliable decision solution. However, the reasons for independency can be related to ease of maintainability, analyzability, context flexibility, decision-making modularity, sensitivity of aggregate knowledge, decision consistency, etc. The problem of integrating multiple FESs involves combining their final individual crisp outputs. Every FES participates in judging the

problem based on a predefined match between problem context and the required specific expertise's. In this research, a consensus-based heuristic will be used as the basic mechanism to combine the outputs of multiple FESs. The heuristic exploits the consensus-relevant information among FESs decision outputs, and is guided by two sets of consensus measures or indicators. The first set is a group of six similarity-based indicators. The second includes voting's and voting's weights derived measures. The arithmetic mean (AM) combining rule is utilized as additional decision guide. The weights are to be computed utilizing the widely used decision making tool, the Analytical Hierarchy Process (AHP). The proposed consensus based heuristic is particularly adequate for the nature of the binary decision problem in which decisiveness is much more required than a compromise. The presented results can also improve decision-making process in many other disciplines, as e.g. environmental risk assessment, new product launching decision, food quality tracking, etc.

Keywords: Fuzzy Expert Systems, Outputs combination/aggregation, Analytical Hierarchy Process, Multiple Parallel Processing, Binary Group Decision-Making, Consensus Analysis.

References

Aly, S., & Vrana, I., "Integrating Multiple Fuzzy Expert System under Varying Requirements"; Proceedings of the 5th International Congress, on Information Technology in Agriculture, Food and Environment, Adana, Turkey, Oct., 2005.

Vrana I. & Aly S.: Multiple parallel fuzzy expert systems utilizing a hierarchical fuzzy model. Agricultural Economics, 53, 2007 (2): PP. 89-93.

Bryson, N., "Group Decision-Making and the Analytic Hierarchy Process: Exploring the Consensus- Relevant Information Content", Computers Ops Res., Vol. 23, No 1, PP. 27-35, 1996.

Ngwenyama, O. K., Bryson, N., Mobolurin, A. "Supporting facilitation in group support systems: Techniques for Analyzing Consensus Relevant Data"; Decision Support System, 16, PP. 155-168, 1996.

Saaty, Thomas L. (1980): The Analytical Hierarchy Process; New York: McGraw-Hill. Giorgio Guariso, Hannes Werthner, Environmental Decision Support Systems, John Wiley & Sons, 1989, ISBN:0-470-21431-7.

Ulises Cortès, Miquel Sànchez i Marrè, Luigi Ceccaroni, Ignasi R-Roda, Manel Poch, Artificial Intelligence and Environmental Decision Support Systems, Applied Intelligence, 13 (1), 2000.

A.E. Rizzoli and W.Y. Young, "Delivering environmental decision support systems: Software tools and techniques," Environmental Modeling and Software, vol. 12, no. 23, pp. 237--249, 1997)

Andrea Emilio Rizzoli, "Environmental decision modeling": http://www.idsia.ch/~andrea/edss/edss.html

iEMSs 2008: International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Taking the OpenMI forward

Roger V. Moore[a] **and Peter J. A. Gijsbers**[b]

[a] *Centre for Ecology and Hydrology, Wallingford, Oxon, OX10 6HU, UK (rvm@ceh.ac.uk)*

[b] *Deltares, Rotterdamseweg 185, 2629 HD Delft, The Netherlands (Peter.Gijsbe@deltares.nl)*

Abstract: The pressure to take a more integrated approach both to science and to management increases by the day. A consequence of this situation is that there is strong encouragement in the science world not just to understand and to be able to predict the response of individual processes but also to predict how those processes will interact. The manager is similarly encouraged to think in the widest possible terms about the likely impact of any policy before it is implemented. One response to this situation has been an increased interest in integrated modelling and a consequent interest in technologies that allow models of different processes, based on different concepts and from different suppliers to be linked. A European Commission co-funded group has developed such a mechanism, the OpenMI, and is now transforming their research output into an operational standard. However, a technical solution alone is not sufficient. It must be delivered in a way that makes its use acceptable across the world. This paper sets out for discussion in Workshop 13, a strategy for the future development, delivery and use of the OpenMI.

Keywords: Model linking; Model coupling, Integrated Modelling, Integrated water management, OpenMI, Open Modelling Interface.

1. INTRODUCTION

The pressure to take a more integrated approach both to science and to management increases by the day. At almost any scale from local to global, it is no longer possible to consider issues in isolation; to do so runs a high risk of creating more problems than are solved. The consequence of this situation is that there is strong encouragement in the science world not just to understand and to be able to predict the response of individual processes but also to predict how those processes will interact, one with another. The manager is similarly encouraged to think in the widest possible terms about the likely impact of any policy before it is implemented. A new reservoir might solve a water supply problem but will it adversely affect the fishing and hence the local tourist trade? Biofuels may be renewable but will they create a food shortage? One approach to predicting such impacts would be to create new models simulating more and more processes. This, however, is neither feasible nor useful and would make poor use of the huge investment to date in the current models. A better approach, bringing with it many extra benefits, would be to find a way of linking existing models and modelling components such as databases or visualisation systems.

