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 2

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 7: Community modeling, data and model interoperability **711**
(organised by Alexey Voinov, Chris Duffy, David Arctur, Ilya Zaslavsky and Ralf Seppelt)

Addressing Cultural and Institutional Barriers to Data and Model Interoperability
P. Dibner and D. Arctur 711

An European database for integrated assessment and modeling of agricultural systems
S. Janssen, E. Andersen, I.N. Athanasiadis and M.K. van Ittersum 719

A Design for Framework-Independent Model Components of Biophysical Systems
M. Donatelli and A.E. Rizzoli 727

Conceptual Model of Single Information Space for Environment in Europe
J. Hřebíček and K. Kisza 735

Bridging the Gap between Geohydrologic Data and Distributed Hydrologic Modeling
G. Bhatt, M. Kumar and C.J. Duffy 743

Using the GEONAMICA® software environment for integrated dynamic spatial modelling
J. Hurkens, B. Hahn and H. van Delden 751

PTF: an Extensible Component for Sharing and Using Knowledge on Pedo-Transfer Functions
M. Acutis, M. Donatelli and G. Lanza Filippi 759

Community-based software tools to support participatory modelling: a vision
A. Voinov, D. Arctur, I. Zaslavskiy and S. Ali 766

Sharing emergency information between Emergency Control Centres: the Project REACT
U. Delprato, M. Cristaldi and A. Gambardella 775

Session 8: Integrated modelling of the land system **783**
(*organised by Rüdiger Schaldach and Jennifer Koch*)

Towards a simulation-based study of grassland and animal diversity management in livestock farming systems
G. Martin, M. Duru, R. Martin-Clouaire, JP. Rellier, JP. Theau, O. Thérond, L. Hossard 783
792

An Agent-Based Model for Exploring Land Market Mechanisms for Coastal Zone Management
T. Filatova, A. van der Veen and A. Voinov
800

An activity based cellular automaton model to simulate land use changes
J. van Vlieta and H. van Delden
808

Integrated modeling of agricultural land use change in Romania: From retrospective causal analysis to future developments
T. Lakes and D. Müller
817

Go East: A Residential Land Use Model for the Periphery of Rome

G. L. Ciampaglia and B. Tirozzi

827

Xplorah, A multi-scale integrated land use model
H. van Delden, E.R. Gutiérrez, J. van Vliet and J. Hurkens

Session 10: Participatory and Conceptual Modelling – Making Models Amenable to Stakeholders' Concerns **835**
(*organised by Jaroslav Myšiak, Raffaele Giordani, Carlo Giupponi, Farmani Raziyeh, Nigel Quinn, Carmel A. Pollino, Anthony J. Jakeman and Piotr Magnuszewski*)

Multi-Agent Models for Teaching, Extension and Collaborative Learning
T. Berger and P. Schreinemachers 835

Addressing stakeholder concerns using the Integrative River Rehabilitation Model (IRRM)
M. E. Borsuk, S. Schweizer and P. Reichert 842

Participatory Approaches in Developing a Model to Assist Water Resource Management in a Catchment in the Solomon Islands
T. Chan, B. Powell, S. Hoverman and H. Ross 850

Improving communication in urban planning using TOPIC-COHESIE
H. van Delden, P. Valkering, P. Luja, M. van de Lindt and V. Neevel 858

ARDI: a co-construction method for participatory modelling in natural resources management
M. Etienne, D. Du Toit and S. Pollard 866

An Evolutionary Bayesian Belief Networkbased Methodology for Adaptive Water Management
R. Farmani and D. Savic 874

The integration of Expert and Stakeholder Cognitive Models to support Environmental Monitoring
R. Giordano, S. Liersch, M. Vurro and V.F. Uricchio 880

Participatory process management
J. Krywkow and M. Hare 888

The Event Bush as a Potential Complex Methodology of Conceptual Modelling in the Geosciences
Pshenichny C. A., Nikolenko S. I., Carniel R., Sobissevitch A.L., Vaganov P. A., Khrabrykh Z. V., Moukhachov, V. P., Shterkhun V. L., Rezyapkin A. A., Yakovlev, A.V., Fedukov, R.A. and Gusev, E.A.
900

Connecting Science and Decision-making: A Conceptual Framework through Organisation Knowledge Management
T.F. Siew 913

Session 11: Bridging the gaps between design and use: evaluating decision and information support tools **925**
(*organised by Brian S. McIntosh, Carlo Giupponi, Keith Matthews and Martin van Ittersum*)

A note on attitudes towards and expectation from the Decision Support Systems
J. Mysiak 925

Assessing the impact of environmental decision and information support tools
B. S. McIntosh and E. Diez 932

DSS Success Measures: Evaluating the SCaRPA DSS
N. Herron and S. Cuddy 940

Raising the Bar – Is evaluating the outcomes of decision and information support tools a bridge too far?
K. Matthews 948

User driven application and adaptation of an existing Policy Support System to a new region
H. van Delden 956

Usability assessment of an integrated land use model in Portugal
C. Alves 964

Participary modelling on fire prevention and urbanisation in southern France: from co-constructing to playing with the model
M. Etienne 972

Opportunities offered by the design and the use of a gridded multi-scale European Soil Information System in support of European soil policy
P. Panagos and M. van der Liedekerke 980

Eururalis – discussion support tool for rural Europe
P. Verweij 988

Application of multi criteria assessment in Bereg floodplains
Z. Flachner 999

Sustainability Impact Assessment Tools (SIAT): A meta-model design for cross-sectoral European Impact Analysis
S. Sieber 1008

From INGIS to NAMOS – Operationalisation and contextualisation of sustainability at the local level
D. Rink 1017

Session 12: Routine checks of model consistency: methodology and information technologies **1025**
(*organised by Georgii A. Alexandrov*)

A Methodology for Building Credible Models for Policy Evaluation
C. A. Aumann 1025

Evaluating consistency of biosphere models: software tools for a web-based service
G. A. Alexandrov and T. Matsunaga 1033

The Carbon-Land Model Intercomparison Project (C-LAMP): A Protocol and Evaluation Metrics for Global Terrestrial Biogeochemistry Models
F.M. Hoffman, J.T. Randerson, I.Y. Fung, P.E. Thornton, Y.-H. Lee, C.C. Covey, G.B. Bonan and S.W. Running 1039

Reducing Uncertainty in Ecological Niche Models with ANN Ensembles
A. P. Kirilenko and R. S. Hanley 1047

Towards model evaluation using Self-Organizing Maps
M. Herbst and M.C. Casper 1055

A Software Component for Model Output Evaluation
G. Bellocchia, M. Donatellia, M. Acutisb, E. Habyarimanaa 1063

Using SOAP to Develop Software for Air Trajectory Calculation
J. Zeng, T. Matsunaga and H. Mukai 1070

Web Components for Development of Computational Methods: Example with Fuzzy
G. Bellocchi, F. Foscarini, M. Canonico and G. Van den Eede 1077

Session 13: OpenMI applications and innovations **1086**
(*organised by Roger Moore and Peter Gijsbers*)

OpenMI Compliant Import of Initial and Boundary Data into a Numerical 3D Model
P. Schade, G. Lang and J. Jürges 1086

Demonstration of Integrated Modelling using the OpenMI in the Scheldt River Basin
N. Devroede, Y. Ronse, J. Van Assel and H. Vereecken 1094

The Migration of the UTHBAL Hydrologic Model into OpenMI
A. Loukas, K. Kokkinos, L. Vasiliades and A. Liakopoulos 1102

Session 14: Environmental modeling applications towards quality of life information services **1110**
(*organised by Kostas Karatzas*)

An Application of Data Mining to PM10 Level Medium-Term Prediction
E. Pasero, A. Montuori, W. Moniaci and G. Raimondo 1110

Neural Network based method for predicting regional visitor attendance levels in recreational areas
T. Räsänen and M. Kolehmainen 1118

Neural Networks and Co-Kriging techniques to Forecast Ozone Concentrations in Urban Areas
C. Carnevale, G. Finzi, E. Pisoni, V. Singh and M. Volta 1125

Studying and predicting quality of life atmospheric parameters with the aid of computational intelligence methods
K D. Karatzas and D. Voukantsis 1133

A multiscale Air Quality Forecast System for Europe and selected regions
H. Jakobs, M. Memmesheimer, E. Friese and A. Ebel 1140

The "Demokritos" web-based air quality forecasting system for the Greater Athens Area
Vlachogiannis D., Sfetsos A., Andronopoulos S., Gounaris N., Yiotis A.and Stubos A.K. 1146

The Use of Third Generation Air Quality Modeling Systems for Web Operational Real-time Forecasting Decision Support Systems: Spain Case 1154

R. San José, J.L. Pérez, J.L. Morant, F. Prieto and R.M. González

Web-system for air quality assessment of the city area based on the mathematical modeling data
A. Titov, D. Belikov, E. Gordov and A. Starchenko 1162

Runoff projection sensitivity to rainfall scenario methodology
F. S. Mpelasoka and F. H. S. Chiew 1169

Assessing Activity-Related Vehicle Emissions through an Integrated Activity-Based Modelling Framework
C. Beckx, T. Arentze, L. Int Panis, D. Janssens and G. Wets 1177

A storm surge model implementation and identification of coastal areas in risk of inundation, in the Mediterranean Sea
Y. Androulidakis, Y. N. Krestenitis, Y. Kontos and G. Georgakopoulos 1187

Development of a Web based Decision Support System for trans-boundary environmental risk assessment and management
Pediaditi K., Stanojevic M., Kouskouna M., Karydas, C. Zianis D., Boretos, N. 1195

Estimating the Pollution Risks to Urban Groundwater from Industry and Sewers
D.N. Lerner, B.N. Chisala and N.G. Tait 1203

Spatial Information System for Risk Assessment of Dust Transport in the Neighbourhood of the Surface Coal Mine
L. Matejicek and Z. Janour 1209

Session 15: Renewable resources use for energy production **1215**
(*organised by Giorgio Guariso, Riccardo Minciardi, Roberto Sacile, Michela Robba and Giulia Fiorese*)

Cost Effectiveness Analysis for Renewable Energy Sources Integration in the Island of Lemnos, Greece
A. Angelis-Dimakis, P. Trogadas, G. Arampatzis and D. Assimacopoulos 1215

Implementation of a GEOdatabase to administrate global energy resources
M. Biberacher, S. Gadocha, S. Gluhak, N. Dorfinger, D. Zocher 1223

Integrated Model Platform for Sustainable Energy Planning. Colombian case
R. Quijano and J. Domínguez 1231

GIS modelling of forest wood residues potential for energy use based on forest inventory data: Methodological approach and case study application
L. Panichelli and E. Gnansounou 1239

Research activities in renewable energy sources integration with GIS at CIEMAT
J. Domínguez, I. Pinedo and J.A. González 1247

An Environmental Decision Support System for an Efficient Monitoring and Planning of Woody Biomass for Energy Production, with Remote Sensing Aid
F. Frombo, M. De Martino, R. Minciardi, M. Robba and S.B. Serpico 1255

Rapid Sizing of Renewable Energy Power Components in Hybrid Power Plants for Reverse Osmosis 1264

Desalination Process
A. Kartalidis, G. Arampatzis and D. Assimacopoulos

Particle Swarm Optimization for the biomass supply chain strategic planning
J. Izquierdo, R. Minciardi, I. Montalvo, M. Robba, M. Tavera 1272

Forest Biomass Sustainable Use for Energy Production: a dynamic optimization problem
F. Frombo, R. Minciardi, M. Robba, F. Rosso and R. Sacile 1281

Optimizing biogas production: an application to an Italian farming district
G. Fiorese, G. Guariso, A. Polimeni 1289

Decision support models for scheduling hydro and thermal power plants in the liberalized electric energy market
M.T. Vespucci and M. Innorta 1297

GIS based Model to optimize possible self sustaining regions in the context of a renewable energy supply
M. Biberacher, S. Gadocha and D. Zocher 1306

Session 16: Modelling and Support tools for management and treatment of the integrated wastewater system **1314**
(*organised by Ignasi Rodríguez-Roda, Krist V. Gernaey, Eduardo Ayesa, Joaquim Comas, Manel Poch, Miquel Sànchez-Marrè, Ulises Cortés, Ion Irizar, Jean Philippe Steyer, Juan M. Lema, Mark van Loodstrecht and Gürkan Sin*)

The Benchmark Simulation Models – A Valuable Collection of Modelling Tools
J.B. Copp, U. Jeppsson and P.A. Vanrolleghem 1314

Global sensitivity analysis of biochemical, design and operational parameters of the Benchmark Simulation Model no. 2
L. Benedetti, D. J. Batstone, B. De Baets, I. Nopens and P.A. Vanrolleghem 1322

Uncertainty and Sensitivity Analysis of Control Strategies using the Benchmark Simulation Model No1 (BSM1)
X. Flores-Alsina, I. Rodriguez-Roda, G. Sin and K. V. Gernaey 1331

Assessment of data availability influence on integrated urban drainage modelling uncertainty
G. Freni, G. Mannina and G. Viviania 1339

Comparison of control strategies for multiobjective control of urban wastewater systems
G. Fu, D. Butler and S.-T. Khu 1347

AqquaStudy – enabling remote access to model-based software tools for supporting the design and operation of WWTPs
I. Irizar and M. De Gracia 1354

Towards an Agent-Based design for the management of Urban Wastewater Systems
M. Aulinas, M. Poch and U. Cortés 1362

Improving waste water treatment quality through an auction-based management of discharges
J. Murillo, D. Busquets, J. Dalmau, B. López and V. Muñoz 1370

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

Addressing Cultural and Institutional Barriers to Data and Model Interoperability

P. Dibner[a], D. Arctur[b]

[a]*OGC Interoperability Institute, Inc., Los Altos, California USA, pdibner@ogcii.org;*
[b]*OGC Interoperability Institute, Inc., Austin, Texas USA, darctur@ogcii.org*

Abstract: Despite recent technological advances, proliferation of online databases and community data collection and modeling efforts, the environmental observatory and modeling communities remain fragmented due to the lack of summaries of available observations data, differences in information models and metadata, lack of common data discovery and access protocols tuned to model requirements, and significant semantic differences in data description. However, purely technical solutions for interoperability are insufficient for establishing a shared interoperable infrastructure. The establishment of any technology in a population of potential adopters takes place within a context of social and economic processes. A community cannot become instantaneously aware of all the available data and software resources available to it, nor can consensus arise uniformly about which information models, data access mechanisms, and encodings are most appropriate. The theory of technology diffusion describes the process by which innovations are accepted by successive subgroups within a population, and suggests a series of activities to encourage adoption. The related phenomena of increasing returns and path dependence can facilitate the spread of interoperable software services and harmonized data models, but may also lead to lock-in of suboptimal solutions. Fortunately, the low cost, consensus-driven development and minimally invasive nature of interoperable web service interfaces, and the abstract foundations of well-conceived data models minimize this risk, and also further promote acceptance of these technologies.

Keywords: Interoperability; Semantics; Data sharing; Information models; Technology diffusion; Increasing returns.

1. INTRODUCTION

The holy grail of current-day integrative scientific studies is seamless interoperability across arbitrary data sets from arbitrary domains of knowledge. The purely technical aspects of this challenge have been the topic of years of effort in a broad range of studies in the domains of information science and data communications, resulting in various solutions that have been called technical or syntactic. Such solutions consist of suites or libraries of function calls, messages, or service operations that can be implemented by software components distributed throughout the Internet, and whose parameters and results are expressed using well-defined and well-known data structures.

Equal in importance to technical interoperability is semantic or data interoperability, wherein the requests and results passed among the components of a distributed software system carry content that is understood in some broad sense to convey meaning, to carry concrete values for concepts that are interpreted equivalently by all the elements of the system, and also by the system's human users. Technology for expressing such notions has also been developed and is the topic of much current research and a growing range of applications; it falls under the domain of knowledge representation, is expressed as ontologies (controlled vocabularies and, in brief, the definitions of their elements), and forms a fundamental component of the emerging semantic world-wide web.

In addition to current-day requirements for integration of information contained in widely distributed data sources is a growing awareness of the need to integrate data *processing* capabilities within and across domains of knowledge, or put another way, a need to enable interoperability among software *models*. The term "model" is overloaded even within the context of data processing and computer science. For purposes of this discussion, a model is a software program, representing some physical process or phenomenon, that requires a set of parameters as input, possibly at many points in time during its execution, and generates a result or set of results. We specifically exclude from this definition the term "information model" or, equivalently, "data model," which refer to the conceptual view and semantic content of elements within a domain of knowledge.

Technologies for structural and semantic aspects of these issues are still evolving, but have achieved a certain degree of maturity. Yet, they are only the first requirement for a truly interoperable and universally accessible solution to data sharing and integration. The ultimate goal of universal interoperability for data resources and models depends upon propagation of knowledge about the various solutions, acceptance and awareness of the advantages that it confers, agreement about which interoperability technologies are appropriate and desirable, concurrence about the representation and specific content of semantic descriptions, and generation of motivation and development resources for implementation and adoption. These are all social rather than technical processes, and subject to many factors, including limitations and outright barriers that characterize their cultural and institutional milieus.

Coming to grips with these issues in a generic, intellectually defensible, and broadly comprehensible manner requires expertise and experience both within a comprehensive sample of the domains of integrative scientific research, and also well beyond them, touching on topics that embrace economics, psychology, business administration, and a range of other esoteric domains, each potentially a life's work in its own right. We offer here a few perspectives from the standpoint of our particular experience in data modeling, environmental and biological science, and the process of standards development and advocacy from within the context of one international standards body, the Open Geospatial Consortium (OGC).

2. PRELUDE: INTEROPERABLE SOLUTIONS AS A SINGLE ENTITY

Much of the following discussion refers to the uptake of interoperable technologies as if they were one monolithic entity, or treats one such technology as a proxy for all of them. In point of fact, even if we restrict our attention to the OGC alone, we encounter a substantial and growing number of specifications and standards, of a variety of types. The best known, and probably still the most widely represented in deployed systems, are data services, which allow any compliant client application to request and obtain information maintained by the archive and supporting software behind the service interface. A significant and growing number of data models have also flowed from OGC efforts, starting with the foundational Geography Markup Language (GML) [OGC, 2007c] and proceeding to a series of domain-specific application schemas. Most of these data models are specified in terms of their structural or syntactic encodings in the XML Schema language, but the semantic concepts underlying them are also part of the specifications and are well understood within their respective information communities. OGC specifications also include catalogues and associated methods for discovery of resources, as well as an entire set of documents that support sensing devices and human-mediated data collection programs. Specifications that can provide interoperable interfaces to computational models and workflows include the Sensor Planning Service (SPS) [OGC, 2007b], which is more a generic tasking interface than a sensor-specific interface protocol, and the Web Processing Service (WPS) [OGC, 2007a].

For all this, from the standpoint of potential adopters, there is no deep distinction between interoperable geospatial services, data and metadata encodings and meanings, and interoperable modeling paradigms. Essentially the same considerations apply to adoption of any or all of these technologies, and they amount in sum to a substantial component of the "whole product," a concept introduced by Moore [1991] and revisited below.

3. TECHNOLOGY DIFFUSION AND THE PROCESS OF ADOPTION

Any examination of factors that delay or prevent adoption of interoperable technologies would be incomplete without consideration of patterns and processes understood to govern or describe technology adoption in the absence of any particular cultural or institutional barriers. Study of these phenomena has increased during the past several decades, and they now constitute one of the pillars of marketing strategy.

3.1 Technology Diffusion: Background

The adoption of new technology is not and can never be instantaneously universal. There is a dynamic to the spread of information about an innovation, the way it garners the attention of interested parties, and the degree to which it provokes a sense of necessity within members of the target audience.

We consider this process in the light of the theory of diffusion of innovations throughout a population of potential adopters, as originally described by Rogers (see, e.g., Rogers [2003]), and applied and popularized by Moore [1991]. In brief, the theory describes a Technology Adoption Life Cycle in which five different groups within a target population successively adopt innovative technologies. In general, an innovation must be accepted by one group before the next group starts to consider it seriously. The first group, the Innovators, is typically a small minority of the target population. Innovators are technologists, enthusiastic about exploring new technology on its own merits. The next group are the Early Adopters, still a minority, but visionaries who see new technology as a way to establish a competitive advantage, or more generally, produce a major step forward for the sake of their organization or its mission. Their purchasing decisions are based more on their imagination and vision than on established references, and they are willing to take risks in order to gain the benefits that they foresee. The Early Majority are the next group to adopt a technology, comprise a large proportion of the target population, and are key to wholesale acceptance and operational deployment of innovations. Individuals who belong to this group are driven by pragmatic concerns such as remaining abreast of proven solutions and not falling behind beneficial practices that competing organizations have embraced. They are interested only in solutions that are stable and well-supported, and they base purchasing decisions on references from sources that they trust, which necessarily depend on the accumulation of time and experience, and are not available until later in the product's lifecycle. The final two groups are the Late Majority, who adopt technology only after it has been incorporated into standard operations, and the Laggards, who resist adoption. These latter two groups are not of concern in this discussion except to note that they are the final adopters (or non-adopters), in any population.