Against this background, the European Commission, as part of its research programme to facilitate the introduction of integrated water management, commissioned HarmonIT, a project to find a generic solution to the linking of simulation models at run time. The outcome of this work was the Open Modelling Interface (OpenMI) together with a set of supporting software. The work has now moved to a second phase, where the aim is to transform the research output into a supported operational standard. It is this second phase that will form the basis for the discussion in Workshop W13.

2. DICUSSION TOPICS

2. 1 Introduction

Learning from the projects that preceded HarmonIT, its leaders appreciated from the outset that success would depend on the project delivering two things: a technical solution and a plan for its support that would make its use acceptable across the world.

HarmonIT developed and tested version 1.0 of the OpenMI. This successfully demonstrated the feasibility of linking not merely models but any modelling components. It did so with components sourced from different suppliers, using different modelling concepts representing a variety of domains. However, there were many known requirements that the available resources did not allow to be addressed, for example, linking a model running in the .Net environment with one running in a Java environment. Further, although a set of supporting software had been developed to help make models OpenMI compliant, link and run them, it was fully appreciated that this would have to be made much more user friendly before it would be accepted by even sophisticated end users. Although the testing involved the use of the OpenMI in a variety of model linking situations, it was primarily aimed at checking that it functioned correctly at the IT level. It did not really demonstrate the benefits sort from integrated modelling.

As the OpenMI became more widely known about and the possibility that it might represent a solution to the problem emerged, so the focus of l user interest shifted to how it would be made available and supported in future. To be acceptable, the OpenMI would have to be open, non-proprietary, international, backed by a credible organisation and able to support its users. Substantial demonstrations based on real world problems under operational conditions would also be required.

While the benefits of integrated modelling and the OpenMI are becoming accepted amongst the modelling community, this is far from being the case amongst managers; indeed, it is probably true that there are some modellers still to be convinced. If the OpenMI is to succeed, then managers must be convinced first of the benefits of integrated modelling and then the OpenMI. It is not just managers who have to be convinced. The model developers must also be convinced. Potentially, the OpenMI changes the modelling market. The OpenMI's creators, who include three leading developers, all believe that the OpenMI creates new opportunites not only for themselves but also for individuals, small companies and universities, because the OpenMI dramatically reduces the cost of entering the market. Similarly, the consultants who run models for government agencies must be convinced. Interestingly and at this stage unexpectedly, the first commercial applications of the OpenMI have been successful not just technically but commercially – the work was completed in less time for fewer resources.

To address these and many other issues, a second project, OpenMI-Life, has been created with the help of the European Commission. The objectives of this project are to develop version 2.0 of the OpenMI and to plan and put in place a support organisation. However, the EC funding rules only cover demonstrations and this has led to an interesting and ultimately beneficial approach to the problem. OpenMI-Life has been set up as an operational demonstration not just of the OpenMI but of the whole support organisation. By the end of the project, if all goes well, the support organisation should have been shown supporting seven operational demonstrations based on real-world problems. During these demonstrations, it will have responded to user requests for enhancements and in response, it will have developed and released Version 2.0. In parallel, it will have undertaken all the dissemination and promotion activities that would be expected of such a support organisation.

The support organisation is now in place and called the OpenMI Association. The remainder of this paper will raise a number of discussion points relating to the future development of the OpenMI and the association, on which the Association would be most interested to hear the modelling community's views, so that it can shape its future strategy.

3. THE OPENMI ASSOCIATION'S FORWARD STRATEGY

The aspirations in the OpenMI Association's Charter have been translated into:

- A vision of the role of modelling in relation to the integrated management of land and water
- A mission statement for the Association
- A strategy for achieving the mission

To stay relevant, all three will need to evolve over time, as needs and technology change. The Association therefore welcomes discussion on its strategy, which is outlined below.

3.1 OpenMI Association Objectives according to its Charter

The OpenMI Association is an association ('vereniging') established under Dutch Law. Its Charter and Standing Orders are available on the Association's website. The Association is open to international membership for both organisations and individuals. According to its Charter, the objective of the Association is:

> *"The promotion of the development, use, management and maintenance of the Open Modelling Interface (the OpenMI), a standard for the exchange of data between computer software in environmental management."*

The Association seeks to achieve this goal, among other things, by:

- *"Exchanging information with regard to the standard specified in the objective, by word, writing and by means of electronic devices such as a website, both within the Association and with other organisations that pursue a similar objective*
- *"Stimulating the maintenance and development of the OpenMI Standard and its supporting software and managing their release*
- *"Stimulating the provision of information and promoting discussion on the OpenMI in Europe and across the world by the organising of and participating in events*
- *"All that which may be further conductive to the objective"*

These aspirations have been translated into a vision of the role of modelling within the management of land and water, a mission statement for the Association and a strategy for achieving the mission.