In the case of a discontinuous innovation, i.e., one that involves a substantial change in strategy, employee behaviour, and new capabilities with new implications for an entire business, there is a particularly difficult gap that new technology must cross in order to succeed and be accepted by the entire population. This "chasm" is the divide between the Early Adopters and the Early Majority. It is hard to traverse because the Early Majority depend on relationships, proven reliability, and references, and none of these are available early in a product's lifecycle. Experience at this point lies with the Early Adopters, but their criteria are not of the sort valued and trusted by the pragmatist. The theory goes on to suggest a four-step process that maximizes the opportunity for "crossing the chasm", i.e., enabling the widespread adoption of the new technology by the Early Majority. We discuss these below in the context of interoperable geospatial technologies.

3.2 Applicability

This theory was conceived in the context of commerce and marketing of technical products by individual organizations. The case we are considering here involves the global spread of open interoperability standards for data and models, and engagement in technical development as well as operational deployment of these models by many cooperating as well as competing organizational entities. Many of the same considerations apply, but the

cases are really quite different and merit further exploration, as well as examination of the recent history of the development of standards-compliant implementation and uptake in actual deployments.

3.3 Process for "Crossing the Chasm" in the Arena of Geospatial Interoperability

We return to the four-step process prescribed by Moore [1991] in further evaluating the condition and progress of interoperable solutions in the geospatial domain, with particular reference to the process and guidelines for developing such solutions in the context of programs conducted by the Open Geospatial Consortium. Recall that the aim of this process is to engender the adoption of innovative technology by the pragmatic Early Majority, who insist upon reliable and well-referenced solutions before accepting them as a part of their organizations' infrastructure. Moore poses the response to this problem using the analogy of a military operation.

Target the Point of Attack

The first step in this process is to "target the point of attack," or in more conventional terms, to identify a market segment that truly requires and will measurably benefit from the innovation. This step is vital, particularly in the context of commercial organizations. Experience has shown that it is impractical and generally not productive to attempt to promote adoption by the entire target population. An effort of this magnitude would be diluted by the sheer number of individuals that the program was attempting to reach, and would be a fatal flaw within a group that depends heavily on well-established relationships and references. A coordinated effort in one market segment avoids this pitfall.

In the geospatial domain, the collective pain of what might be considered a single market segment engendered the creation of the precursors to the OGC and ultimately the Consortium itself. These were the government agencies, including the Earth Observation agencies, who were stewards of increasingly large volumes of substantial imagery datasets, and who needed to share their data with other government, academic, or commercial organizations. Replicating such large datasets was often prohibitively expensive, and converting formats to allow integration of these with other sources of information was not only costly, but a significant engineering enterprise in its own right, with its own technical challenges, requirements for specialized development, and other sources of delay. Even as computing power and storage capacity increased, so did the capability and practice of collecting satellite and aerial imagery, as well as the types of imagery and other specialized data with spatial and temporal attributes. The OGC started as an early coalition of Innovators and Early Adopters, who sought and ultimately developed workable solutions to the interoperability problem. The need for these innovations was such that the early efforts at development and experimentation were also blessed with an audience from within the Early Majority among the agency population, awaiting a set of technologies that were sufficiently stable to deploy as operational systems.

The success in this market segment, such as it was, provided the basic essentials for propagation of interoperable geoprocessing capabilities to additional agencies and institutions: a well-defined technical solution, a suite of increasingly reliable commercial products, and a growing set of operational deployments, already tested in one segment of the population of potential adopters. This has become a beachhead for further propagation into the Early Majority within other organizations, including city and county governments and the community of professionals involved with environmental modeling and analysis.

Assemble the Invasion Force

This second step of the Moore paradigm entails making sure that the product, when presented to the customer, provides a complete solution, or what Moore calls a "whole product." In practical terms, this means not only that the product is functional and reliable

as an isolated entity, but that other services and capabilities essential to support it are also ready and available to potential adopters.

In the context of interoperable data services, this could include documentation and training, a community of developers and consultants, and an entire suite of products that can be compared and assessed on the road to implementation as operational deployments.

In the context of interoperable data models, it means not only that the model is supported by documentation and a community of implementers, but that the paradigms supported by the data model are truly representative of the target population of adopters. In order for an information model to be useful in the context of interoperable web services, there must also be ancillary technology such as client software that understands the model sufficiently to render or analyze the data, format conversion tools for dataset translation or on-the-fly conversion so that legacy data available via interoperable services, catalog services to enable discovery.

In the context of OGC technology, this step is far from complete, and its maturity varies considerably across the spectrum of services and information models that comprise interoperable geospatial technology. However, the "invasion force" is sufficiently well assembled to gain acceptance and credibility in some domains, representing the leading edge of the "assault." In particular, the OGC data services, including the Web Map Service (WMS) [OGC, 2006], Web Coverage Service (WCS) [OGC, 2007d], and Web Feature Service (WFS) [OGC, 2005] have been incorporated into numerous commercial and open-source products and enjoy increasingly broad use in the geospatial analysis community. Well-defined and broadly accepted information models compatible with the WFS are also emerging in increasing numbers, but still await broad deployment. Approaches to interoperability of simulations and computational models are still undergoing experiment and development. However, their ultimate adoption will be facilitated by the technologies that are already enjoying success in the marketplace, and upon which their technical underpinnings are substantially based.

Define the Battle

This step refers to the need to introduce a product in the context of other competing solutions. If the goal is interoperability of geospatial data and services, then the battle is already defined. The "competition," such as it is, consists of datasets and models that do not interoperate. The costs of non-interoperability are typically well articulated by outreach organizations and individuals who promote interoperable solutions in their own organizations. Thus the real challenge is to elucidate these factors in the context of actual communities that are just starting to encounter the need for interoperable solutions to their data sharing and service operations.

Launch the Invasion

The final step in Moore's paradigm is to initiate the actual marketing and sales campaign, with channels to the customer appropriately selected to reach and match the requirements of the decision makers. The invasion in this case has been in process for several years, and it is ongoing. The initial success of the WMS, WFS, and WCS has already been mentioned. These are in effect the leading edge of the wave of web services in a growing and maturing architecture.

4. INCREASING RETURNS AND PATH DEPENDENCE

Technology markets are different from traditional commodity- or resource-based economic models like those formulated in the nineteenth century. Such conventional models are based on the phenomenon of diminishing returns, i.e., that each increment of investment is accompanied by a reduction in the value gained or profits realized as return on the incremental investment. It is particularly easy to see how this model applies when the amount of a desired resource is limited: as the resource becomes scarcer or more difficult to obtain, its cost increases and thus reduces the profitability of products or processes based upon it. Thus the dynamic is characterized by negative feedback, a stabilizing force. In the

case of competing enterprises based upon different resources, a random increase in the market share of one enterprise increases the costs associated with it, and its competitors become incrementally more attractive by comparison. The classical result is a stable equilibrium in which a combination of the solutions persists in the market, in proportions dependent solely on the marginal costs and returns of the various options, and independent of the history or timing of random fluctuations in market dominance by any of the competitors.

The situation is very different if returns increase rather than decrease on marginal investment. Far from acting as a stabilizing force, economic activity generates positive feedback. In the case of competing enterprises, a slight advantage gained by one is amplified, giving it a greater advantage and increased market share. The end result is a situation where the entire market is dominated by one of the competitors. Moreover, the selection of which one is very sensitive to the history of fluctuations in market share or other advantages, and not determined exclusively by the cost-value propositions offered by the competing enterprises. Therefore an increasing-returns market may lead to broad adoption of suboptimal products that are difficult to displace.

The world of technology is frequently characterized by an increasing-returns dynamic. This is particularly true of the software industry, especially in the domain of application programming interfaces (APIs). Software developed to use a particular set of APIs promotes the acquisition of systems that support those APIs, which therefore assume a greater market share and so become an increasingly desirable platform upon which to base new software. Suites of compatible protocols that enable interoperability among diverse services, data sources, and computational installations are of course prime candidates for triggering this behavior in the marketplace.

To the extent that OGC standards are appearing in operational deployments, this would seem to be a benefit to the goal of promoting interoperability, but it is important to consider and somehow address the danger of locking in suboptimal or even poor solutions. We consider this issue below.

5. DISCUSSION AND ANALYSIS

It is almost legend that some corners of the research community are concerned that adherence to standards may have negative impacts on creativity, or may not satisfy requirements in some way essential to scientific progress. But in our experience, this legend ultimately holds little substance. Every community with which we have engaged includes a contingent of informatics experts who are well versed in the characterization of data in their own domain, who are also seeking, if not already knowledgeable about, the proper approach to integrate with and make use of standard services and data paradigms. However, scientific communities too have their populations of Early Adopters and Early Majority, and as in the commercial domain, the latter are more concerned with maintaining a reliable and productive research facility than with investing resources in promising but still risky propositions for enhancing productivity.

Even when the relevant interoperability technologies have achieved a level of stability and generality that would be acceptable to the Early Majority population segment, barriers remain. They may include extensive investment and training in existing or historical practices, legacy datasets, priorities for proximal goals that have little requirement for or benefit from interoperability, and others, all in combination with limited and previously dedicated resources.

OGC standards were designed be compatible with existing installations and on existing operations, and we posit that this would be a feature of any suite of interoperable service protocols or data models designed for widespread adoption. Many of the requirements that lead to wide acceptance are already built into OGC standards because they were developed by representatives from many domains, whether technical, scientific, or commercial, and intended to be accessible to very broad implementation. It was a given from the outset that they must work with all major software and hardware platforms, they must be neutral with respect to data storage formats, management and custodial practices, and they must also be

reasonably straightforward to install and operate.

The protocols themselves therefore tend to be minimal, and to make few assumptions about the installations that they support. They can in general be implemented as façades for proprietary or more complex systems for data retrieval or analysis, and they need not be exclusive: for a given service they can act as one among many interfaces to the worldwide web.

The same is true of the data models promoted by OGC technology. They are based upon the unadorned, abstract ISO Feature model as specified in the ISO 19101 [2003] and ISO 19109 [2006] specifications, and are not inherently dependent either on GML encoding or OGC service protocols. The utility of data models harmonized to one another and compatible with OGC technology extends beyond OGC-compliant installations.

If indeed these characteristics are essential for any suite of technologies that enable very broad-based deployment of interoperable services, then not only do they present a low barrier for implementation by technologists and acceptance by pragmatists within their respective domains, but they also serve as a counterweight to the lock-in phenomenon that can characterize increasing-returns markets. Service installations should not represent a huge investment, and should not be prohibitively expensive to replace if need be.

6. CONCLUSIONS AND RECOMMENDATIONS

The primary conclusion of this assessment is that data and model interoperability appear poised to undergo an adoption process similar to that of the simpler geospatial data services and products in other highly technical domains. The picture is a bit confused and substantially blurred because of the great diversity of potential adopter populations and the urgency with which they need to integrate their data assets and modeling capabilities with each other and with those of other communities. Additional factors include different rates of uptake for the broad array of service and data models that have been defined, and differences between the populations of end users, service implementers, and modelers.

An increasing returns market and associated path dependency seem to favor development of interoperable service, data, and software modeling in a consistent direction, but raise the specter of inflexibility and broad acceptance of suboptimal solutions. Service definitions that impose minimal restrictions upon existing or new implementations, and carefully conceived and broadly vetted data models both mitigate these potential negatives and facilitate acceptance within communities wary of impacts upon their operations or resources.

The recommendations that flow from these observations are aimed at promoting the further spread of interoperable data services and computational models, and they fall into two main categories.

First, engage in activities that lower the barriers to technology diffusion. In particular, focus on issues that are appropriate to the current stage of development of the standard of concern. For technologies that are still experimental, seek involvement and feedback from the Innovators and Early Adopters in target communities of interest, including your own organization. As the technologies mature, leverage and participate in efforts to refine the product offerings, whether through implementation, development of tools or documentation, or participation in beta tests for commercial or open-source developers.

In terms of communication with hesitant parties, it is important both to maintain the mindset and communicate the truth of the fact that supporting effective interoperability standards is an opportunity, not a restriction. Data standards must arise within the community that understands and uses the underlying models.

The second general recommendation is to leverage the phenomenon of the increasing returns market.

First and foremost, develop, adapt, or incorporate existing standards-compliant products in your operations. To the extent that experimental or pilot implementations are available, use them, and do not overlook the substantial and growing framework of geographically extensive and increasingly high resolution datasets with spatial and temporal content, and the network of OGC services that provide access to them. Early in the history of the OGC Web Services (OWS) initiatives, there were only a few service implementations that supported broad datasets such as the USGS transportation networks, hydrological features, political boundaries and similar assets, and satellite imagery were available, but they added immense value to experiments and demonstrations. Today, many more standards-compliant services are available, and access to them is made all the easier due to the wide availability of software tools, both commercial and open-source, that support the same standards.

A small or partial solution is in general preferable to no solution. If you are able to publish a portion of your data using an interoperable service, or are able to access only one of many interoperable web-based resources, do so.

For data modeling efforts, engage with the broadest possible segment of the relevant community. It is not necessary for a community to produce only one data model. Sometimes there are different, even incompatible models produced by different segments of the same community. But usually, even though the models cannot be mapped successfully to one another, there is an overarching, often simpler construct to which the mutually incompatible models can be successfully mapped. Regardless, consider implementing a data service that supports the model you prefer. At a minimum, you will be making it accessible to the market of potential users, which promotes experimentation and allows broader assessment.

REFERENCES

Arthur, W. Brian, *Increasing Returns and Path Dependence in the Economy*, U. Michigan Press, 201 pp., Ann Arbor, 1994.

ISO 19101:2003, *Geographic Information – Reference Model*, 2003.

ISO 19109:2006, *Geographic Information – Rules for application schemas*, 2006.

Moore, Geoffrey A., *Crossing the Chasm*, Harper Business, 227 pp., New York, 1991.

OGC, *OpenGIS® Web Feature Service (WFS) Implementation Standard*, OGC Document 04-094, 2005.

OGC, *OpenGIS® Web Map Server (WMS) Implementation Specification*, OGC Document 06-042, 2006.

OGC, *OpenGIS® Web Processing Service (WPS) Implementation Specification*, OGC Document 05-007r7, 2007a.

OGC, *OpenGIS® Sensor Planning Service (SPS) Implementation Specification*, OGC Document 07-014r3, 2007b.

OGC, *OpenGIS® Geography Markup Language (GML) Encoding Standard*, OGC Document 07-036, 2007c.

OGC, *OpenGIS® Web Coverage Service (WCS) Implementation Standard*, OGC Document 07-067r5, 2007d.

Rogers, Everett M., *Diffusion of Innovations*, 5th ed., Free Press, 518 pp., New York, 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

An European database for integrated assessment and modeling of agricultural systems

Sander Janssen[1,2], Erling Andersen[3], Ioannis N. Athanasiadis[4] and Martin K. van Ittersum[1]

[1] Wageningen University, Plant Production Systems Group, Wageningen, sander.janssen/martin.vanittersum@wur.nl

[2] Wageningen University, Business Economics Group, Wageningen

[3] University of Copenhagen, Forest and Landscape Group, Copenhagen, eran@life.ku.dk

[4] Dalle Molle Institute for Artificial Intelligence (IDSIA), Lugano, ioannis@idsia.ch

Abstract: Integrated Assessment and Modelling (IAM) can be used to assess socio-economic and environmental indicators, which generally require the linkage of models from different domains. To integrate a set of different models for an IAM, the data required by each of the models as inputs from a range of data sources also needs to be consistently integrated. This paper describes the process of development of a database integrating different data sources for an IAM project, and the human factors involved in the process of reaching consensus across peers with clashing requirements and needs. We adopted a structured process using a shared ontology as a means to one integrated relational database serving a set of models of a highly multi-disciplinary nature. The relational database covers data on agricultural systems, e.g. soil, climate, farm, agricultural management and agricultural policy data. The integrated database has been coupled to a range of quantitative models. The database schema and the shared ontology are distinct products that can be reused for or extended by other IAM projects requiring a similar set of data. It is recommended for any IAM project in which several models are coupled to adopt an explicit, collaborative and iterative process to specify an adequate data structure for storing data used in the project. For such a process to succeed it has to focus on the relevant domain knowledge captured across the data sources and this paper offers a proposal for such a process.

Keywords: Community modeling, agricultural systems, database, European Union

1. INTRODUCTION

Integrated Assessment and Modelling (IAM) is increasingly used to assess the impacts of policies, technologies or societal trends on the environmental, economic and social sustainability of systems (Parker, et al., 2002). IAM is a methodology to combine several quantitative models representing different systems and scales into a framework for Integrated Assessment (Parker, et al., 2002). Consequently, IAM can cover several organisational and spatio-temporal scales to provide quantitative assessment of impacts. To integrate a set of different models for an Integrated Assessment and Modelling project, the data required by each of the models as inputs and produced as outputs generally need to be consistently integrated. Each of the quantitative models used in an IAM is derived from a different discipline, requires different and to some extent overlapping data-sources, and is operational on different spatial and temporal scales.

SEAMLESS is an IAM research project (Van Ittersum, et al., 2008), which aims to provide a computerized framework to assess the sustainability of agricultural systems in the European Union at multiple scales. This aim is achieved by combining micro and marco level analysis, addressing economic, environmental and social issues, and facilitating the re-use of models and providing methods to conceptually and technically link different models together (Van Ittersum, et al., 2008).

Within SEAMLESS we faced a difficult data-integration challenge. Data have to serve dynamic biophysical models, static bio-economic farm models and partial computable general equilibrium market models. This required the integration multiple data-sources (including data related to European agriculture, including economic, biophysical, climatic data, model simulation input and output data, scientific workflow configurations and visualization of indicators) into a single relational database schema.

The objective of this paper is to describe the process of development of the SEAMLESS database, and the human factors involved in the process of reaching consensus across peers with clashing requirements and needs. The SEAMLESS European database on agricultural systems is presented. We adopted a structured process using a shared ontology as a means to arrive at one integrated relational database serving a set of models of a highly multi-disciplinary nature. This process is re-usable for other IAM projects, whereas the end result in terms of the database is re-usable for IAM of agricultural systems in Europe.

The next Section will describe firstly some theory behind ontologies and process of ontology engineering and the data sources of relevance to the SEAMLESS project. Consequently the results will be presented in the third Section as a description of the European database on agricultural systems, as the links between ontology and database and as the process used to construct this database with a group of researchers. Finally, conclusions and recommendations are provided.

2. MATERIAL AND METHODS

2.1 Ontologies and relational databases

In the context of integrated modelling, ontologies are useful to define the shared conceptualization of a problem, as ontologies consist of a finite list of concepts and the relationships between these concepts (Antoniou and van Harmelen, 2004) and as ontologies are written in a language, e.g. Web Ontology Language (McGuinness and van Harmelen, 2004), that is understandable by computers. In research aiming to integrate different models, scientists from various disciplines can define a common ontology that their domains share. A common ontology serves as a knowledge-level specification of the joint conceptualization, in our case of the data-sources used in the Integrated Assessment and Modelling project. Our efforts focused on the development of such a high-level ontology for the SEAMLESS data.

The common ontology is subsequently transcribed to the integrated relational database scheme, based on the conventions of the Semantic-Rich Development Architecture (SeRiDA) (Athanasiadis, et al., 2007). The SeRiDA combines object-oriented programming, relational databases and ontologies as three separate layers each with a distinct role: OWL ontologies for expressing rich domain semantics, Enterprise Java Beans™ for end-user application development, and normalized relational databases for persistence storage (Athanasiadis, et al., 2007). Through the SeRiDA the mapping of object-oriented models to ontologies is facilitated, while it provides an Object Relational Mapping (ORM), thereby acting as a bridge between different programming paradigms.

The use of ontologies has as advantages that the ontologies are richer in their representation of relationships between concepts than relational database schemas, have a strong implementation of inheritance, can be used as documentation tool for metadata, can be used for source code generation and allow to capture knowledge on the system under study as a distinct product.

2.2 Process of ontology engineering

In developing a common ontology, a group of scientists should agree and adopt one tight, well-reasoned and shared conceptualization. The development of a common ontology by a

group of researchers is a complex, challenging and time-consuming task (Gruber, 1993, Holsapple and Joshi, 2002). Tools are available that help in ontology development and to store the ontology once it has been developed (e.g. Protégé OWL) (Knublauch, 2005). In developing the common ontology for the different data sources in our project, a collaborative approach was used. A collaborative approach is based on 'development as a joint effort reflecting experiences and viewpoints of persons who intentionally cooperate to produce it' and it thus requires a consensus-building mechanism (Holsapple and Joshi, 2002)'. As part of this collaborative approach, an inductive approach was used (Holsapple and Joshi, 2002). In our inductive approach, the common ontology was developed by examining and analyzing the data-structures of the initial data-sources and extracting relevant properties or discussing the relationships.

2.3 Data sources

The data sources of relevance to the model-based assessments in the SEAMLESS project are:

(i) The Farm Accountancy Data Network (FADN) (EC, 2008a) is a source for evaluating the activities and income of agricultural holdings and the impacts of the Common Agricultural Policy. It consists of an annual survey carried out by the Member States of the European Union. The member states in the Union collect every year accountancy data from a sample of the agricultural holdings in the European Union (EC, 2008a). The data collected are, for example, physical and structural data, such as location, crop areas, livestock numbers, labor force, and economic and financial data, such as the value of production of the different crops, sales and purchases, production costs, production quotas and subsidies.

(ii) The European Soil Database (ESBN, 2008) on soils in Europe aims to provide a harmonised set of soil parameters, covering Europe (the enlarged EU) and bordering Mediterranean countries, to be used in agro-meteorological and environmental modelling at regional, national, and/or continental levels. Its scale is 1: 1.000.000 and it contains Soil Geographical Database of Eurasia, PedoTransfer Rules Database, Soil Profile Analytical Database of Europa and Database of Hydraulic Properties of European Soils (ESBN, 2008).

(iii) The European Interpolated Climate Data (JRC, 2008) provides interpolated daily data for a grid of 50 x 50 km covering Europe and Maghreb (average period 1975 -today). The majority of the original observations data originates from around 1500 meteorological stations across the European continent, Maghreb countries and Turkey.

(iv) Farm management data have been collected through dedicated surveys as part of the SEAMLESS project (Borkowski, et al., 2007). In the SEAMLESS project a lack of European data on agricultural management was identified. With agricultural management data is meant the use of inputs (fertilizers, pesticides, irrigation) and the timing of input use on farms. Surveys (Borkowski, et al., 2007) were developed as part of the SEAMLESS project. Data collected in these surveys are timing and amounts of inputs, crop rotations, machinery, labour requirements and costs.

(v) The COCO/CAPREG dataset (Britz, et al., 2007) is based on NewCronos (Eurostat, 2008) and FAOSTAT (FAO, 2008). It contains complete and mutually consistent time series for hectares/herd size, output coefficients, production, market balances, economic accounts and unit value prices (incl. consumer prices). For SEAMLESS, the relevant part of the COCO/CAPREG is the data on agricultural policies in the European Union.

The datasets from the Farm Accountancy Data Network, European Soil Database and European interpolated Climate data have been categorised into typologies (Metzger, et al., 2005, Andersen, et al., 2007, Hazeu, et al., 2007) to enable modelling of homogenous spatial units and to allow for characterization and sampling. The data sources haven been aligned with existing administrative categorizations like the Nomenclature of Territorial Units for Statistics (NUTS) (EC, 2008b).

3. RESULTS

3.1 Method to develop the integrated database

Initially, the data from the different sources were stored in eight different databases. To develop a common ontology for all data-sources, three scientists (a computer scientist, a landscape and forest ecologist and database expert, and a systems analyst) engaged in an integration process. These three scientists involved other domain experts in the integration process, when additional knowledge was required.

As a kick-off, a three day meeting was organized with experts on the database content and database set-up. Data-modeling was used to create a data-schema during the meeting. The result of this meeting was a database schema for some of the databases, which was subsequently translated into an ontology using Protégé (Knublauch, 2005). Next step was to extend this ontology by including all the relevant data sources required for running the models. This process lasted for over a period of six months with frequent discussions through email and web-meetings. During this period two additional face-to-face meetings were required of only one day. This first version of the common ontology was exported to the first version of the SEAMLESS relational database schema using the SeRiDa-framework. The SEAMLESS database schema v.1.0 was discussed and improved between the three scientists involved in the project in roughly 3 iterations, leading to a first stable version of the database schema. Subsequently, the data from the original sources were entered into the database, which led to new improvements of the ontology and database schema v2.0.

When this database schema v2.0 was filled with data, the models were coupled to it using the Enterprise Java Beans™ generated by the SeRiDA framework. In coupling and running the models, some errors and required extensions of the common ontology were identified. These errors and required extensions were discussed and solved as part of the review of the database schema v2.0. During the review the three scientists tried to simplify and improve the schema as much as possible. This review lasted about two months and was organised through web-meetings and phone calls. Other domain experts were involved for their opinion on parts of the schema, which led to database schema v3.0. The data could be entered without requiring revisions into this version of the database schema. As part of the fourth version of the database schema, metadata will be included as part of the ontology.

3.2 European database on agricultural systems

Figure 1 provides an overview of the ontology developed for the European database on agricultural systems as developed in the SEAMLESS project. As can be seen from Figure 1, which shows the part of the database of relevance to soil, farm and climate data, there are concepts which classify the data, for example Farm Specialization, Farm Size and NUTS region and there are concepts that hold the actual data, like Representative Farm, Soil Characteristics and Daily Climate.

A central concept of the ontology is the concept of Representative Farm, which defines a FarmType in an FADN region in Europe for a specific year. A FarmType is specified according to the dimensions of farm size, farm intensity and farm specialization (Andersen, et al., 2007) (Fig. 2). As an example of a classifying concept, Farm intensity is a classification of farms according to their total output of agricultural produce per hectare (Andersen, et al., 2007). If the total output is below 500 euros per hectare, then the farm falls in the class of low intensity, if it is between 500 and 3000 euros, then it is medium intensity and if is more than 3000 euros, then it is high intensity. While a FarmType is not linked to a specific region or year, a Representative Farm is specific to a region and a year.

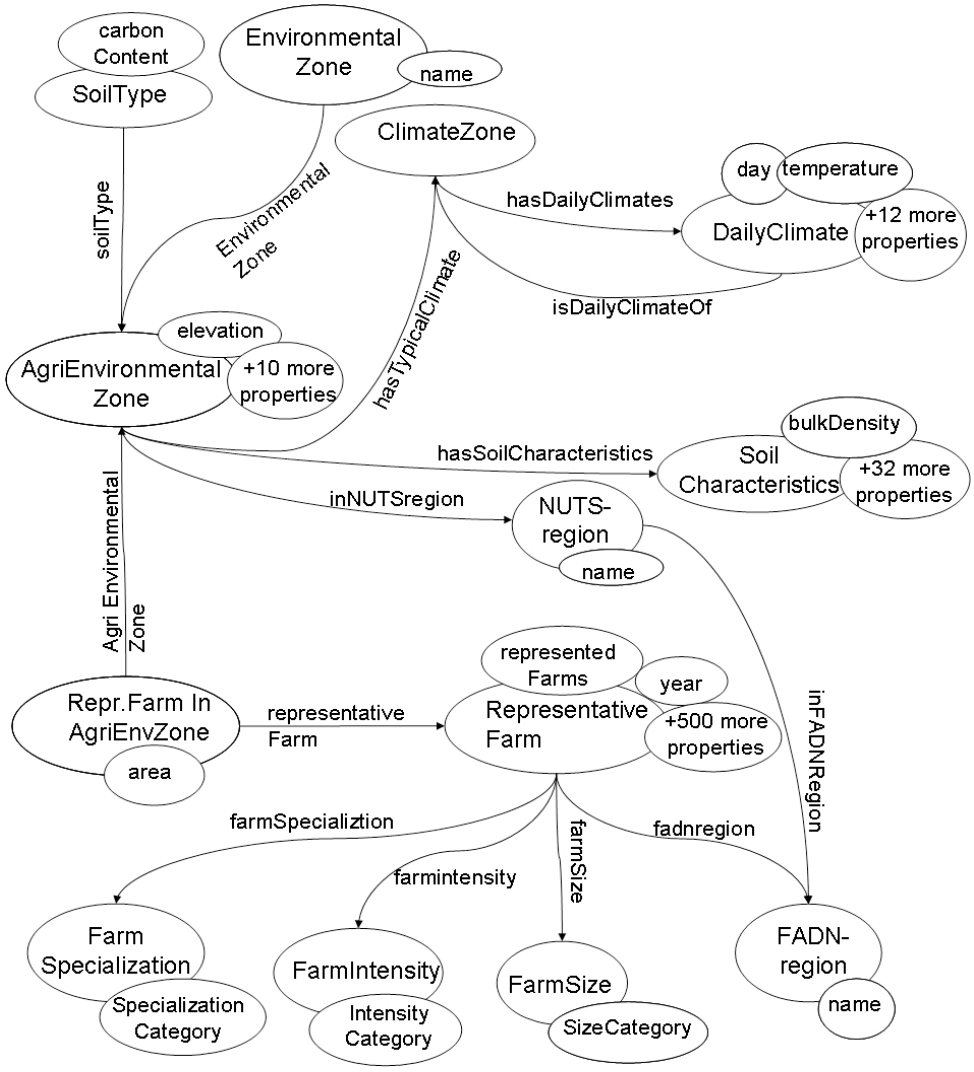


Figure 1. An ontology-schema of the European database on agricultural systems showing the parts on farms, soils, climate and their links.

As can be seen in Figure 1, AgriEnvironmental Zone is a central concept, in that it links to soil and climate data. An AgriEnvironmental Zone is a unique combination of an Environmental Zone, the soilType and NUTS region. AgriEnvironmental Zones are the smallest homogenous units in a region in terms of climate and soil data. Environmental zones are used to stratify the diverse European Union climate in zones with a similar climate (Metzger, et al., 2005). The Environmental Zones cover more than one region, and a Climate Zone is thus a unique combination for a NUTS-2 region and Environmental Zone for which a set of climate data is available. A Climate Zone provides the daily climate data for a 30-years time period for a region and Environmental Zone, so one record for every day. Examples of properties of daily climate data are rainfall in mm per day, average daily temperature in degrees Celsius per day and wind speed at 10m in m/s.

Each AgriEnvironmental zone is linked to a set of soil data, as classified according to Soil Types. Six different Soil Types were defined according to topsoil organic carbon classes (Hazeu, et al., 2007). For each unique combination of a Soil Type and a NUTS-region a set of soil data is available as stored in the concept of Soil Characteristics. Examples of properties of the soil characteristics are thickness subsoil and topsoil, depth to rocks and saturation top soil.

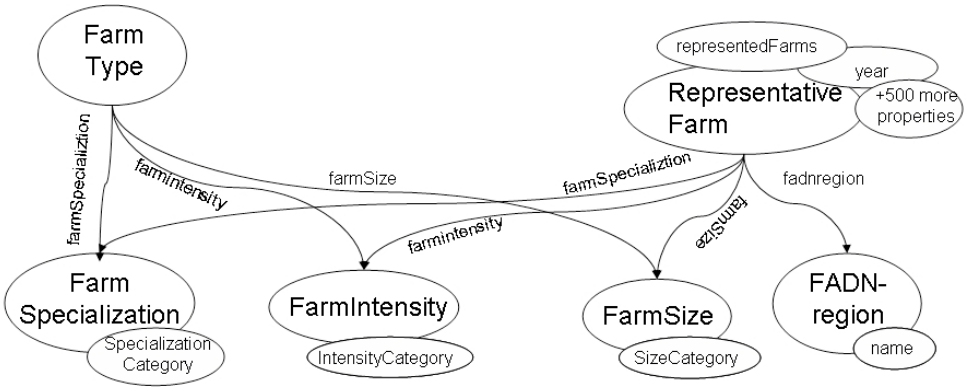


Figure 2. The concepts Farm Type and Representative Farm and the relationships to their classifying concepts.

The link between AgriEnvironmental Zones and Representative Farms is made through allocating an area of an AgriEnvironmental Zone to each Representative Farm. This implies that each AgriEnvironmental Zone is allocated to one or more Representative Farms and each Representative Farm can be found in one or more AgriEnvironmental Zones. As can be seen from Figure 1, Representative Farms and AgriEnvironmental Zones are based on different administrative regions e.g. AgriEnvironmental Zones refer to NUTS-2 regions (EU25 has 270 NUTS-2 regions) and Representative Farms refer to FADN-regions.

In this paper the link between agricultural management data, policy data, AgriEnvironmental Zones and Representative Farms will not be explained in detail. As the agricultural management differs within regions, Regional Agricultural Management Zones were created. A Regional Agricultural Management Zone has a distinct set of agricultural management data and is linked to one or more Agri-Environmental Zones. Finally, data on agricultural policies and prices are linked to each NUTS-region. The current version of the database consists of 329 tables including 2 035 fields and with 379 relations between the tables. The number of records in the database now exceeds 7.4 million.

3.3 Links between ontology and relational database

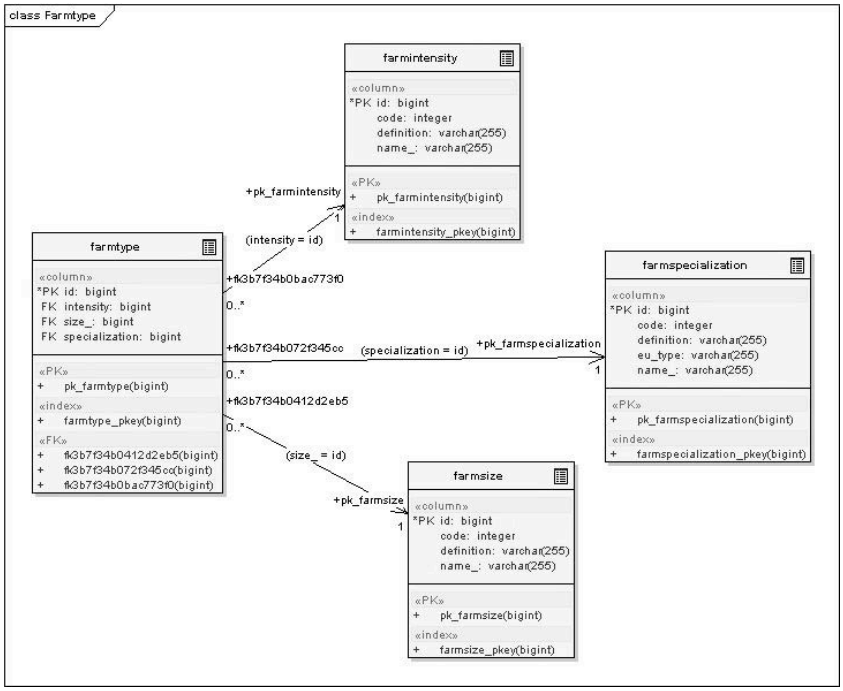


Figure 3. a relational database schema of the FarmType, FarmIntensity, FarmSize and FarmSpecialization-tables (made with Enterprise Architect©).

Figure 3 presents part of relational database schema related to FarmType as it is generated from the ontology schema from Figure 2. From Figure 3 it can be seen that all the relationships between the tables are enforced through foreign and primary keys. The FarmType table is linked by many-to-one foreign keys to the classifying tables FarmIntensity, FarmSize and FarmSpecialization. These many-to-one foreign keys represent the relationships farmSpecialization, farmIntensity and farmSize from Fig. 2, which describe that each FarmType has one and only one reference to the classifying concepts of FarmIntensity, FarmSize and FarmSpecialization. This example demonstrates the translation of the ontology into the relational database schema that is usable for persistent data storage. More examples can be found in Athanasiadis, et al.(2007).

4. CONCLUSION AND RECOMMENDATIONS

By using ontologies in a collaborative process of conceptual modelling, we managed to derive a common database schema that integrates a range of data sources from different domains specified at different spatial and temporal scales. This common database schema and the common ontology on which it is based are distinct products that can be reused for or extended by other research projects requiring a similar set of data. The integrated database has been linked to a range of quantitative models and can be coupled to other models with similar data requirements. It is recommended for any Integrated Assessment and Modelling project in which several models are linked or complex models are developed to adopt an explicit process to specify an adequate data structure for storing data used in the project. This paper provides a proposal for such a process, which should be collaborative and iterative. Using a framework like SeRiDA for mapping between programming paradigms allows the programmers to benefit from the strengths of each of the programming paradigms. Also, adopting an explicit process to specify an adequate data structure and a framework likes SeRiDA helps scientists to focus on the domain content of the data structure, while not loosing focus in details of technical implementation in different programming paradigms.

ACKNOWLEDGMENTS

We thank all scientists in the SEAMLESS project who contributed to development of the common ontology on projects and scenarios. This work has been carried out as part of the SEAMLESS Integrated Project, EU sixth Framework Programme, Contract No. 010036-2.

REFERENCES

Andersen, E., B. Elbersen, F. Godeschalk and D. Verhoog, Farm management indicators and farm typologies as a basis for assessments in a changing policy environment, *Journal of Environmental Management*, 82(3), 353, 2007.

Antoniou, G. and F. van Harmelen, *A Semantic Web Primer*, The MIT Press, 238 pp., Cambridge, Massachusetts; London, England, 2004.

Athanasiadis, I. N., F. Villa and A. E. Rizzoli, Enabling knowledge-based software engineering through semantic-object-relational mappings, paper presented at 3rd International Workshop on Semantic Web Enabled Software Engineering, 4th European Semantic Web Conference, Innsbruck, Austria, 2007.

Borkowski, N., P. Zander, G. Stokstad, Sandra Uthes, F.-J. Reinhardt, G. Bellocchi and M. Hecker, Conceptual Approach to Identify and Assess Current Activities, PD3.3.9, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, 23 pp. pp., Wageningen, 2007.

Britz, W., I. Pérez, A. Zimmermann and T. Heckelei, Definition of the CAPRI Core Modelling System and Interfaces with other Components of SEAMLESS-IF, SEAMLESS Report No.26, SEAMLESS integrated project, EU 6th Framework Programme, contract no.010036-2, 116 pp., 2007.

EC, Farm Accountancy Data Network (FADN) Source: EU-FADN-DG AGRI-G3, European Commission. Retrieved on 15 February 2008 from http://ec.europa.eu/agriculture/rica/.

EC, Nomenclature of territorial units for statistics - NUTS Statistical Regions of Europe, European Commission. Retrieved on 15 February 2008 from http://ec.europa.eu/comm/eurostat/ramon/nuts/home_regions_en.html.

ESBN, European Soil Database, European Soil Bureau Network, European Commission-Joint Research Centre (JRC). Retrieved on 15 February 2008 from http://eusoils.jrc.it/esbn/Esbn_overview.html.

Eurostat, European Dissemination Database (Formerly New Cronos), Retrieved on 15 February 2008 from http://epp.eurostat.ec.europa.eu/.

FAO, FAOSTAT, Food and Agriculture Organisation of the United Nations. Retrieved on 15 February 2008 from http://faostat.fao.org/.

Gruber, T. R., A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition*, 5, 199-220, 1993.

Hazeu, G., B. Elbersen, E. Andersen, B. Baruth, C. A. van Diepen and M. J. Metzger, The SEAMLESS biophysical typology: a spatial agri-environmental modeling framework, In: Brouwer, F. and M. K. van Ittersum (Ed.) *Unknown; forthcoming*, Springer Academic Publishing, 2007.

Holsapple, C. W. and K. D. Joshi, A collaborative approach to ontology design, *Communications of the ACM*, 45(2), 42-47, 2002.

JRC, Meteorological data Source JRC/AGRIFISH Data Base - EC - JRC, AGRIFISH unit-Institute for the Protection and Security of the Citizen (IPSC) European Commission - Joint Research Centre. Retrieved on 15 February 2008 from http://agrifish.jrc.it/marsstat/datadistribution/.

Knublauch, H., Protege OWL, Stanford Medical Informatics. Retrieved on 24 January 2008 from http://protege.stanford.edu/.

McGuinness, D. and F. van Harmelen, OWL Web Ontology Language Overview, WWW Consortium. Retrieved on 24 January 2008 from www.w3.org/TR/owl-features/.

Metzger, M. J., R. G. H. Bunce, R. H. G. Jongman, C. A. Mucher and J. W. Watkins, A climatic stratification of the environment of Europe, *Global Ecology and Biogeography*, 14(6), 549-563, 2005.

Parker, P., R. Letcher, A. Jakeman, M. B. Beck, G. Harris, R. M. Argent, M. Hare, C. Pahl-Wostl, A. Voinov, M. Janssen, P. Sullivan, M. Scoccimarro, A. Friend, M. Sonnenshein, D. Barker, L. Matejicek, D. Odulaja, P. Deadman, K. Lim, G. Larocque, P. Tarikhi, C. Fletcher, A. Put, T. Maxwell, A. Charles, H. Breeze, N. Nakatani, S. Mudgal, W. Naito, O. Osidele, I. Eriksson, U. Kautsky, E. Kautsky, B. Naeslund, L. Kumblad, R. Park, S. Maltagliati, P. Girardin, A. Rizzoli, D. Mauriello, R. Hoch, D. Pelletier, J. Reilly, R. Olafsdottir and S. Bin, Progress in integrated assessment and modelling, *Environmental Modelling & Software*, 17(3), 209, 2002.