3.2 Vision

Today's world is one where the fair and sustainable allocation of diminishing natural resources between competing demands is increasingly important to avoid conflicts. Making such decisions, however, is far from easy. Although we have considerable knowledge of individual environmental, social and economic processes in the land and water domains, our ability to predict how those processes will interact is limited; hence, it is difficult to foresee the wider implications of natural events and management policies as is required by integrated management.

It is neither feasible nor particularly useful to build a single model of all the processes in a catchment. However, emerging technologies, such as the OpenMI, are making it possible to link together new and existing models from different suppliers, based on a variety of modelling concepts and representing different domains. There is already preliminary evidence that this linked or 'integrated modelling' approach will offer an effective tool for developing and assessing integrated management options and that it leads to better decisions, i.e. decisions that make the best use of the resources.

The OpenMI Association believes that integrated management in some form or other is the only option for the future management of our resources. Because of the complexities inherent in integrated management, managers will demand ever more sophisticated decision support systems. These are essentially predictive models. As the need to understand the wider impacts of any decision increases, so the models will have to take account of more and more processes. The OpenMI Association, therefore, foresees a future where the concept of integrated modelling is widely accepted by authorities as leading to better decisions, and where it is the adopted practice in optimising the management of complex systems.

3.3 Mission

The attainment of the vision will require the combined efforts of developers, modellers and users. Within this context, the role, and hence the mission, that the OpenMI Association has set itself, is to:

- Promote integrated modelling as a means of achieving better management decisions, so that resources are more fully exploited and impacts are better understood;
- Develop and support the OpenMI Standard so that it becomes the first choice for model integration.

The OpenMI Association has defined for itself the following boundaries:

- Technical scope:
    - The exchange of data between independent models and modelling components either at run-time or when the models are run sequentially
    - Facilitating the embedding of linked models and components in decision support systems (DSS)
- Application domain: environmental management with the priority on land and water systems.
- Intended user groups of the OpenMI and its applications: model developers, the modelling community and the competent authorities.
- Geographic focus of dissemination activities: although the OpenMI Standard can be applied world-wide, the priority regions for the OpenMI Association in the first years will be Europe and the United States of America.

3.4 Implementation Strategy

To achieve its mission, the OpenMI Association will focus on the following key actions:

- Creating a culture that eases the adoption and acceptance of integrated modelling as a concept and the use of the OpenMI as the recognised option for integrated modelling
- Ensuring that the OpenMI remains relevant, easy to use, of high quality and available under acceptable conditions
- Supporting the community of OpenMI users and providing a compliancy service for developers who have adopted the OpenMI Standard in their products
- Disseminating necessary and useful information
- Enabling the user community to influence the development of the OpenMI and its Association
- Being the legal body that holds the ownership of the OpenMI
- Securing the necessary resources to fulfil its mission and implementation strategy

3.4.1 Creating a culture for integrated modelling

Changing attitudes to integrated management and modelling will require the provision of information and the creation of new skills at both high and low levels in user, modelling and developer communities.

The top-down approach will be to ensure that key decision-makers (competent authorities) are aware of the added value of integrated modelling and the role of the OpenMI Standard. Their needs will be identified so that they can be provided with the information that will enable them to make well-informed decisions about the use of integrated modelling and the OpenMI.

In parallel, the OpenMI Association will identify the most widely used model codes and decision support systems and will encourage their developers to make them OpenMI compliant (if not yet so).

The bottom-up approach will be to create a new body of scientists, engineers and IT professionals with integrated modelling and OpenMI skills. This will be achieved by liaising with respected higher educational institutions and encouraging the inclusion of material on integrated modelling and the OpenMI in appropriate undergraduate, MSc and PhD courses, workshops and seminars. The Association will co-operate with the educational institutions to create course material. The development of training services by the Association's members will be similarly encouraged.

3.4.2 Ensuring the OpenMI remains relevant

Following the Open Source model, the Association will maintain, develop and release the OpenMI standard. The Association will establish a group to monitor the driving forces, user needs and new technologies. Using the combined resources of the group and the user community, it will periodically propose changes and, after consultation with the user community and their agreement by the OpenMI Executive Committee, implement them. In this way, the OpenMI Association will seek to ensure the continuing relevance of the OpenMI Standard.