Van Ittersum, M. K., F. Ewert, T. Heckelei, J. Wery, J. Alkan Olsson, E. Andersen, I. Bezlepkina, F. Brouwer, M. Donatelli, G. Flichman, L. Olsson, A. Rizzoli, T. van der Wal, J.-E. Wien and J. Wolf, Integrated assessment of agricultural systems- a component based framework for the European Union (SEAMLESS), *Agricultural Systems*, 96, 150-165, 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

A Design for Framework-Independent Model Components of Biophysical Systems

M. Donatelli[a,,†], A.E. Rizzoli[b]**

[a] Agriculture Research Council, 40128 Bologna, Italy;
[**] currently seconded to European Commission Joint Research Centre, Institute for the Protection and Security of the Citizen, Agriculture Unit, Agri4cast Action, Ispra, Italy
[b] Dalle Molle Institute for Artificial Intelligence, Manno-Lugano, Switzerland;
[†] Corresponding author, marcello.donatelli@jrc.it

Abstract: Most efforts in the design of software frameworks for biophysical systems simulation have focused on the compromise between domain specificity and use flexibility. Models in such frameworks fall in two main categories: either framework-specific, or "legacy" code. In the former case, models implemented as software components can take full advantage of the framework services, but they depend on the framework. In the latter case, components are seen as discrete units of software, in general of coarse granularity in modelling terms, and the dependency on the framework is minimal, but the potential for composition and reuse is limited. Thus, modellers who want to use a modelling framework are faced with two choices: if the framework is extensible, implement a framework specific component (i.e. not reusable outside the specific framework); else, the alternative is to provide a component as a black-box, taking little or no advantage of the framework itself. We argue that component design choices, rather than being driven by the specific framework architecture, should rather promote re-usability by including design traits that represent a compromise between generality and specificity in order to maximize the adaptability of components. This paper present: 1) a software design of non-framework specific components, and 2) real-world applications of the design presented.

Keywords: Modelling; Component-oriented programming; Software components.

1. INTRODUCTION

In the past decade there has been an increasing demand for modularity and replaceability in biophysical models (e.g. [Jones et al., 2001]; [David et al., 2002]; [Donatelli et al., 2003, 2004, 2006]), both aimed at improving the efficiency of use of resources and at fostering a higher quality of modelling units. Rather than having generalist modellers working on all details of complex integrated models, it is more efficient and effective to assemble sub-models developed by specialists in the specific sectors.

The modular approach developed in the software industry is based on the concept of encapsulating the solution of a modelling problem in a discrete, replaceable, and interchangeable software unit. Such discrete units are called components. A software component can be defined as "a unit of composition with contractually specified interfaces and explicit context dependencies only. A software component can be deployed independently and is subject by composition by third parties" [Szypersky et al., 2002].

Component-oriented designs actually represent a natural choice for building scalable, robust, large-scale applications, and to maximize the ease of maintenance in a variety of domains, including agro-ecological modelling [Argent, 2004]. This concept has been applied to biophysical simulation and has led to the development of several modelling frameworks (e.g. Simile, MODCOM, IMA, TIME, OpenMI, SME, OMS, as listed in Argent and Rizzoli [2004], and in Rizzoli, Leavesley et al. [in press]), which allow the use

of components by linking them either directly, or through a simulation engine, if they expose their interface requesting a numerical integration service. However, targeting model component design to match a specific interface requested by a modelling framework decreases its re-usability. This possibly explains why modelling frameworks, although a great advance with respect to traditional model code development, are rarely adopted by groups other than the ones developing those [Rizzoli et al., in press].

A possible way to overcome the problem of scarce reusability of components is to adopt a component design that targets the intrinsic re-usability and interchangeability of model components (e.g. Donatelli et al., [2006a, b]). Such components can be used in a specific modelling framework via an application of the design pattern Adapter [Gamma et al., 1994]. Such classes act as bridges between the framework and the component interface.

The communication between a component and its client (a framework as above), and more in general the ease of re-use, can be enhanced by addressing the following requirements:

1. the component must target the solution of a sufficiently widespread modelling problem;
2. the published interface of the component must be well documented and it must be consistent;
3. the configuration of the component should not require excessive pre-existing knowledge and help should be provided in the definition of the model parameters;
4. the model implemented in the component should be extensible by third parties,
5. the dependencies on other components should be limited and explicit;
6. the behaviour of the component should be robust, and degrade gracefully, raising appropriate exceptions;
7. the component behaviour should be traceable and such a trace should be scalable (browseable at different debug levels);
8. the component software implementation should be made using widely accepted and used technologies.

In the following we present the choices we made in the design of some components developed by an informal network of biophysical modellers.

2. THE DESIGN

The design of a software component implementing a biophysical model requires the ability to identify the right trade-offs in order to deliver a software product that complies with the requirements, while respecting given performance constraints. In particular, we identified a number of design choices, which helped us in the realisation of the software components. We can classify the choices in three broad categories: structure, interaction, and quality. Choices regarding the model granularity and the model interface pertain to the design of the component structure. Choices on the component extensibility and the component dependency are related to the design of the component interaction with a modelling framework. Choices on the component reliability, on the tracing of its execution and even on the technology for its implementation are strictly connected to the measure of the component quality.

2.1. Design of component structure

2.1.1 Model granularity

An essential part of modelling is choosing the model resolution power, that is, the model *granularity*. As a simple example, think of a population dynamics model, where the fractional growth rate is a parameter. This model can be refined, letting the fractional growth rate become a function of the population density. The fractional growth rate is therefore described by two models with different granularity: a constant parameter, or a limiting function. A similar situation holds for the alternative formulation of the limiting functions, which can be exchanged to originate different population growth curves.

"Fine grained" components are more likely to be reused for specific computations, in the context of larger modelling problems. The formulations of the limiting functions can be exchanged among different models, even in different contexts, from the growth of population of bacteria, to the growth of plant biomasses. In this sense, fine grained components match the requirement of providing a solution to a sufficiently widespread modelling problem. Yet, the complexity of a fine grained model can rapidly grow and become unwieldy, requiring a lot of domain specific knowledge to make an actual reuse of the basic components. On the other hand "coarse" modelling solutions have the advantage of hiding the complexity of several modelling approaches by providing a pre-made composition of those, yet their reuse in larger systems limits the flexibility of the overall model; in fact a large parte of it, represented by the "coarse" component, cannot be modified, with problems in terms of maintenance and further development.

Design patterns provide us with a compromise: using an implementation of the Façade pattern we can still adopt a "coarse" modelling solution, hiding the complexity of each model solution while preserving a modular structure based on simple model units. Hence, simple model units can either be used in isolation or they can be composed to develop other modelling units. Examples are the CLIMA components [Donatelli et al. 2005] which implement fine-grained models that generate synthetic weather variables. The component architecture adopts the Strategy design pattern in order to allow for the plugging-in of alternative model formulations to generate the model output, since various models can be used for the same purpose. Each component exposes an interface which must be implemented by each model unit, both internal and from extensions (see 2.1.2) of the component (see Figure 1). The choice of inheriting behaviour via an interface, rather than inheriting implementation from a base class, maximizes flexibility and still allows for inheritance from a class in platforms like Java or .NET which do not allow for multiple inheritance. The CLIMA components referenced above match such requirements.

2.1.2 Model interface

A component implements one or more modelling solutions for a specific domain. A software component can be seen as a block box that processes inputs and returns outputs. Also a dynamic model can be seen as identified by its state transition function and its output transformation function [Zeigler, 1995]. It is therefore straightforward to package the model in a software component exposing the model inputs and outputs as interface. The use of software components is equally valid for static and dynamic components, in the latter case interface variables are often declared as types with names such as States, Rates, Auxiliary, Exogenous etc.). It is also evident that modelling choices define an abstraction of the domain being modelled by selecting which inputs and outputs to consider and how to describe and detail them. Del Furia et al., [1995] proposed to use object-oriented data structures called *domain classes* to describe and implement such data. Each attribute of a domain class classes has, in the design presented, beside its value, a set of attributes such as minimum, maximum, and default value; units; description.

The value of domain classes goes beyond their meaning as software implementation items, in fact they describe the domain of interest providing information about each variable used in the interfaces, making the interfaces semantically explicit. (Athanasiadis et al. 2008) therefore extended the concept of a domain class including the possibility to refer to a publicly available ontology via the attribute URL. There is a direct mapping between domain classes and ontologies, since the domain class code can be automatically generated from the ontology. The clear advantage is that the ontology is language and platform (and framework) independent and it is therefore possible to provide a description of the domain class, which is totally abstracted of the specific technological solution which is adopted. A simple application to generate the code of domain classes in .NET is currently available [DCC, 2006].

An interesting consequence is that if domain classes and interfaces are implemented in a separate unit from models, the model software unit can be replaced without affecting the client using the Domain classes which are part of the signature of the interface methods (the interface that all strategies must implement). An example of such interfaces, for static component, is shown in Figure 1. A component implementing domain classes and

interfaces and another implementing models are a *unit of reuse*; the model component alone can be defined as a *unit of interchangeability* (see Figure 2).

The interface to access models is used both for simple models ("simple strategies") and for composite models ("composite strategies"), which make use of other simple models by implementing an association to simple strategies. This is an implementation of the Composite design pattern (Bishop, 2008). At the same time, composite strategies hide the complexity of the system, providing a single point to access articulated modelling approaches, hence making an implementation of the Façade pattern, as previously mentioned. Also, the structural implementation of the pattern Composite allows an easy implementation of the behavioural pattern Strategy. In fact, context strategies can be built encapsulating the logic to select alternate models (strategies) according to various criteria, e.g. input data availability.

```
/// <summary>
    /// Interface that all ClimIndices strategies must implement
    /// </summary>
    public interface IClimIndicesStrategy : IStrategy
    {
        /// <summary>
        /// Makes estimate
        /// </summary>
        /// <param name="d">Data-type for daily and monthly weather data,
        /// and site data</param>
        /// <param name="u">Univariate statistics</param>
        /// <param name="dc">Clim Indices</param>
        void Estimate(DataWeather d, UnivariateDataWeather u,
                                                    DataClimIndices dc);
        /// <summary>
        /// Tests pre-conditions
        /// </summary>
        /// <param name="d">Data-type for daily and monthly weather data,
        ///  and site data</param>
        /// <param name="u">Univariate statistics</param>
        /// <param name="callID">Client ID for specific call</param>
        /// <returns>Result of pre-conditions test</returns>
        string TestPreConditions(DataWeather d, UnivariateDataWeather u,
                                                        string callID);
        /// <summary>
        /// Test post-conditions
        /// </summary>
        /// <param name="dc">Clim Indices</param>
        /// <param name="callID">Client ID for specific call</param>
        /// <returns>Result of post-conditions test</returns>
        string TestPostConditions(DataClimIndices dc, string callID);
        /// <summary>
        /// Resets output to NaN / smallest integer possible
        /// </summary>
        /// <param name="dc">Clim Indices</param>
        void ResetOutputs(DataClimIndices dc);
        /// <summary>
        /// Set parameters default
        /// </summary>
        void SetParametersDefault();
    }
```

Figure 1. The interface of a strategy in a software component that computes weather indices.

An example of an online ontology browser for the domain classes and components implementing the design described and referenced in this paper, is available at: http://www.apesimulator.it/ontologybrowser.aspx. The software components can also be inspected using a Windows application [MCE, 2006].

We have thus targeted the requirement that the interface of the component must be well documented and it must be consistent. Moreover, thanks to the ontology, the configuration of the component does not require excessive pre-existing knowledge, and help is provided in the definition of the parameters, since they can have default, minimum and maximum values.

2.2 Design of the component interaction

2.2.1. Component API and component extensibility

A software component exposes an application programming interface (API) that allows the use and integration of the component in a software architecture. The API of the component must implement a pattern like the Create-Set-Call [Cwalina and Abrams, 2006]: objects are created via a default constructor with no parameters, some attributes are set, and finally the model is called.

A component implementing a dynamic model must be iteratively called to compute the updated values of the state and output variables, as specified by the state transition and the output transformation equations. Our design choice was to provide a unique method to call all models. The signature of the method can be like the one that follows:

```
Update(DomainClass d, IStrategy s);
```

where `d` is an input-output object (a domain class) and `s` is a strategy, that is, a particular modelling choice to be used in the state transition or output transformation. Being `DomainClass` and `IStrategy` public, and being inheritance from `DomainClass` allowed, the component method can be extended both from the domain class and the strategy viewpoint. In fact, if an extension of the domain class is needed, the new domain class will inherit from `DomainClass`, whereas a new strategy will implement `IStrategy`. This allows using in clients the same API, usable for both the original models and extended models. This design choice answers the requirement of extensibility of the component. Implementation tests were made both in Java and in C# comparing this solution to direct calls to algorithms, resulting in a negligible difference in performance.

Finally, components should be stateless, to simplify their use in different systems. Sample clients, inclusive of code which show how to use and extend the component, must be made available.

2.2.2 Component dependencies

While dependencies should be kept at a minimum, we found necessary and particularly useful to introduce a dependency to another component, named Preconditions, available both in C# and Java (http://www.apesimulator.it/help/utilities/preconditions) which provides the essential services required by this software architecture:

- it contains the base interfaces for models (`IStrategy`) and domain classes (`IDomainClass`);
- it provides a type (`VarInfo`) used to describe the attributes of each variable (name, min and max values, default value, etc.);
- It is used to guarantee the quality of the implemented solution via the design-by-contract approach.

Given that the instances of the VarInfo type contain information on the variables, the Precondition components uses this information to run a suite of built-in tests. More tests can also be defined using the built-in ones. This component outputs tests results to screen, to a default TXT file, to an XML file, and to a listener to be defined by the client using the component; other output drivers can be developed implementing the interface `ITestsOutput` made available by the component. Other dependencies to specific libraries (e.g. for numerical calculus) can be included, but no dependency to any specific frameworks is implemented. The UML diagram of Fig.1 shows the main components used according to the design presented. The design choices we made in the realization of the Precondition component implement the requirement for limited and explicit dependencies.

2.3. Design of component quality

2.3.1 Components reliability

The robustness of a software component is greatly enhanced if the implementation follows the design-by-contract approach [Meyer 1992], which requires that a clear contract between “client” and “server” is established. The server is the software component, the client is an application (or another component) making use of the component resources, and the contract is the respect of the pre- and post-conditions that must hold after the invocation of one of the component’s services.

Each component, implementing a model, must therefore come with its "contract", that is the conditions that identify its domain of applicability, and the limits to the use of its results. This allows developing a more specific suite of unit tests (the documentation of the components must contain at least some of the unit tests performed during implementation). All of this contributes to the transparency of the modelling solution.

Test of pre- and post-conditions is implemented by overloading the component API (client driven choice); however, if an unhandled exception occurs, the test of pre-conditions is run and an informative message describes the error and model and component source of the exception, allowing for continuing execution of the client according to a user choice. It is evident that this design choice satisfies the requirement for components to be robust, degrade gracefully, and raise appropriate exceptions.

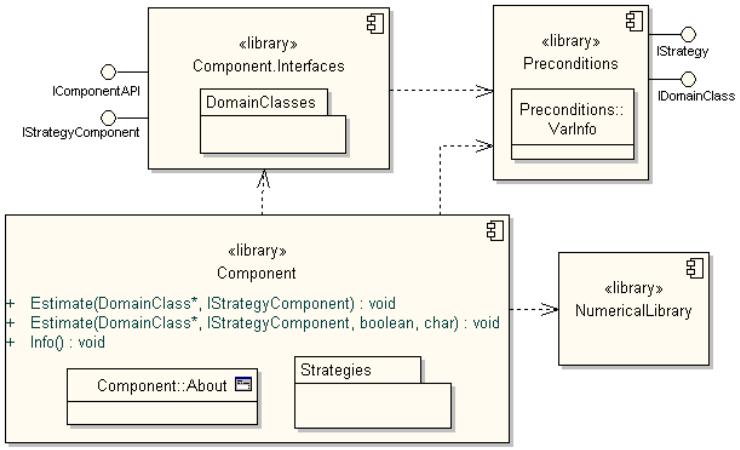


Figure 2. Example of a component diagram for the proposed architecture.

2.3.2 Tracing

Being able to closely inspect the behaviour of a component is a powerful tool to ascertain its quality. Traceability is therefore a major quality requirement. The traceability of component behaviour is implemented in the version of the components implemented in C# and running on the NET platform using the `System.Diagnostic.TraceSource` class, in one implementation that allows setting the listeners by the client. Various levels of tracing (critical, error, warning, information, and verbose) can be hence pooled in one or more listeners with all traces from other components and from the client. In Java components this is obtained using a logger. This satisfies the requirement for traceability of the component execution.

2.3.3 Technology

Sometimes, once a solid design has been cast, technology is seen as of secondary importance. Unfortunately this is not the case, and technology and its evolution often drives the design process in a tightly connected feedback loop

We used the Microsoft .NET 2.0 framework to implement our components. However, the object model of .NET allows easy migration to the Sun Java platform. Such migration has been actually made for some of the components referenced. Although the same design can be implemented under different platforms, the advantages of a memory managed environment combined to the object model of .NET or Java makes these platforms a first choice. Specific routines can still be written using "unsafe" C++ (as opposed to "managed" C++ in .NET) to optimize performance, but this should be seen as the exception not to give up all the advantages of a managed memory in complex systems using components with different origin.

3. PROOFS OF CONCEPT

Components implementing the solutions above have been made available for public use and other are being developed. The design has been tested on static [Acutis et al., 2008];

[Carlini et al., 2006]; [Confalonieri et al., 2008]; [Donatelli et al., 2006a and b] and dynamic biophysical models [Acutis et al., 2007]; [Trevisan et al., 2007], on agro-management models [Donatelli et al., 2007] and on statistical indices [Bellocchi et al., 2008]. Use of components has been done on applications (desktop and web, including web services) and via frameworks such as TIME (not published) and Modcom [APES, 2007]. Components can be used directly in applications as shown in the sample projects provided in the software development kits, of via adapters in modelling frameworks, as shown in the component diagram of Fig. 3.

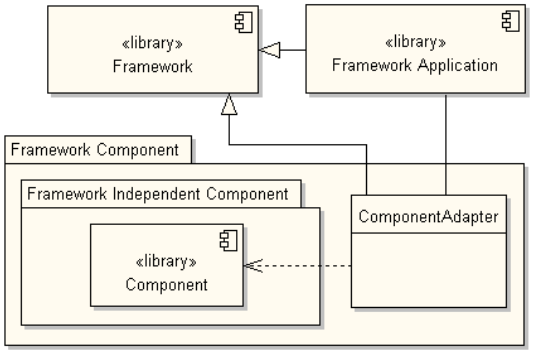


Fig. 3 Using a framework independent component in a model framework

4. CONCLUSIONS

Shifting the focus from modelling frameworks to a component design that follows the component oriented design results in a greater chance for model re-use. Framework independency stimulates model developers who do not feel constrained by a dependency to groups which develop specific frameworks; at the same time components can be easily used in several frameworks via simple wrappers. Design choices related to the modularity of model implementation and the provision of an explicit ontology for interfaces increases the transparency of the model construct and allows sharing knowledge in quantitative and usable terms. Several functionalities can be easily used to various extents by different clients. The design presented allows for extensibility by third parties which can then build on the domain description and models made available.

ACKNOWLEDGEMENTS

This publication has been funded by the SEAMLESS integrated project, EU 6th Framework Programme for Research, Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract no. 010036-2). We are in debt to Riccardo Wenger for the implementation of the CLIMA and the Preconditions components in Java. We also gratefully acknowledge Rob Knapen for his work in testing the performance of options for components interfaces in the Java platform.

REFERENCES

Acutis, M., P. Trevisiol, A. Gentile, D. Ditto, and L. Bechini, In M. Donatelli, J. Hatfield, and A. Rizzoli, editors, *Farming Systems Design 07, Methodologies for Integrated Analysis of Farm Production Systems*, volume 2, page 201. ESA, 2007.

Acutis, M., M. Donatelli and G. Lanza Filippi, PTF: an extensible component for sharing and using knowledge on pedo-transfer functions. *These Proceedings*.

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)

APES, Agricultural Production and Externalities Simulator, http://www.apesimulator.org; http://www.apesimulator.org/help/apes/ help on-line.

Argent, R.M., An overview of model integration for environmental applications – components, frameworks and semantics. *Environmental Modelling and Software* 19(3), 219-234, 2004.

Argent, R.M. and A.E. Rizzoli, Development of multi-framework model components. In: Pahl-Wostl C., Schmidt S., Rizzoli A.E., Jakeman A.J. (Eds.), *Trans. of the 2nd biennial meeting of the International Environmental Modelling and Software Society*, Osnabrück, Germany, vol. 1, 365-370, 2004.

Bellocchi, G., M. Donatelli, M. Acutis, and E. Habyarimana,. A Component and Software Application for Model Output Evaluation. *These proceedings*.

Bishop, J., 2008. C# 3.0 Design Patterns, O'Reilly, 314 pp.

Carlini, L., G. Bellocchi, and M. Donatelli, Rain, a software component to generate synthetic precipitation data. *Agronomy Journal*, 98, 1312-1317, 2006.

Confalonieri R., G. Bellocchi, and M. Donatelli, A software component to compute agro-meteorological indicators. Submitted to: *Environmental Modelling and Software*, 2008.

Cwalina, K., and B. Abrams, Aggregate components. In: *Framework Design Guidelines: Conventions, Idioms, and Patterns for Reusable .NET Libraries*. Addison-Wesley, Courier in Westford, Massachusetts, USA. 235-271, 2006

David, O., S.L. Markstrom, K.W. Rojas, L.R. Ahuja and W. Schneider. The object modelling system. In: Ahuja L.R., Ma L., Howell T.A., (Eds.), *Agricultural system models in field research and technology transfer*. Lewis Publishers, Boca Raton, FL, USA, p. 317-344, 2002.

Del Furia, L., A. Rizzoli and R. Arditi, Lakemaker: A general object-oriented software tool for modelling the eutrophication process in lakes, *Environmental Software*, 10(1), 43-64, 1995.

DCC, Domain Class Coder. http://www.sipeaa.it/ASP/ASP2/DCC_GuestBook.asp, installation; http://www.apesimulator.it/help/tools/dcc/ help on-line, 2006.

Donatelli, M., G. Bellocchi, and L.Carlini, A software component for estimating solar radiation. *Environmental Modelling and Software*, 21(3), 411-416. 2006a.

Donatelli M., G. Bellocchi, and L. Carlini, Sharing knowledge via software components: models on reference evapotranspiration. *European Journal of Agronomy* , 24(2), 186-192, 2006b.

Donatelli, M., J. Bolte, F. van Evert and W. Wang. Which software designs for evolution. In: van Ittersum M.K., Donatelli M. (Eds.), Modelling cropping systems: science, software and applications. European Journal of Agronomy 18, 193-195, 2003.

Donatelli, M., L. Carlini, G. Bellocchi, M. Colauzzi, CLIMA: a component based weather generator. In Zerger, A. and Argent, R.M. (eds) *MODSIM 2005 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, December 2005 ISBN: 0-9758400-2-9, 627-633, 2005

Donatelli M., E. Ceotto, and M. Acutis, 2007. A software component to simulate agro-management, In M. Donatelli, J. Hatfield, and A. Rizzoli, editors, *Farming Systems Design 07, Methodologies for Integrated Analysis of Farm Production Systems*, volume 2, page 217. ESA, 2007.

Jones, J.W., B.A. Keating, C.H. Porter. Approaches to modular model development. Agricultural Systems 70, 421-443, 2001.

MCE, Model Component Explorer. http://www.sipeaa.it/ASP/ASP2/MCE_GuestBook.asp installation; http://www.apesimulator.it/help/tools/mce/ help on-line, 2006.

Meyer, B., Applying "Design by Contract", *IEEE Computer*, 25(10), 40-51, 1992.

Rizzoli A.E, M. Donatelli, I. Athanasiadis, F. Villa, and D. Huber, . Semantic links in integrated modelling frameworks. *Mathematics and Computers in Simulation*, doi:10.1016/j.matcom.2008.01.017, in press.

Rizzoli A.E., Leavesley G. et al. Integrated modelling frameworks for environmental assessment and decision support. In Jakeman, A.J., Voinov A., Rizzoli A.E., Chen, S. (eds) *Environmental Modelling and Software*. Elsevier, in press.

Szypersky, C., D. Gruntz and S. Murer, *Component software - beyond object-oriented programming*. 2nd Ed. Addison-Wesley, London, United Kingdom, 2002.

Trevisan M., A. Sorce, M. Balderacchi, A. Di Guardo, 2007. A software component to simulate agro-chemicals fate. In M. Donatelli, J. Hatfield, and A. Rizzoli, editors, *Farming Systems Design 07, Methodologies for Integrated Analysis of Farm Production Systems*, volume 2, page 231. ESA, 2007

Van Evert F., A. Lamaker, 2007. The MODCOM framework for component-based simulation. In M. Donatelli, J. Hatfield, and A. Rizzoli, editors, *Farming Systems Design 07, Methodologies for Integrated Analysis of Farm Production Systems*, volume 2, page 223. ESA, 2007.

Zeigler, B.P., *Object-oriented simulation with hierarchical modular models*. Academic Press, 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

Conceptual Model of Single Information Space for Environment in Europe

J. Hřebíček[a], **K. Kisza**[b]

[a] *Institute of Biostatistics and Analyses, Masaryk University, Kamenice 126/3, 625 00 Brno, Czech Republic (hrebicek@iba.muni.cz)*

[b] *Faculty of Informatics, Masaryk University, Botanická 68a, 602 00 Brno, Czech Republic (kisza@iba.muni.cz)*

Abstract: The paper presents the further development of the conceptual model of the Single Information Space for the Environment in Environment (SISE) and compares this with an upper ontology concept of the SISE. The developed conceptual model of the SISE enables an implementation the vision of development an integrated, modern, common, shared and sustained Single European Information Space infrastructure for environmental information exchange and environmental management in Europe.

Keywords: Environmental data; Environmental information; Knowledge; Information space for environment; Conceptual model.

1. INTRODUCTION

The development of the Single European Information Space has been the first aim of the European i2010 strategy, [i2010, 2005] since 2005. The objective "ICT-2007.6.3: ICT for Environmental Management and Energy Efficiency" of the Seventh Research Framework Programme (FP7) specified this aim for the area of environmental protection and sustainable development. The aim of the development of the Single Information Space for Environment in Europe (SISE) was introduced by Schouppe [2008]: an Information and Communications Technology (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. This aim of SISE is to provide some sort of integrated information space in which environmental data and information are combined with knowledge. This infrastructure will enable a "holistic view" and allow the processing of different types of environmental data and information to extract more knowledge for decision making (correlating information and data) that are not currently possible. Ongoing developments in the context of thematic environmental legislation of EU are increasingly recognising the need to adopt a more modern approach to the production, exchange and use of environmental data and information. Full attention will be on the optimisation of complex data flows across all decision levels, across borders and sectors in developed Shared Environmental Information System (SEIS) by the Go4 team (DG Environment, EEA, Eurostat and JRC) [COM46 final, 2008], [SEIS, 2008].

The development of the complete and complex SISE covering all interactions among environmental data, information and knowledge using current ICT tools is practically impossible [Pillmann et al., 2006]. There is very fast growth of amount of data, information and knowledge all over the world each day. Therefore, it is appropriate to develop a common methodology of building the basic conceptual model of the SISE, which enables the common overview on environmental data, information and knowledge in standardized way. This paper takes into account some of above mentioned challenges of the SISE and the SEIS and presents and compares two concepts: an upper ontology approach and

author's conceptual model approach. The conceptual model of the SISE issues from ideas previously published by Hřebíček and Sluka [2003], Pillmann et al. [2006], Hřebíček et al. [2007], Hřebíček and Ráček [2007], and Schouppe [2008] and the paper summarizes the current results of author's team in the research of this topic.

2. MOTIVATION

Let us consider for a basic model of the SISE using of the upper ontology concept [Niels and Pease, 2001], [Pease, 2003], [Batres et al., 2007], [Villa, 2007]. The upper ontology concept offers fundamental structure and rule sets according to which is to build domain ontology models (for example ontology for medicine, financial engineering, etc) to achieve their compatibility. The domain ontology models for complete area of environment are not developed yet.

The general ontological model includes four basic elements: *individuals*, *classes*, *attributes* and *relations* and they are described in details e.g. in Wikipedia [2008].

Let us have two domain ontology models built according to upper ontology principles, then it is secured their compatibility in the most general form, i.e. there are no different definitions of equivalent classes in these models. The upper ontology concept is not suitable for more specific tasks (for example for the public access to environmental factors or effects), where the detailed level of the solution is needed. In this case it is necessary to use more detailed model, which is of much specific than the upper ontology model. If we have the class that provides the extraction of specific information type, (particularly for environmental information of factors or effects used for the decision making support), it is necessary to find out the primary data used for the extraction of this information. There are two basic possibilities for solving above problems:

- To define the class containing both primary data, and procedures for the information extraction in its attributes (this way is very unpractical – it could be represented by huge data and information aggregations which are very difficult to process – apart from data duplicity).
- To define special relations and classes enabling to form the information structure in the model (during the processing of such model it is necessary to know which relations and which classes were used and their context).

We developed the conceptual model formalized this second approach. Generally, there are many ways of formalization. We started from the object-oriented model of environmental information described by Hřebíček and Sluka [2003]. Hřebíček and Ráček [2007] have developed the basic conceptual model of the SISE and presented it on the conference ISESS 2007 in Prague. This version was generalised by Hřebíček et al. [2007], and the last renewed one is presented in the next chapter.

2. CONCEPTUAL MODEL

2. 1 Single Information Space for Environment in Europe

Let us consider a *network of constituents* consisting of four principal sets $\underline{I}$, $\underline{M}$, $\underline{O}$ and $\underline{A}$ which represent *universum* of classes of *information*, *methods*, *objects* and *attributes*, [Hřebíček and Ráček, 2007], [Hřebíček et al, 2007]. Every class of attributes, objects, methods and information includes also *meta-data description*. All four sets have a *tree structure* (continuous acyclic directed graphs), where nodes (classes) in lower tree layers inherit their structure (including also meta-data) from nodes in upper layers (Figure 1),.

Definition: Let us define the SISE as a quintuple: $S = [I, M, O, A, R]$ where I, M, O and A are domains of classes – tree structured – and R is a set of constituent inheritance and constituent aggregation relations (relation of inheritance, parent, child, predecessor, successor).

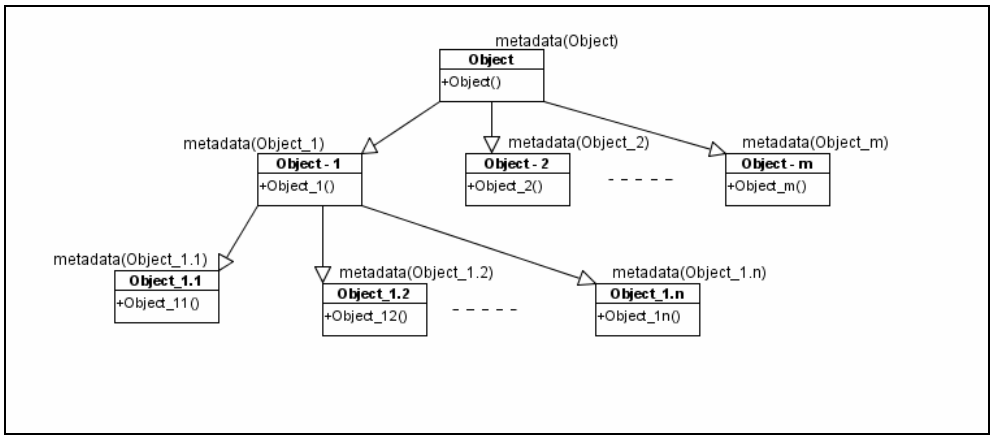


Figure 1: Common structure of object tree.

Following terms are used in the definition of the SISE:

- The tree is the continuous acyclic directed graph – the radical tree.
- The top element *v* is denoted as the *root* of the radical tree. The root *v* is the element, to which any edge is headed, there are only edges headed from it (at least one edge).
- The root is an element that represents the basic (the most general) type, from which are derived all other elements in the tree – *relation of inheritance.*
- For each element *w* is defined the *parent* element *p* as the element which is connected by *edge (p, w)* heading from *p* to *w* (every element different from the root element has only one parent). For the element *w* is defined the *child* element as any element *q*, which is connected by *edge (w, q)* heading from *w* to *q*. Each element except the leaf has at least one child (is able to have also more children). The leaf has no child.
- Any tree element in the tree with basic type of the root *v* is either empty structure or element of the type *v*, which is connected with the finite number of disjoint tree structures with the basic type *v* (mark them as *subtree*).

Lemma 1: If *p* is the child of *q* and at the same time *q* is the child of *r*, then *p* is child of *r*.

Assumption 1: Let sets *I*, *M*, *O* and *A* are defined as domains of *information, methods, objects* and *attributes* and we assume, that the constituent aggregation is hidden in next mandatory rules:

- For every $i \in I$ exist set M' and relation *r* where $M' \subseteq M$ and *r(i, M')* is valid;
- For every $m \in M$ exist set O' and relation *r* where $O' \subseteq O$ and *r(m, O')* is valid;
- For every $o \in O$ exist set A' and relation *r* where $A' \subseteq A$ and *r(o, A')* is valid.

The simple example of the above structure is presented on Figure 2.

2.2 Representation of information (knowledge)

Let us consider:

- Information is represented as a continuous acyclic directed graph – *four-leveled radical graph.*
- The top element *v* is denoted as the root of this radical graph. The *root v* is the element, to which any edge is headed, there are only edges headed from it. The top element belongs to *I* set.

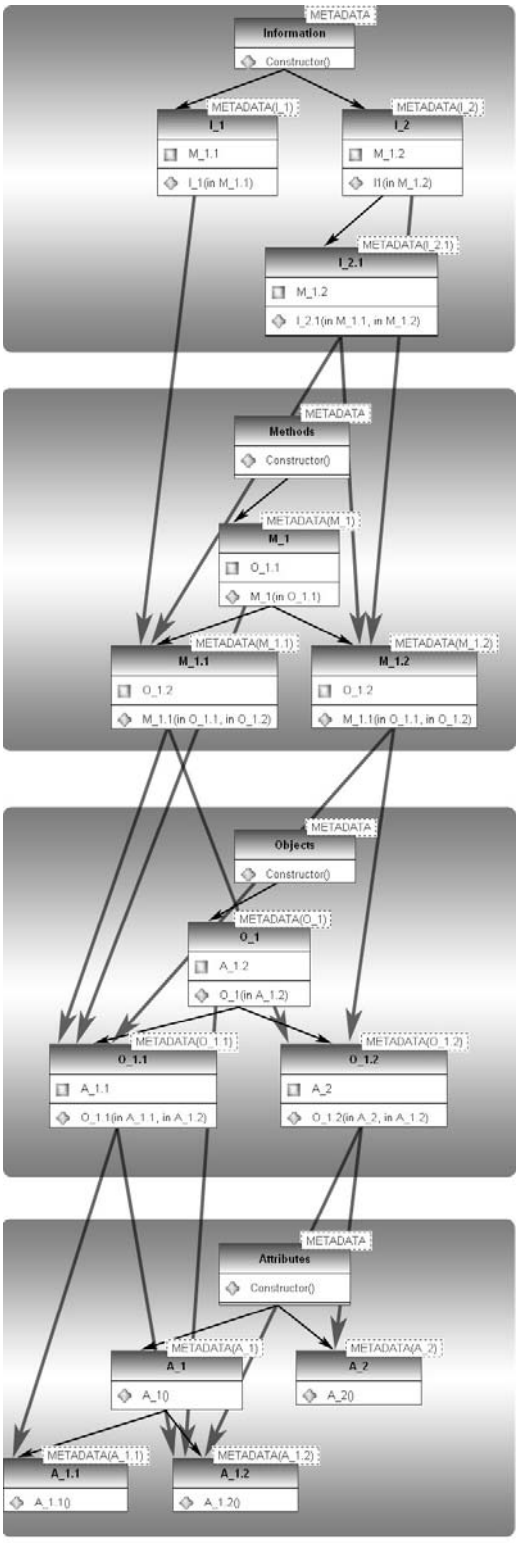


Figure 2: Basic structure of Environmental Information Space.

- There will not be used the relation of inheritance for purposes of this graph, but new relation *R* is defined: *R:* M^n -> *I*, or O^n -> *M*, or A^n -> *O*.

- The root element of each relation is just one element from set *I* (i.e. the first graph level). Its successors (i.e. the second graph level) are elements from set *M*. The successors of these elements from set *M* (i. e. the third graph level) are elements from set *O*. The successors of these elements from set *O* (i.e. the fourth graph level) are elements from set *A*.
- For each top element *w* is defined the *predecessor* element *p* as the element, which is connected by the *edge (p, w)* heading from *p* to *w*. For each top element *p* is defined the *successor* element *w* as the element, which is connected by *edge (p, w)* heading from *p* to *w* (is able to have also more successors).

Lemma 2: It is valid:

- If *w* belongs to the *A*, then all its successors belong to the *O*.
- If *w* belongs to the *O*, then all his successors belong to the *M*.
- If *w* belongs to the *M*, then all his successors belong to the *I*.
- If *w* is successor of *u*, then for no child is *w* his successor (i.e. w is not successor of no element which is element of the subtree, where the root is *u*).

3. IMPLEMENTATION IDEA

3.1 Semantic web services

The base sets of elements of the SISE are elements on the different implementation level in the graph representing concrete information. Implemented elements are the same (constant) for all users (primary data, a mathematical definition of methods, etc). Unimplemented elements are represented like interfaces which are necessary to implement according to concrete user possibilities, without necessity to know whole "domain knowledge". Both implemented elements and interfaces represent "domain knowledge". They are interconnected with the ontology representing "operational knowledge".

Information is represented by the graph that contains implemented elements and unimplemented elements. Unimplemented elements are defined as common interfaces with input and output limitations. Every such element can be implemented like service. Type of those services refers to the type of the element (belonging to base set):

- *M – method* – a service enables obtaining information from aggregate data (elements of set *O*).
- *O – objects* – a service enables data access.
- *I – information* – a service enables the aggregation of basic information – outputs of methods – for obtaining information in the way requested by user.

Any information can be present like a process *T* that has defined the plan, which has to be executed to achieve the requested result (information). The process *T* is the complex process that is able to contain sub-processes that must be scheduled and executed (elements of sets *M, O, A*). Each of these sub-processes has its input and output limitation (characteristics). Each of those subtasks has preconditions and post conditions that must hold, as does *T* itself. Those preconditions and post conditions are properties that must hold either prior to or posterior to the planning, scheduling, or execution of that task/subtask. There are constraints on those phases of the complex service enables to get complex information (the element of *I*).

We can apply this model to describe any information, data or service and their relations in SISE. It enables contextual reasoning and the automated transition (coercion) of semantics from one context to another. This approach enables integration of the data and metadata associated with a service together with specifications of its properties and capabilities, the interface for its execution, and the prerequisites and consequences of its use.

For example in SISE we can describe the configuration for any information from its components according to a required specification of information as a quintuple *[R‘, I‘, M‘, O‘, A‘]*, where objects are connected with relations and:

- *A‘⊆A (i.e. I is a subset of I);*
- *O‘⊆O;*
- *M‘⊆M;*
- *I‘⊆I;*
- *R‘⊆R.*

These configurations cover the domain knowledge. They describe attributes, objects, methods and relations among them, necessary to get any information defined in SISE.

If we consider some area of interest we can take it as a quintuple *[R‘, I‘, M‘, O‘, A‘]* where elements are on various level of implementation. Implemented elements are the same for any user who use the SISE (mathematical definitions, common data …). Unimplemented elements are represented as interfaces which are necessary to implement according to possibilities of concrete user without need to know whole domain knowledge. We suggest implementing these interfaces as semantic web services. Both implemented parts and unimplemented parts represent the domain knowledge. They are connected by ontology relations represent the operational knowledge.

Generally, ontology defines a common vocabulary for researchers whose need to share information, data and services in domain. It includes machine-interpretable definitions of basic concepts in the domain and elations among them. The main advantages of using ontology in our model for SISE are that ontology enables mainly:

- to analyze domain knowledge and make domain assumptions explicit;
- to share common understanding of the structure of information among people or (and) specialized software technologies (computer agents). For example, we can suppose several different web sites (each with its own database) contain environmental information. If these web sites implement the same underlying ontology (defined in SISE), then computer agents can easily extract and aggregate related information from these different sites. This aggregated information can be used as an answer to user queries or as an input data to other applications.
- to reuse of this domain knowledge (after making analysis);
- to separate domain knowledge from the operational knowledge;
- to provide a large extent of flexibility and expressiveness, the ability to express semi-structured data and constraints, and support types and inheritance.

As an instrument for implementation of services we suggest use Web Ontology Language (OWL-S) [OWL-S, 2004]. OWL-S (formerly DAML-S) is an ontology of services that makes these functionalities possible. In this submission we describe the overall structure of the ontology and its three main parts: *the service profile for advertising and discovering services; the process model, which gives a detailed description of a service's operation*; and *the grounding, which provides details on how to interoperate with a service, via messages*, [OWL-S, 2004].

OWL-S takes a mostly reactive planning view of the semantics of web services. The reactive planning view means that services and their complex compositions are generally viewed as a three-phase operation: *planning, scheduling*, and *execution*. There is some set of objectives or goals that a developer or user wants to achieve. This set might be viewed as the rationale for the desired web service. One might have multiple plans (various compositions of web services) that could achieve those desired goals. A given plan is selected or composed from a library or registry of services/plans. That plan can be represented as a more-or-less complex task or process model, [SWS, 2005].

This model is suitable for our idea of tree representation of any information defined in S which was mentioned above.

To realize our vision of the SISE, it is necessary to implement theoretical model S in some pilot project and fully describe formalization processes and form recommendations for used technologies. If we concern SISE as an abstract upper model *S*, it will be necessary to define common Semantic constraints (preconditions and postconditions) for defined interfaces of services.