3.4.3 Supporting the OpenMI user community

Through its Technical Committee, the OpenMI Association will provide and supervise a forum for discussion and questions (http://sourceforge.net/projects/openmi/); this forum is intended primarily for developers. Additionally, the Association will provide and maintain a website and a 'wiki' site (wiki.openmi.org). The web site will be the entry point for the Association and the OpenMI. It will contain final copies of all key documents, news and announcements and will be designed with end users and those new to the OpenMI in mind. The site will link to the wiki, which will be the home page of the Technical Committee. The wiki will have a more informal style and will provide access to detailed technical information and reports of user experiences with the OpenMI.

The Technical Committee is not resourced to maintain a staffed help desk and cannot guarantee that all questions on the discussion forums will be answered by them. However, the Association will foster the creation of a mutually supportive OpenMI Community in the open source spirit.

The Association will always recommend end users to go first to their software supplier for support. The OpenMI Association will encourage software suppliers to build up OpenMI skills among their support staff, so that they can help their users with OpenMI-related queries. Hence, from an end user's (modeller's) point of view, the preferred channels of communication are the same as those that exist in today's single-domain modelling. All users seeking support will be welcome to explore the information available through the website and the wiki. All users, but particulalrly developers, should feel welcome to report bugs and make proposals for change to the OpenMI.

The OpenMI Association will supervise the co-ordination and organisation of training sessions for model developers to help them make their modelling software OpenMI-compliant. The Association will not provide the training itself but will collaborate with dedicated third-party trainers, who can advertise their services on the Association website. Registration for training sessions will be co-ordinated through the website. Links to published training materials and tutorials will also be provided.

End users will be asked to contact their own software providers for specific OpenMI-related training. Generic examples and general information on the functionality and use of the OpenMI User Interface, as offered by the Technical Committee, will be provided by the Technical Committee as soon as resources permit.

3.4.4 Dissemination of information

The main medium for disseminating information will be the OpenMI Association website (www.openmi.org), through which all formal information about the OpenMI Standard and the OpenMI Association will be made available. This website will also provide information on events, training and where to find OpenMI-compliant software.

Members will be allowed to advertise their products and services on the website. More informal information, primarily focused at developers, will be held at the OpenMI Association's wiki, which will be accessible via the website. In addition, the Association will provide a forum on Sourceforge for debate, feedback and the informal exchange of information and experience relating to the OpenMI, model integration and integrated modelling.

The OpenMI Association will convene specific sessions on the OpenMI at conferences. It will also publish papers and articles in the scientific and popular press and maintain a newsletter.

Basic material for creating OpenMI-related presentations will be made available through the website.

3.4.5 Availability of the technology

The Association will impose no restrictions on the use of the OpenMI for research or commercial purposes, and will not impose any royalty charges or licence fees.

The OpenMI Association will release the OpenMI Standard and provide access to it through its website. The Standard will be released under the Lesser General Public Licence (LGPL) (http://ww.gnu.org/licenses/lgpl.html). The Association will provide all necessary documentation relating to the Standard.

New releases will be timed to achieve a balance between the need for stability and the need to move with the times.

In order to maintain the high quality and reliability of the Standard, the OpenMI Association will adopt quality assurance procedures and quality control standards.

The OpenMI Association will ensure that there is always at least one documented Software Development Kit (SDK) and Graphical User Interface (GUI) available under suitable open-source licence conditions. It will also allow and encourage third parties to develop, deliver and maintain other versions of the SDK and GUI, or any other tools that will simplify the migration (making modelling components compliant), linking and running of integrated models.

3.4.6 Compliancy

The OpenMI Association will develop a procedure for establishing that components comply with a particular release, and will allow developers to label such components as 'OpenMI version *n.m* Compliant'. The Association will hold a register of OpenMI

compliant components. The Association will not make any statement of warranty about OpenMI-compliant components. The Association will work towards establishing an auditing service.

3.4.7 The Association as a legal entity

The Association will be an international, non-proprietary and not-for-profit organisation.

New members will always be welcome, particularly those willing and able to contribute actively.

The OpenMI Association, as a legal entity, will ensure and safeguard the copyrights and intellectual property rights (IPR) of the OpenMI Standard and any related products created by the OpenMI Association. (The OpenMI logo and the name 'OpenMI' have been registered as a trademark and the domain names www.OpenMI.org and www.OpenMI.com have also been registered.)

3.4.8 Secure resources

The OpenMI Association will seek the resources to support its activities by membership fees, contributions in kind, donations, project funding or any other resource that is approved by the General Assembly.

In order to keep the required resources to a minimum, the OpenMI Association will actively encourage services to be carried out by third parties, thus creating business development opportunities for the OpenMI community.