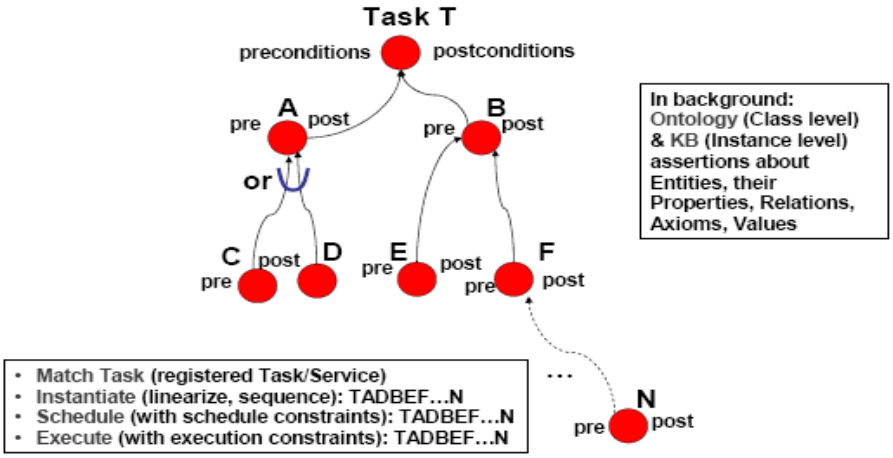


Figure 3. Planning, Task, Process Representation. Source [SWS, 2005].

4. CONCLUSION

Interaction among various digital data, information and knowledge sources are necessary for building SISE, e.g. ensure so - called semantic interoperability. Semantic interoperability is conceptual formulation of metadata structure that allows semantically combine data elements from different schematics, dictionaries and other sources and makes possible to search information across heterogeneous distributed data source. By the help of semantic interoperability are solved e.g. problems, when individual sources use various terms for description of the same term or on the contrary use same terms for various notions.

The technologies supporting semantic interoperability are very popular and exploited nowadays (semantic technology), especially then ontology, which compared to for example UML (Unified Modeling Language) offer some other benefits.

The amount of existing ontological models is rising very quickly. Mostly they are focused on some specified domain (domain ontology) or systems in companies (application ontology). To ensure interoperability of these systems, it is necessary to solve compatibility of systems and it covers as a first step – to standardize model for environmental data, information and knowledge.

The usage of the developed model will enable more exactly find out (identify) which elements (constituents) and in what way they have been used in the extraction of information any time, without necessity to know more details about the meaning of single relations and classes in the whole model context.

Nowadays it is just theoretical model which flows from earlier research and we try to verify it on implementation pilot in biodiversity and waste management area.

ACKNOWLEDGEMENTS

The paper is supported by the Ministry of Environment of the Czech Republic (project No. SP/4i2/26/07), further by the Ministry of Education and Youth (project INCHEMBIOL, No. MSM0021622412).

REFERENCES

Batres, R., Matthew West, M., Leal, D., Price, D., Masaki, K., Shimada, Y., Fuchino, T. and Naka, Y., An upper ontology based on ISO 15926, *Computers & Chemical Engineering*, 31(5-6), 519-534, 2007.

COM46 final, Towards a Shared Environmental Information System (SEIS), Commission of the European Communities, 2008.

Hřebíček, J., and Sluka, J., Model of the Object-Oriented Environmental Information - A New Approach in the Environmental Data and Information Management. In: *17. International Conference Informatics for Environment Protection. The Information Society and Enlargement of the European Union*. Marburg: Metropolis Verlag, Germany, 381-388, 2003.

Hřebíček, J., and Ráček, J., Environmental Information Space I. Theory and Practice. In: *6th International Symposium on Environmental Software Systems (ISESS 2007).* Guelph: IFIP WG 5.11, 123-133, 2007.

Hřebíček, J., Kisza, K., Štefaník, M., and Ráček, J., Environmental Information Space II. *3rd Internationl Summer School on Computational Biology*. Masaryk University, 158-170, 2007.

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>, 2008/04/30

Maedche, A., Ontology Learning for the Semantic Web. Kluwer Academic Publishers, N.Y., 2002

Niles, I., and Pease, A., Towards a Standard Upper Ontology. In: Proceedings of the 2nd International Conference on Formal Ontology in Information Systems, Chris Welty and Barry Smith, eds, Ogunquit, Maine, 2001

OWL-S, http://www.w3.org/Submission/OWL-S/, 2004

OWL, <http://www.w3.org/TR/owl-features/>, 2008/04/30

Pease, A., The Sigma Ontology Development Environment. In Working Notes of the IJCAI-2003 Workshop on Ontology and Distributed Systems, August 9, Acapulco, Mexico, 2003.

Pillmann, W., Werner Geiger, W. and Voigt K., Survey of environmental informatics in Europe, *Environmental Modelling & Software*, 21(11), 1519-1527. 2006

SEIS, http://ec.europa.eu/environment/seis/index.htm, 2008

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>, 2008/04/30

SWS, Frameworks for Semantics in Web Services Workshop, MITRE Position Paper, 2005

Villa, F., A semantic framework and software design to enable the transparent integration, reorganization and discovery of natural systems knowledge. *Journal of Intelligent Information Systems*, 29(1), 79-96, 2007

Wikipedia, Ontology (information science), <http://en.wikipedia.org/wiki/Ontology_(computer_science)>, 2008/04/30

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

Bridging the Gap between Geohydrologic Data and Distributed Hydrologic Modeling

G. Bhatt [a], **M. Kumar** [a] **and C.J. Duffy** [a]

[a] *Civil and Environmental Engineering, Pennsylvania State University, University Park, Pennsylvania, USA (gxb913/muk139/cxd11@psu.edu)*

Abstract: This paper outlines and demonstrates a strategy for coupling of integrated hydrologic model and Geographic Information System (GIS) to meet pre/post processing of data and visualization. Physically based fully distributed integrated hydrologic models seek to simulate hydrologic state variables and their interactions in space and time. The process requires interaction with a range of heterogeneous data layers such as topography, soils, hydrogeology, climate, and land use. Clearly, this requires a strategy for defining topology definitions, data gathering and development. Traditionally GIS has been used for data management, analysis and visualization. Integrated use and streamlineed development of sophisticated numerical models and commercial Geographic Information Systems (GISs) poses challenges inherited from proprietary data structures, rigidity in their data-models, non-dynamic data interaction with pluggable software components and platform dependence. Independent hydrologic modeling systems (HMSs), GISs and Decision Support Systems (DSSs) not only increase model setup and analysis time but they also result in data isolation, data integrity problems and broken data flows between models and the tools used to analyze their inputs and results. In this paper we present an open-source, extensible and pluggable architecture, platform independent "tightly-coupled" GIS interface to Penn State Integrated Hydrologic Model (PIHM) called PIHMgis. The tight-coupling between the GIS and the model is achieved by the development of PIHMgis shared-data model to promote minimum data redundancy and optimal retrievability [Kumar et al., 2008]. The procedural framework of PIHMgis is demonstrated through its application to Shaver's Creek Watershed located in Susquehanna River Basin in Pennsylvania.

Keywords: Geographic Information Systems (GIS); Hydrologic Model; Shavers Creek; Susquehanna River Basin.

1. INTRODUCTION

Physically based distributed hydrologic models simulate the spatio-temporal dynamics of the important hydrologic processes using spatially distributed watershed's physical properties and forcing fields [Feeze et al., 1969]. These models better represent natural heterogeneities [Entekhabi et al., 1989; Pitman et al., 1990] with the goal of enhancing our understanding and prediction of the spatio-temporal dynamics of hydrologic processes. Clearly, a key challenge in the development and use of distributed, physically based modeling frameworks is the large number of physical parameters that must be incorporated into the model. Geographic Information Systems (GISs) with their ability to handle both spatial and non-spatial data, and to perform data management and analysis operations have a strong potential to advance development and use of more complex modeling frameworks if used appropriately. A major deficiency of GIS that has been recognized is the lack of sophisticated analytical and modeling capabilities [Maidment, 1993; Wilson, 1996; Camara, 1999]. Likewise many, existing hydrologic models are not developed with data structures that facilitate close linkage to GISs and decision support systems (DSSs) [National Research Council, 139-63, 1999].

Prior efforts have implemented a range of different levels of coupling between a GIS and hydrologic models helping to elucidate the relative advantages and disadvantages of alternative coupling approaches in terms of representation of the watershed, watershed decomposition, sensitivity/uncertainty analysis, and parameter estimation as highlighted by Watkins et al. [1996]. Current GISs have limitations that impede coupling with hydrologic models [Abel et al., 1994; Kopp 1996]. Also since many of the advanced GISs are platform dependent, running mostly on Windows platform personal computers (PCs), they limit users from taking advantage of high performance computing architectures. Many commercial GIS framework suffer from closed data structures for GIS features, making it difficult to develop customized data manipulation/visualization tools that evolve with a modeler's/user's needs. Moreover, hydrologic models generally need other software support for pre- and post-processing tasks such as sensitivity analysis or decision support. The diverse needs of hydrologic research motivate the importance of developing coupled GIS and physical modeling systems able to incorporate more flexible tools and formats [Deckmyn et al., 1997].

In this paper, we demonstrate an integration methodology for an open source GIS framework and an integrated hydrologic model that enables users to take advantage of object oriented programming (OOP) to provide direct access to the GIS data structure, to better support efficient query and data transfer between the hydrologic model and GIS [Kumar et al. 2008]. The data structure has been designed to be flexible for modification and customization of the model or GIS, rich enough to represent complex user defined spatial relations and extensible to add more software tools as the need be. The "tightly-coupled" integrated GIS interface to Penn State Integrated Hydrologic Model (PIHM) has been created in the Open Source Quantum GIS [www.qgis.org]. The software framework used to create the tightly coupled PIHMgis system is generic and can be used in other model applications. Beyond describing the software framework for PIHMgis, this paper also demonstrates the importance and use of the framework for representing, modeling, visualizing and analysis data to Shaver's Creek Watershed in Susquehanna River Basin in Pennsylvania as case study.

2. INTEGRATION METHODOLOGY

2. 1 Introduction

Goodchild [1992], Nyerges [1993] and Sui et al. [1999] have discussed software integration strategies for GIS frameworks and models that range from loosely coupled to fully integrated systems. As discussed in Table 1, loose coupling, where a distinct GIS and model system exchange information using files, may be prone to data inconsistency, information loss and redundancy, leading to increased model setup time and post-processing. On the other extreme embedded coupling, where the model itself is developed in the GIS framework leads to a large and complex source code structure, which leaves the code inertial to change results it in being closed and isolated. Tight coupling preserves identity of GIS and hydrologic model behind the shared user interface and allows data exchange using shared data and method base.

Table 1. Different levels of integration between a GIS and a hydrologic model

Coupling Level → / Characteristics ↓	Loose	Tight Integration	Embedding
Shared User Interface	×	√	√
Shared data and method base	×	√	√
Intra-simulation Model Modification	×	×	√
Intra-simulation Query and Control	×	×	√

In this study we follow a tight coupling methodology [Kumar et al., 2008] based on the classification listed in Table 1. Tight coupling has the advantage of (1) preserving the advantage of independent development of various tool boxes as is the case for loose coupling and (2) the shared memory access to GIS data and model data linked through a carefully designed object oriented programming strategy for both the GIS and hydrologic model mimics many of the advantages of an embedded coupling system. As listed in Table 1, one of the major pre-requisites for tight coupling between GIS and a hydrologic model is to have a shared data and model structure base. Developing a shared data model for a GIS and a hydrologic model requires a careful consideration of both software systems and identification of connection points between them.

2.2 GIS Framework

The architecture of the GIS data model determines the ease of coupling a GIS with a hydrologic model. Generally a data model for a GIS includes constructs for spatial data, topological data and attribute data [Nyerges, 1987]. Data structures and associated descriptive constructs used in the data-management subsystem of GIS can lead to efficient data storage, editing and retrieval, and definition of new customized feature object representations within a GIS and integrated hydrologic model. This implies that a data model with its data, rules and relationships base can be a suitable basis for supporting GIS applications as well as hydrologic modeling. The data structure of the integrated hydrologic model will be then determined by the type and properties of the data models used in GIS. One of the pre-requisites for a tightly coupled integration, based on a shared data model will be access to the GIS architecture. Open source access to a GIS's architecture facilitates the development and use of GIS classes and methods while also providing the interface and linkages necessary for tight coupling. In this study, Quantum GIS (QGIS) is used as the base GIS system which is tightly coupled with PIHM. QGIS is open-source GIS and has been developed in C++, C and Qt (http://trolltech.com/), which makes it attractive as a base framework to develop a model interface.

2.3 Hydrologic Modeling Framework

The hydrologic processes incorporated in the model require data coverage sets of physical properties and system states at time t to predict system states the results at $t + \Delta t$. Δt is adaptively determined depending on the time scales of the interacting processes at each time t. In this study, we present PIHM [Qu and Duffy, 2007] tightly coupled with QGIS. PIHM is a physically-based, distributed hydrologic model that uses a finite volume formulation for the governing physical equations and constitutive relationships interacting on and across the unit elements of the decomposed domain. The governing physical equations generally represented by partial differential equations (PDEs) are discretized in space using the 'method of lines' [Leveque, 2002] approach to reduce them to ODEs. Figure 1 shows a typical "kernel" defined for triangular and linear prismatic elements along with the interacting physical processes to be coupled in the model. The kernel is designed to capture dynamics of multiple processes while maintaining the conservation of mass at all cells, as guaranteed by the finite volume formulation [Leveque, 2002].

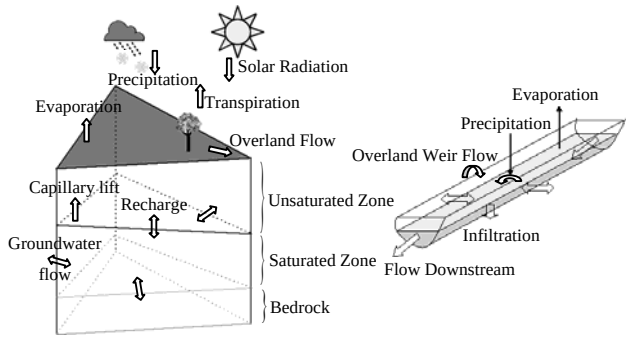


Figure 1. Interacting hydrologic processes on each prismatic element (left) and on each linear river element (right) [modified from Qu and Duffy, 2007]

The PIHMgis framework developed in this study supports the organization, development and assimilation of the extremely large set of spatial and temporal data for each model cell and its neighbors. With the shared data model, relationships and schemas between GIS and hydrologic model, tight coupling leads to an integrated system where GIS is simply another option to generate addition state and output variables in the model and to provide additional management, analysis and visualization options while the hydrologic model becomes one of the analytical functions of the GIS.

3. PIHMgis INTERFACE

PIHMgis is an integrated and extensible GIS system with data management, data analysis, data modelling, unstructured mesh generation and distributed PIHM modelling capabilities. The underlying philosophy of this integrated system is a shared geo-data model between GIS and PIHM thus making it possible to handle the complexity of the different data models, representation structures and model simulations. PIHMgis has been developed using basic QGIS source code. The Graphical User Interface (GUI) component have been designed in Qt [http://trolltech.com/products/qt], which is a standard framework for high-performance, cross-platform graphical widget toolkit development while the algorithms for several modules and the hydrologic model PIHM have been implemented in C and C++.

PIHMgis interface is procedural and interactive. Figure 2 shows a snapshot of PIHMgis interface. More snapshots are available at http://www.pihm.psu.edu/ >> PIHMgis >> Documents >> Users Guide. “Help” guides the user in selecting control parameters, the underlying algorithm through each PIHMgis module. Modularity is achieved via the plugin architecture which provides a mechanism for third parties to extend the QGIS core application.

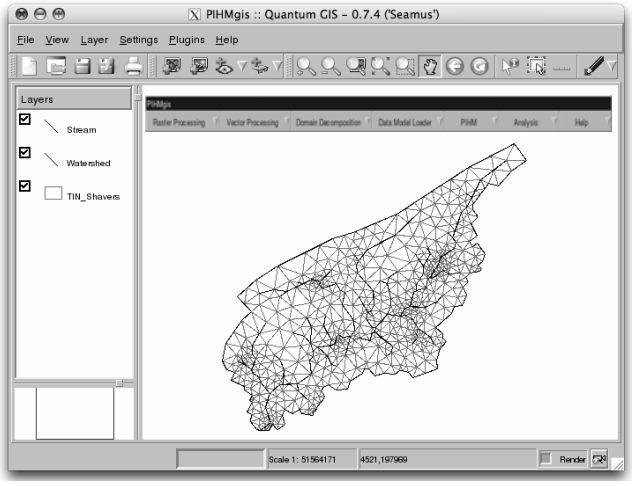


Figure 2. PIHMgis Interface. The data frame shows the unstructured domain decomposition for Shaver’s Creek Watershed.

Architectural framework of PIHMgis shown in Figure 3 outlines the functionalities provided by the framework. Tight coupling shares the user interface between the GIS and the modelling framework. Direction of the arrow in shows the possible data flow within the framework. All the modules of PIHMgis have been organized in a procedural structure. Procedural framework of PIHMgis has been categorized into six processing stages. Raster Processing modules facilitate stream

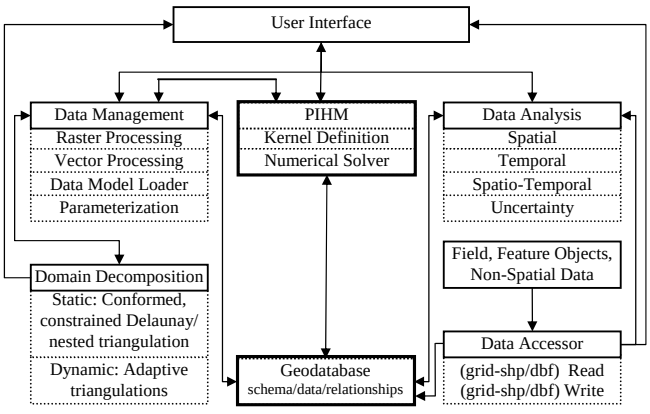


Figure 3. Architectural Framework of PIHMgis showing functionalities and conceptualization; direction of arrow shows the possible data flow within the

definition and watershed delineation. The Vector Processing aids users in defining watershed properties using nodes, polygons and polylines which eventually serve as domain constraints. The domain constraints are used to generate constrained Delaunay triangulations with certain mesh quality criteria. Before solving the finite volume based system of ODEs using RunPIHM module, the model parameters associated with soil, land cover as well as forcing and boundary conditions are assigned to each triangular and stream element in automated fashion in Data Model Loader modules. Finally, statistical and other kind of spatial and temporal data analysis and visualization can be performed to the model output using Analysis modules.

4. PIHMgis APPLICATION: CASE STUDY

PIHMgis takes advantage of the fact that modern geohydrologic datasets are stored and distributed in the form of a geodatabase [Arctur and Zeiler, 2004]. PIHMgis facilitates easy and accurate data development leading to easy model setup, model run, analysis and visualization. To demonstrate the procedural framework a case study application of PIHMgis to Shavers Creek Watershed located in Susquehanna River Basin is discussed in this section.

4.1 Raster Processing

Raster Processing facilitates stream definition and watershed delineation from the Digital Elevation Model (DEM) of the modelling domain. It is executed in a procedural framework involving computation of: (1) Fill Pits Grid; (2) Flow Grid; (3) Flow Accumulation Grid; (4) Stream Grid; (5) Link Grid; and (6) Catchment Grid. Figure 4A shows the 30 meter DEM of the Shavers Creek and Figure 4B shows the catchment polygon and stream polyline feature obtained after Raster Processing modules. A threshold of 2000 grids was applied to Flow Grid computed using d8 algorithm [Tarboton, 1991] for stream definition.

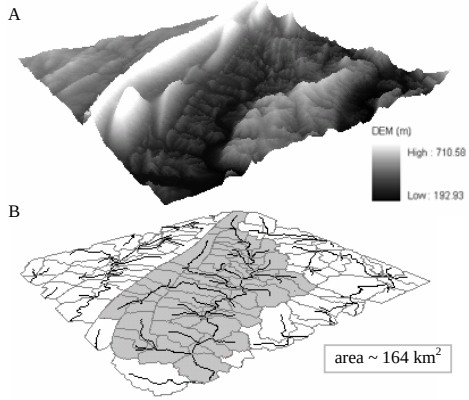


Figure 4. A shows DEM of Shaver's Creek. **B** Catchment and stream feature obtained after Raster Processing

4.2 Vector Processing

Geohydrologic features such as soils, land cover and other physiographic coverages can be used as constraining layers for the purpose of decomposition or mesh generation of modelling domain in addition to features generated using Raster Processing modulated by the modelling purpose. However, Stream polylines and catchment polygons obtained using Raster Processing retains the signature of the grid used, which of course depends on the DEM resolution. Vector Processing modules address issues specifically pertaining to modelling exercise as it allows development of a GIS layer which contains all the information of preferentially simplified constraining layers enabling efficient and quality

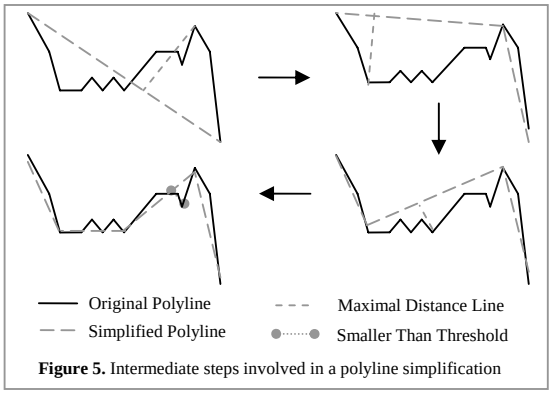


Figure 5. Intermediate steps involved in a polyline simplification

domain decomposition for modeling.

Features of the type polygon or polyline contain fluctuations or extraneous bends. Preferential simplification is a crucial module, part of Vector processing which simplifies the feature by eliminating nodes responsible for those fluctuations and still preserving the essential shape of the feature using simplification algorithm [Douglas and Peucker, 1973] as shown in Figure 5.

4.3 Domain Decomposition

PIHM uses vertical projection of triangular irregular mesh to form a local control volume which facilitates better representation of terrain [Kumar et al., 2008]. TRIANGLE [Shewchuk, 1996] has been integrated to decompose the domain into a high-quality, constrained, conforming Delaunay triangulation. TRIANGLE uses Ruppert [1995] and Chew [1993] algorithms for triangulation to generate non-skinny triangles and enforces the user selected quality constraint to the constraining layer prepared after Vector Processing modules.

In this study only the external boundary of the watershed is considered. A simplification tolerance of 200 meters was applied to both the watershed and stream feature. The decomposed unstructured mesh for the modelling domain is shown in Figure 2, where a 23 degree minimum-angle quality constraint was used.

4.4 Data Model Loader

The shared geo-data model contains all the topological and relational information needed to represent the modelling domain as well as the geohydrologic data needed for model parameterization. The Data Model Loader modules enrich the geodatabase defined by the classes and relationship of the shared geo data base. Several algorithms have been incorporated in Data Model Loader modules to facilitate topology assignment and model parameterization related to each triangular element and river segment.

An element is defined by the collection of three nodes in relation to the decomposed domain. Each element is assigned with a representative parameter value corresponding to each geohydrologic data layer along with nodes and neighbour information as shows in Figure 6A. Where as, a river segment is identified as one of the edge of an element, therefore defined by two nodes. Topology for channel segments is defined by From Node, To Node, Downstream segment, Left triangular element, Right triangular element [Figure 6B].

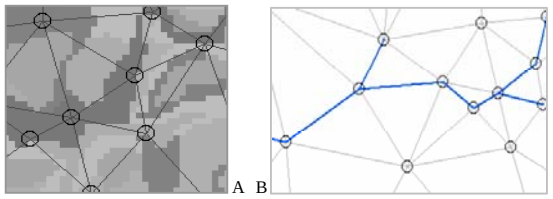


Figure 6. A. Data parameterization of each layer for each representative element. **B.** Topology for each stream segment using nodes and elements

4.5 Run PIHM

RunPIHM module embraces the PIHM and facilitates its execution right from the GIS framework. PIHM uses semi-discrete finite volume approach to reduce the governing PDEs into ODEs. The local system of ODEs defined on the each unit element and linear stream segments are assembled over the entire modelling domain forming a global system of ODEs. A state-of-art stiff-ODE solver SUNDIALS [http://www.llnl.gov/CASC/sundials/] is used to solve the global system of ODEs. RunPIHM module directly interacts with the geodatabase previously enriched by the Data Model Loader modules to retrieve all the topologic and geohydrologic model parameters. As simulation progress all the spatio-temporal model simulated data feeds back the geodatabase in the Network Common Data Form (NetCDF) format.

4.6 Analysis

PIHMgis modules discussed in section 4.1 to 4.5 provides easy data development, efficient model setup, and model execution. PIHMgis also provides modules to meet specific need for analysis and visualization of model simulated data in addition to basic GIS functionalities of QGIS. RunPIHM provides several optional parameters for the purpose of model calibration. However in this paper no calibration has been performed as part of model simulation. That is, data used in the simulation can be considered as a-priori information from independent sources. The Time Series module allows visualization of time series of model simulated parameter [described in Figure 1]. Figure 7 shows a time series plot of saturation averaged over the whole domain. Spatial Plot module allows creation of spatial maps as time series doesn't provide any information regarding spatial distribution of any simulated parameter. Figure 8 shows the spatial distribution for the annual average soil saturation. Since the motivation behind analysis of simulated results may vary widely depending on modelling interest, it is necessary that PIHMgis have extensible and pluggable architecture which allows easy addition of customized analysis and visualization modules.

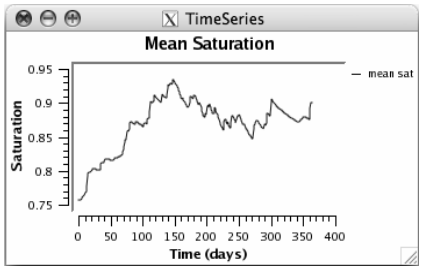


Figure 7. Spatial mean saturation over the entire domain for 1yr.

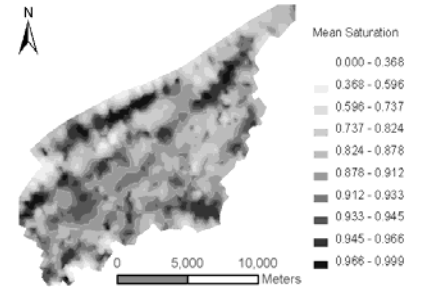


Figure 8. Mean annual saturation over the entire domain.

5. CONCLUSIONS

Isolated and independent hydrologic models and pre-processing (input data preparation) and post-processing (analysis and visualization), leads to increased model setup time and errors due to broken data flow. PIHMgis uses a tightly-coupled GIS framework which is based on shared-geo-data model to bridge hydrologic model and geohydrologic data (GIS framework). It offers a strategy for integration of modelling, analysis and visualization of complex multidimensional geohydrologic and land surface information.

Open source development of PIHMgis provides transparency, free access, modification to the source code. PIHMgis source code documentation is available at http://www.pihm.psu.edu/pihmgis_documents.html. The tight coupling strategy leaves the frameworks extensible and allows independent development. Moreover, the procedural framework of PIHMgis provides ease of use and preserves independence of each module at the same time.

REFERENCES

Abel, D.J., and Kilby, P.J., The systems integration problem, International Journal of Geographical Information Systems, 8(1):1-12, 1994.

Arctur, D., and Zeiler, M., Designing Geodatabases: Case Studies in GIS Data Modeling, ESRI Press, 2004.

Camara, G., Souza, R., Monteiro, A., Paiva, J., and Garrido, J., Handling complexity in GIS interface design, In Brazilian Symposium in Geoinformatics, Geoinfo, 1999.

Chew, P.L., Guaranteed-Quality Mesh Generation for Curved Surfaces, Proceedings of the Ninth Annual Symposium on Computational Geometry (San Diego, California), pages 274-280, Association for Computing Machinery, 1993.

David, D., Peucker, T., Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, The Canadian Cartographer 10(2), 112-122, 1973.

Deckmyn, J., Paniconi, C., Giacomelli, A., and Kleinfeldt, S., A Modular Approach to Integrated Environmental Modeling Systems Incorporating GIS and Visualization Functionality, CRS4-TECH-REP-96/26, 1997.

Entekhabi, D. and Eagleson, P.S., Land surface hydrology parameterization for atmospheric General Circulation models including subgrid scale spatial variability, Journal of Climate, 2(8), 816-831, 1989.

Freeze, R.A., and Harlan, R.L., Blueprint for a physically-based digitally simulated, hydrologic response model, J. Hydrol., 9, 237– 258, 1969.

Goodchild, M.F., Geographical information science. International Journal of Geographical Information Systems 6(1): 31-45, 1992.

Kopp, S.M., Linking GIS and hydrological models: where have we been, where are we going? In: HydroGIS 96: Application of Geographic Information Systems in Hydrology and Water Resources, 13-21. IAHS Publ. no. 235, 1996.

Kumar M., Bhatt, G. and Duffy, C.J., An Object Oriented Shared Data Model Coupling. International journal of Geographic Information System, In preparation, 2008.

Kumar M., Bhatt, G. and Duffy, C.J., An efficient domain decomposition framework for accurate representation of geodata in distributed hydrologic models, In preparation, 2008.

Leveque, R.J., Finite Volume methods for hyperbolic problems. Cambridge University Press, 2002.

Madsen, N.K., The method of lines for the numerical solution of partial differential equations (abstract), ACM SIGNUM Newsletter, v.10 n.4, p.5-7, 1975.

Maidment, D.R., GIS and hydrological modeling. In M. F. Goodchild, B. Parks, & L. Steyaert, (Eds.), Environmental modeling with GIS (pp. 147–167). New York: Oxford University Press, 1993.

Nyerges, T.L., GIS Research Issues Identified During a Cartographic Standards Process, Proceedings, of International Symposium on Geographic Information Systems, 1987.

Nyerges, T.L., Understanding the scope of GIS: Its relationship to environmental modeling. In Environmental Modeling with GIS, edited by Goodchild, M.F., Parks, B.O. and Steyaert, T. (New York: Oxford University Press), pp. 75-93, 1993.

Pitman, A.J., Henderson-Sellers, A., and Yang, Z.L., Sensitivity of regional climates to localized precipitation in global models, Nature, 346, 734-737, 1990.

Qu, Y., Duffy, C.J., A semidiscrete finite volume formulation for multiprocess watershed simulation, Water Resour. Res., 43, W08419, doi:10.1029/2006WR005752, 2007.

Ruppert, J., A Delaunay Refinement Algorithm for Quality 2-Dimensional Mesh Generation, Journal of Algorithms 18(3):548-585, 1995.

Shewchuk, J.R., Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator, in Applied Computational Geometry: Towards Geometric Engineering, volume 1148 of Lecture Notes in Computer Science, pages 203-222, Springer-Verlag (Berlin), 1996.

Sui, D.J. and Maggio, R.C., Integrating GIS with hydrological modelling: practices, problems and prospects, Computers, Environment and Urban Systems, 23, pages 33-51, 1999.

Tarboton, D.G., Bras, R.L., Rodriguez-Iturbe, I., On the Extraction of Channel Networks from Digital Elevation Data, Hydrologic Processes, 5(1): 81-100, 1991.

Watkins, D. W., McKinney, D. C., Maidment, D. R., and Lin, M., Use of geographic information systems in ground-water flow modeling, J. Water Resour. Plng. and Mgmt., ASCE, 122(2), 88–96, 1996

Wilson, J.P., Inskeep, W.P., Wraith, J.M., and Snyder, R.D., GIS-based solute transport modeling applications: Scale effects of soil and climate data input. Journal of Environmental Quality 25: 445-53, 1996.

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 the GEONAMICA® software environment for integrated dynamic spatial modelling

Jelle Hurkens[a], Bernhard Hahn[a] and Hedwig van Delden[a]
[a] Research Institute for Knowledge Systems (RIKS) BV, Maastricht, The Netherlands (jelle@riks.nl)

Abstract: GEONAMICA® is a framework developed by RIKS to support the development of integrated spatial decision support systems (ISDSS). This paper firstly presents an outline of the development process of an ISDSS, signalling potential obstacles and emphasising the need for cooperation between the involved parties. Secondly, it gives an explanation of the way in which a framework helps to reduce development costs and improve the quality of the final product, combined with a study of the requirements needed for the framework to be effective. Thirdly, the paper presents an overview of the framework, relating it to existing modelling paradigms. In the conclusion, we review the adherence of the framework to the requirements we set out and give a look into the future.

***Keywords**:* modelling framework; model integration; modular model component; integrated assessment; DSS development

1. INTRODUCTION

Spatial planning processes are changing by the possibility to make better informed decisions on the basis of available spatial data and insight in spatial dynamics. This leads to an increase in the demand for Spatial Decision Support Systems (SDSS) [Densham, 1991]. However, the unfamiliarity of decision makers with SDSSs makes it hard to communicate the possibilities of such a system. Previous projects with similar objectives can help exemplify the benefits and elicit the kinds of questions or problems it can help answer, but the SDSS developer must make sure he can deliver what he promised with the often limited budgets that are available for such projects.

In this light, an application framework to support the development of a SDSS can prove a valuable advantage. A *framework* is a reusable, semi-complete (software) application that can be specialised to produce custom applications [Fayad, et al. 1999]. The primary benefits stem from two types of reuse: design reuse and implementation reuse. By delivering a useful set of patterns as a documented design, as well as a partial solution in the form of a skeleton application, a framework may save lots of costs for rediscovery and reinvention [Hahn & Engelen, 2000].

Besides reusability, *modularity*, *extensibility* and *inversion of control* [Fayad, et al. 1999] are benefits of an application framework that can reduce development costs. Moreover, in the domain of integrated SDSSs, these properties help to support a more advanced development process, in which a relatively simple prototype system is iteratively improved and expanded based on the users needs. Modularity eases the implementation of adaptations by localising their impact on the system. Extensibility allows a system to become more complex and specialised on the users needs without the need for a complete redesign. Inversion of control keeps the interaction procedures between components stable from one iteration in the development process to the next.

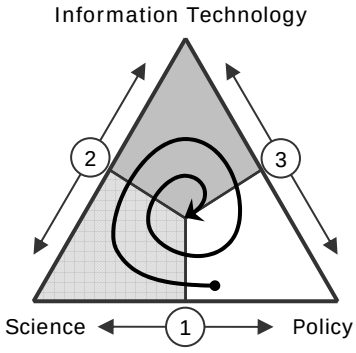
In the next part, we will elaborate on this iterative development process and the roles of the involved parties. From the insights in this process, we will derive requirements that a framework must meet in order for it to be a suitable base to build on. Section 2 gives an

explanation of the **GEONAMICA®** framework, which has been developed for this purpose, and relates it to the theoretical foundation it builds on. The last section shows how it meets the requirements we have set out and presents a look into the future.

1.1 Developing an Integrated Spatial Decision Support System (ISDSS)

The development of an ISDSS can best be described as an iterative process of communication and social learning amongst three involved parties – as depicted below. First, there are *policy makers*, the end-users of the system. They provide the policy context and define the problems, functions and usage of an ISDSS. Second, there are *scientists* responsible for the main model processes and choices of scale, resolution and level of detail. Third, there are *IT-specialists* who design the system architecture and carry out the software implementation of the models and user interface.

Note that for complex ISDSS projects, a fourth role is vital for success. The *DSS architect* has the main responsibility for integration, communication and management and assures the quality of the integrated model underlying the system. As a generalist, the architect bridges methodological and knowledge gaps between policy makers, scientists and IT-specialists – both between and within groups. In order to fulfil this role, the architect needs a solid and intuitive understanding of the application domain and the purpose of the system, as well as very good communication skills.



The interaction between the three groups involved is as important for the quality of the final product as the tasks carried out by each group individually. Policy makers and scientists can select policy-relevant research and models capable of answering the problems set out by the policy makers, translate policy options and external factors into model input and translate model output into policy-relevant indicators (1). Scientists and IT-specialists can work together to implement new models, link existing models and ensure consistency throughout the system (2). IT-specialists need to work with policy makers to set up a user interface that represents the relevant input and output in a comprehensible manner without overwhelming the user with the wealth of available information and possibilities (3).

The interaction helps each group to gain a better understanding of the needs the others have and the possibilities they offer. For this understanding to take full effect, an iterative approach is best suited. After the initial goals and requirements of the system have been established, IT-specialists can set up a prototype system to help asses its usability and show policy makers the possibilities. Their feedback can be used to make the model more robust, add missing input or output to the system and improve the operability of the user interface. The prototype can be used by scientists to train policy makers – and their technicians – in the use of the system. This includes translating real-world problems or questions into interventions in the system and translating an analysis of model output into concrete, valuable recommendations or conclusions.

Besides giving scientists a better understanding of the policy making process, such consultancy can help increase the practical value of the system – that is, make sure that it will be used and that it will be used appropriately. Due to its integrated approach, the utilisation of the ISDSS could even have an impact on the work-practice of the involved organisation. After some time, an extension of the system or refinement of some part may prove desirable and the development process can enter another iteration.

1.2 Perspectives on the system

Not only do the three parties involved in the development of a spatial DSS have different roles, they also have a different view on the system. These different perspectives can cause problems in communication or false expectations. What seems trivial to a policy maker

may be a complicated task for a modeller or IT-specialist. Conversely, small changes to the model or software can make a huge difference in the eye of a user of the system.

To policy makers the system acts as one whole; particular input will produce particular output. Though their understanding of the processes involved may be richer or poorer, their focus is aimed at the use of the system, not at its internal structure. Scientists and IT-specialists need a higher level of system specification [Zeigler, et al. 2000]; they need to know the internal structure. The decomposition of a system into coupled components helps scientists to understand it by dividing research or knowledge into comprehensible parts, often focused per discipline. For IT-specialists, the benefits are even greater, as they have the ability to design, implement and test each component individually. As a result, their decomposition may be more extensive than that of scientists. Finding an appropriate decomposition can be a difficult, time-consuming task. This can be lightened by the use of a framework, however.

1.3 Requirements for a framework / ISDSS

The previous sections give an outline of the process of developing an ISDSS and the common communication problems occurring between the parties involved in such development. From these insights we can derive four requirements for an application framework to successfully support this process, meaning its use will reduce development costs and result in a more stable and better suited product.

1. Modularity
2. Scalability
3. Powerful system in terms of size, speed and model complexity
4. Interactive system

The first two requirements stem from the iterative development process itself and relate more directly to the framework, while the last two are imparted common requirements of the ISDSS and relate also to the skeleton application the framework offers.

To support iterative development, a framework must reduce the costs of subsequent iterations by allowing easy adaptation or improvement of the work in the previous phase and by supporting the extension of the system with new models or functionality. This firstly requires *modularity* of the framework, meaning the entire system can be decomposed in a prescribed way into components that interact through well-defined, stable interfaces. When local changes must be made, modularity helps to keep the impact of those changes localised, thereby keeping development costs limited and allowing parallel work on distinct components. Secondly, the framework and system should be *scalable* in the sense that extensions can be made without the need to revise previous work and without the system growing excessively complex or resource consuming. This prevents the ISDSS developer from having to start from scratch when the user would like to see an extension of the ISDSS's model or functionality.

After the system has been expanded in several iterations, it can grow to a substantial piece of software, particularly compared to the prototype version. To still be able to use this application in a similar way as the first one – for example, running it on a desktop computer with computation times in the order of minutes – the final system has to be *powerful* in terms of size (of data and models) and speed (of data access and model computation). The framework that was used to build the first, lightweight version must support the final, heavyweight version as well. Besides the increased size of data or models, subsequent iterations of the development process can also raise the need for greater feedback between models, thereby raising model complexity. The model framework, upon which the models are built, should be able to support this increased complexity without the need for a complete redesign.

One of the general requirements for any DSS is user-friendliness. This requires from the framework that the DSS developer is able to design the user interface as he deems suitable. Therefore, the framework should not oppose design, layout or functionality on the user interface. Still, a common requirement will be for the user to be able to interact with the

system directly. This means the framework cannot afford to lock up during computation. It must stay responsive to user input and process these changes to display the correct output.

2. GEONAMICA®

GEONAMICA® is the object-oriented application framework [Fayad, et al. 1999] developed by RIKS to build decision support systems based on spatial modelling and (geo)simulation. It has been developed over the past 15 years and has been used to generate integrated spatial decision support systems, such as **WADBOS** [Engelen, et al. 2003a], **ENVIRONMENT EXPLORER** [Engelen, et al. 2003b], **MEDACTION** [Van Delden, et al. 2007], **XPLORAH** [Van Delden, 2008] and **MOLAND** [Barredo, et al. 2003]. Besides these, RIKS has used **GEONAMICA®** to develop **METRONAMICA**, a template SDSS that includes a local dynamic land use interaction model, a regional interaction model and/or a transport model – depending on the exact version. It can be used to set up a specific SDSS without the need for additional software development by filling the system with data, calibrating the model and training the users.

In ISDSSs, such as the ones mentioned above, we can distinguish three major components: a database to store information used by the system – mostly raster or vector map data, time series data and cross-sectional data –, a model base to manage the models that are used and a user interface to enable the user of the system to interact with it. Setting up each of these components and letting them work together should be facilitated by the application framework as much as possible.

GEONAMICA® offers set components for the storage of map data, time series and cross-sectional data. It provides a modelling framework based on the Discrete Event System Specification (DEVS) formalism [Zeigler, et al. 2000] and includes a model controller that manages the models, makes sure they interact properly and tells each model when to perform certain, predefined actions. To create a user interface, **GEONAMICA®** includes a skeleton structure and a rich class library of user interface components, such as map display and editing tools, list and table views and two-dimensional graph editing components.

The strength of **GEONAMICA®** lies in the fact that the modelling framework provides a generic structure for the models that allows them to be integrated more easily, while enabling complex dynamic models to be executed efficiently. The environment is set up in such a way as to enable users to run simulations interactively, by allowing them to intervene in the system and observe the results of their actions directly in a comprehensive manner or save the results to persistent storage for more elaborate analysis or presentational purposes.

2.1 Model blocks

The modelling framework in **GEONAMICA®** builds on the DEVS formalism. Specifically, the entire model is composed of *model blocks*, which are encapsulated parts of a model that can communicate with each other and with the system through a standardised interface. A model block contributes to the entire model – in its simplest representation as a collection of variables and equations or algorithms – with the variables it contains and the procedures associated with it to compute the value of these variables. Data hiding is particularly applied to the variables of a model block that can only be accessed through dedicated ports.

Model blocks communicate with each other through input and output ports. An input port allows a model block to gain read-only access to a variable of another model block, such that it can be used for computation. An output port allows read-only access to one of the variables of a model block. By linking all input ports of a model block to output ports of other model blocks, we create a coupled component. The collection of interlinked components forms the system that represents our entire model. For each particular system, the model is specified in an XML file. The listed model blocks are instantiated and coupled at run-time, thereby allowing a change of the model without the explicit need for additional programming and allowing different versions of a model to be maintained in parallel.

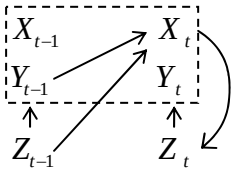
Note that the entire model can be represented by model blocks hierarchically. That is, we can define a composite model block as the collection of a set of model blocks and the links between their input and output ports. Input ports of model blocks in the composite that are not connected to output ports of model blocks in that composite are automatically rerouted to input ports of the composite model block. The set of output ports of the composite consists of all output ports of the contained model blocks.

2.2 Simulation engine

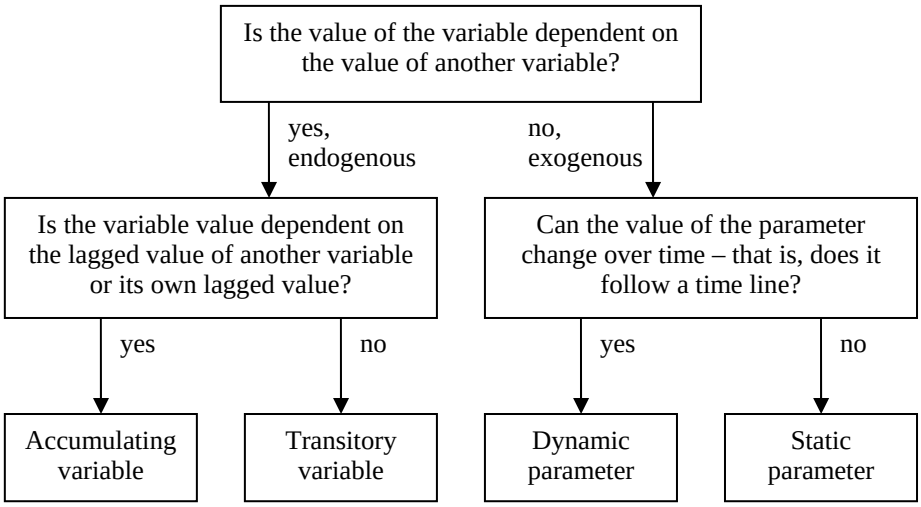
The variables in a model block represent the state of a real-world or imaginary entity at some point in time or over some time period. This means that we cannot store the value of some variable for two moments in time without explicitly making a copy. The *simulation engine* tells each model block when to compute the value of its variables for a specific point in time. The model block responds by telling the simulation engine when it should calculate again. To keep the model output comprehensible for the user, the simulation engine always keeps all model blocks up to date with the global simulation time.

This paradigm can present a problem when we calculate more than one variable in a model block. Consider, for example, a model with two model blocks. Model block A has the variables *X* and *Y* and model block B has the variable *Z*. The value of the variables is calculated as:

$$A: \begin{cases} X_t = f\left(Y_{t-1}, Z_{t-1}\right) \\ Y_t = g\left(Z_t\right) \end{cases}$$
$$B: \left\{ Z_t = h\left(X_t\right) \right.$$



On the right, these relations are depicted schematically. The arrows indicate that a variable (head) is dependent on the value of another variable (tail). We can easily see that *X* should be calculated before *Y* and *Z*, and *Z* should be calculated before *Y*. However, if *X* and *Y* are in the same model block, they should be calculated simultaneously, since calculation procedures are associated with model blocks, not with individual variables. To be able to represent this model in our paradigm, we have to make a distinction between variables that are dependent on a lagged value and variables that are only dependent on current values. We call the former *accumulating* variables and the latter *transitory* variables. So *X* is an accumulating variable and *Y* and *Z* are transitory variables. The complete determination of the type of a variable is depicted in the schema below.

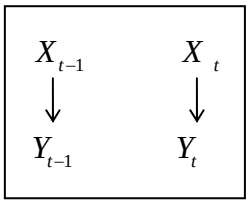


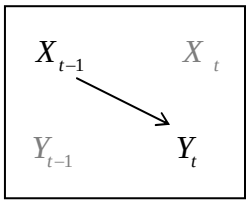
In order to be able to model the above example, we have to split the computation of a model block into two parts; one in which all accumulating variables are calculated and one in which all transitory variables are calculated. First, all accumulating variables are calculated for all model blocks. Next all transitory variables are calculated. This separation

does not resolve all precedence relations. In the example above, we see that the transitory variables of model block A should still be calculated before those of model block B.

Since accumulating variables are dependent on lagged values, their value needs to be specified for the start time of a simulation. In the initialisation of the model, copying this value to the current value replaces the calculation procedure for accumulating variables.

Once we know the dependencies between variables, the remaining precedence relations can be derived automatically, as depicted below. On the left we see a non-lagged relation and on the right we see a lagged relation. Depending on the type of the variables *X* and *Y* we have six different possibilities, as listed in the table below, where T stands for a transitory variable and A for an accumulating variable.





Relation	X_{t-1}	X_t	Y_t	**Meaning**	**Brief**
$Y_t = f(X_t)$	-	T	T	*X* must be calculated before *Y* is calculated.	$X < Y$
$Y_t = f(X_t, Y_{t-1})$	-	T	A	This is impossible, since all accumulating variables must be calculated before all transitory variables.	-
$Y_t = f(X_t)$	-	A	T	This is ok; all accumulating variables are calculated before all transitory variables.	ok
$Y_t = f(X_t, Y_{t-1})$	-	A	A	*X* must be calculated before *Y* is calculated.	$X < Y$
$Y_t = f(X_{t-1}, Y_{t-1})$	T	-	A	This is ok; all accumulating variables are calculated before all transitory variables.	ok
$Y_t = f(X_{t-1}, Y_{t-1})$	A	-	A	*Y* must be calculated before *X* is calculated.	$Y < X$

The relations between variables can be derived from the connections between input and output ports. Hence, the order in which the variables of the model should be calculated can be derived once we know the model coupling. To upscale this order to model blocks, we add equality relations (in terms of precedence) between all accumulating and between all transitory variables of the same model block. Note that, since the precedence relations between variables are relevant to either the accumulating or transitory computation phase, depending on the type of variables involved, the order in which model block calculate can differ in the two phases.

The **GEONAMICA®** model framework builds on the DEVS formalism, but does not comply with it fully. In the DEVS formalism, the distinction is made between rate and state variables. First, the change of all state variables are calculated and stored in rate variables. Next, the state variables are updated with the rate variables. This method allows model blocks to calculate completely independently at the cost of having to store the change of each variable explicitly, even when a variable could be updated directly. The **GEONAMICA®** model framework takes advantage of such redundancies to reduce the strain on resources. This comes at the cost of generality. Note that any specific model that can be implemented using the DEVS formalism, can be also implemented using the **GEONAMICA®** modelling framework. Hence, the consequences are entirely practical.

2.3 Incorporating user input

We mentioned before that the simulation engine always keeps all model blocks up to date with the current simulation time to present a comprehensible output to the user. This means that in the course of a simulation, we iteratively take time steps to advance the state of the system. At the start of a time step, we update parameter values changed by the user and recalculate the transitory variables dependent on these parameters. Next, we can advance the simulation clock and calculate the accumulating variables of each model block in the

order derived from the precedence relations between variables, as explained above. Finally, the transitory variables of each model block are updated in their respective order and the new output is presented to the user. The next simulation step is performed when the user instructs the system to do so or automatically if desired.

If we strictly follow the outset above, the user could interact with the system only after a time step is completed and before the next one has started. This is, of course, an unacceptable characteristic for an interactive system. However, allowing a user to alter parameter values in the middle of a simulation step would result in undetermined behaviour of the system. Hence, we need a mechanism that allows the user to interact with the system during the course of a simulation step, but guarantees the model that parameter values remain constant during this period. This mechanism has been incorporated in the interface ports of model blocks.

Interface ports provide access to the parameter values of a model block, which can be altered by the user between simulation steps. The changes made in the user interface are cached in the interface port to which the user interface is linked. At the beginning of a simulation step, the user interface ports are instructed to relay their cached changes to the actual parameter values. This way, we require no further synchronisation between user interface and model processes, as far as the updating of parameter values is concerned, thereby greatly reducing the overhead caused by such synchronisation issues.

3 CONCLUSIONS

The **GEONAMICA®** framework has been designed to support the development of ISDSSs and to meet the requirements set out in section 1.3. In the next section, we will elaborate on the way in which those requirements are met. The last section will discuss the limitations, possibilities for improvement and future expectations.

3.1 How does GEONAMICA® meet the requirements?

The four requirements set out in section 1.3 are met by the **GEONAMICA®** framework through different concepts incorporated therein. The main requirement of modularity is met by the decomposition of the model into model blocks and the ability to specify the complete model at run-time. Model blocks offer a clear method of decomposition that allows the model framework to implement common procedures, such as reading or writing to file. The input, output and interface ports form a stable interface for interaction between model blocks themselves and between model blocks and other components.

The fact that the model is specified at run-time turns models blocks into components that can be replaced without the need for recompilation. Combined with the possibility to specify the model hierarchically, this functionality allows one to replace an entire sub-model with another one in the blink of an eye. For example, we can replace a simple model block containing scenarios for certain exogenous variables with a sub-model that computes these variables endogenously. Except for the implementation of the sub-model itself, which only has to be done once, there is no extra programming required to incorporate it into the integrated model of the system. Additional feedback loops in the model are handled by the simulation engine, which determines the computation order of the model blocks.

By reducing the storage of redundant data – by storing each variable for only one moment in time and by introducing precedence relations for the computation of model blocks –, the **GEONAMICA®** framework allows extensive, integrated models to be incorporated in an ISDSS running on a desktop computer. By building on the DEVS formalism, the model framework allows a high level of model complexity, while keeping the strain on resources limited.

The user interface that **GEONAMICA®** offers is only an empty shell. Hence, the DSS developer is free to design the user interface as best suitable for the user of the system. One could, for example, create a different user interface for different groups of users, possibly with a different background, objective or authority. The framework does, however, offer

support for synchronisation between user interface and model processes with the caching mechanism incorporated in the interface ports. This significantly lightens the load on the user interface programmer to deal with synchronisation issues, which is one of the harder problems to tackle in the design and implementation of a user interface.

3.2 A look into the future

The **GEONAMICA®** framework has been developed with generality in mind. Its structure is based on general requirements that follow from an understanding of the development process of an ISDSS, as set out in section 1.3. The goal, however, has not been to develop a framework that can be applied in a very wide range of conceivable applications, but to develop a framework to support the actual ISDSS development that happens in practice. As a result, the framework has a limited scope of situations in which it is useful.

GEONAMICA® is a suitable candidate for an application framework to develop a DSS that incorporates a discrete time simulation model. It is designed for systems to be used interactively on a modern desktop computer running a Windows operating system. For reasons of efficiency of execution, the framework has been implemented and can – at this point – only be used in the C++ programming language. All these are technical limitations that will render **GEONAMICA®** unsuitable for a good number of projects. There are, however, a significant number of projects with the goal to develop an (IS)DSS that can benefit greatly from the use of the framework. The examples listed in section 2 demonstrate the practical value.

As the requirements posed for ISDSS development get stricter and competition becomes stronger, we need to keep developing the **GEONAMICA®** framework to improve the available components and overcome the current limitations. At RIKS, we are working towards platform independency to allow a broader range of applications of a system – for example, web-based access to models running on a server. A next step would be greater independence of programming language, which can be achieved by the ability to incorporate models that comply with compatible standards.

REFERENCES

Barredo, J., C. Lavalle, L. Demicheli, M. Kasanko and N. McCormick, *Sustainable urban and regional planning: The MOLAND activities on urban scenario modelling and forecast*, Office for Official Publications of the European Communities, Luxembourg, 2003.

Densham, P.J., Spatial Decision Support Systems. In: *Maguire, D.J., M.F. Goodchild and D.W. Rhind (eds.), Geographical Information Systems: Principles and Applications*, 403–412, 1991.

Engelen, G., I. Uljee and K. van de Ven, WadBOS: Integrating knowledge to support policy-making in the Dutch Wadden Sea, In: *Geertman, S. and J. Stillwell (eds.), Planning Support Systems in Practice*, 513–537, 2003a.

Engelen, G., R. White and T. de Nijs, Environment Explorer: Spatial Support System for the Integrated Assessment of Socio-Economic and Environmental Policies in the Netherlands, *Integrated Assessment, 4*(2), 97–105, 2003b.

Fayad, M.E., D.C. Schmidt and R.E. Johnson, *Building Application Frameworks: Object-Oriented Foundations of Framework Design*, Wiley Computer Publishing, 664pp., New York, 1999.

Hahn, B. and G. Engelen, Concepts of DSS systems in German Federal Institute of Hydrology, *Decision Support Systems (DSS) for River Basin Management*, 9–44, 2000.

Van Delden, H., P. Luja and G. Engelen, Integration of multi-scale dynamic spatial models of socio-economic and physical processes for river basin management, *Environmental Modelling and Software, 22*(2), 223–238, 2007.

Van Delden, H., E. Gutiérrez, J. van Vliet and J. Hurkens, Xplorah: A tool supporting Integrated Spatial Planning in Puerto Rico, *paper presented at this conference*, 2008.

Zeigler, B., H. Praehofer and T-G. Kim, *Theory of Modeling and Simulation*, Academic Press, 510 pp., 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

PTF: an Extensible Component for Sharing and Using Knowledge on Pedo-Transfer Functions

M. Acutis[a], M. Donatelli[b, **, †], G. Lanza Filippi[a]

[a] Department of Produzione Vegetale, University of Milano, Italy
[b] Agriculture Research Council, 40128 Bologna, Italy;
[**] currently seconded to European Commission Joint Research Centre, Institute for the Protection and Security of the Citizen, Agriculture Unit, Agri4cast Action, Ispra, Italy
[†] Corresponding author, marcello.donatelli@jrc.it

Abstract: Soil data availability for modelling purposes is often insufficient for the application of physical or semi-empirical models simulating soil hydrology. Standard soil surveys frequently do not include hydrological characteristics of the soil, such as either parameters of water retention and conductivity functions or, simpler than the former, estimates of soil water content at field capacity and wilting point. Even when at least part of such data is available, a quality control is needed to ensure not only that values fit within expected ranges, but also to check for consistency across parameters in a specific soil. The use of pedotransfer functions (PTF) allows estimating "what we need from what we have", that is, it allows estimating soil hydrological parameters from soil data often available. The literature makes available a large number of PTF, and new ones are being proposed. Such PTF range from very simple empirical functions, to complex soil physical models. Users must select a PTF to be used based on both available data and their a-priori knowledge about the soil to be simulated. Still, the choice of the PTF to be used is at times controversial, and users may want to compare the estimate made by several PTF against the same data. Also, users may want to test their own PTF, may be specific for a set of soils and thus perfectly adequate for application in a specific contest, against well known ones. An extensible and easily reusable library encapsulating a collection of PTF can be an important tool to support development and operational use of soil-related models, and to share the increasing knowledge about PTF. The objective of this paper is to illustrate the free available component PTF (PedoTransfer Functions). The component is available for both Windows .NET and JAVA platforms, and it is made available with some proof of concept applications (inclusive of source code) in C#, VB.NET and Java, which show how to extend the component and how to use it. The software component presented in this paper meets the following main requirement: i) easy to reuse; ii) with a clear ontology of the variables used in each PTF, where units, value range, and significance, are unambiguously defined; iii) extensible by third parties independently, allowing for an open system to which scientists can contribute; iv) freely available for non-commercial use.

Keywords: soil hydrologic characteristics, soil water retention, soil water conductivity, software components.

1. INTRODUCTION

The application of physical or semi-empirical approaches in all models that involve soil as one of the sub-system to be simulated (e.g. crop growth, nutrients and pollutants dynamics, CO_2 sequestration, soil organic matter dynamics) require soil data (in particular hydrological parameters) that are often unavailable. Consequently, since a long time [e.g. Nielsen et al., 1986] there has been an interest in methods to estimate soil hydrological parameters from commonly available soil data. More recently, the term "pedotransfer function" (PTF) was introduced by Bouma and van Lanen (1987), and become of common use after the work of Bouma [1989]. In this paper, the link between soil surveys

and soil hydrology is emphasized, but it is also shown that several soil variables, such as soil erodibility [Renard et al, 1997] or parameters for solute transport [Gonçalves et al, 2001], can be estimated from basic data,. Nowadays a large number of procedures to estimate soil hydrological parameters has been developed, using different methods. Such methods are based on regression as in the classic work of Gupta and Larson (1979) and in more recent EU HYPESS project [Wösten et al, 1999], or the Vereecken et al. PTF [1989], on either physically based or physical-empirical approaches [e.g. Aria e Parish, 1981, Mishra et al, 1989], or on neural networks [e.g. Shaap et al, 2001, Minasny and McBratney, 2002].

It can be very difficult for a user to select what is the more appropriate method for a specific application in a specific environment [Acutis and Donatelli, 2003], considering that the different procedures can produce very different results [Gijsman et al., 2003], so there is a need for a tool that offers the possibility to estimate the unknown variables with several methods on the basis of available data. Assuming that some laboratory or field measurements of the desired hydrological parameters are available, multiple estimates can be evaluated against such data, allowing the selection of the best performing method. Such evaluation can be done using also composite metrics specifically developed for soil PTF [Donatelli et al., 2004].

Some software tools have been developed implementing PTF. Several of them run in a web browser, but offer only a single or few methods of estimation; examples of such software are Pedon-E [Ungaro et al., 2001] and SWLIMITS [Ritchie et al., 1999, Suleiman and Ritchie, 2001]. Some already existing stand alone software, i.e. SoilPAR2 [Acutis and Donatelli, 2003] present a collection of pedo-transfer functions (PTF), with some possibility to store and visualize soil data, compare estimation against measured data, and evaluate the performance of different PTF. Main limit of SoilPar2 and similar software is that they are close applications, hence it is not possible to add other PTF not included in the application. Also, there is no quality control on data used.

To overcome the limitations above, a software component with the following main requirements is presented in this paper: i) easy to reuse, in stand alone or web application, in Windows .NET and Java platforms; ii) with a clear ontology of the variables used in each PTF, where units, allowed value range, and significance are unambiguously defined; iii) extensible by third parties independently, and in a transparent way, without recompiling the code, allowing for an open system to which scientists can contribute; iv) freely available for non-commercial use. The PTF component was developed to match the requirements above and others.

2. IMPLEMENTED PTF METHODS

The component implements a range of PTF methods:

1. to estimate water content at some specific vales (defined Point PTF in Acutis and Donatelli, 2003) and notably for field capacity and wilting point;
2. to estimate parameters of different types of retention functions and for conductivity functions;
3. to estimate saturated conductivity and bulk density.

The modelling background is essentially based on the approaches published in Pachepsky and Rawls [2004]. The list the PTF currently made available in the component, and their inputs and outputs, is reported in table 1. The equations of the implemented PTF are not reported here, but reference to the source is provided with the documentation. When possible, each implementation was tested against other existing software (unit tests were made and are made available in the documentation of the component).

Table 1. Pedotransfer function currently implemented in PTF, with their input and output .

	Pedo-transfer function										
Variable	**H**	**V**	**B**	**BSS**	**MJ**	**H**	**HM**	**J**	**MQJ**	**S**	**C**
IsTopSoil	in	-	-	in	-	-	-	-	-	in	-
Clay	in	in	in	in	in	in	in	in	-	in	in
Silt	in	in	in	in	in	in	in	in	-	in	in
Sand	-	in	in	in	in	-	-	-	-		in
Organic carbon	in	in	in	in	in	-	-	-	in		in
BulkDensity	in	in	in	in	in	in	-	in	out	out	in
SWC field capacity	out	out	out	out	out	out	out	-	-		out
SWC wilting point	out	out	out	out	out	out	out	-	-		out
SWC at various pressures	-	-	out	out	-	out	out	-	-		-
Van Genuchten N	out	out	-	-	-	-	-	-	-		-
Van Genuchten M	out	out	-	-	-	-	-	-	-		-
Van Genuchten theta saturation	out	out	-	-	-	-	-	-	-	out	-
Van Genuchten theta residual	out	out	-	-	-	-	-	-	-		-
Campbell A	-	-	-	-	out	-	-	-	-		out
Campbell B	-	-	-	-	out	-	-	-	-		out
Mualem *l*	out	-	-	-	-	-	-	-	-	-	-
K saturation	out	out	-	-	-	-	-	out	-	out	out

H = HYPRESS (Wösten et al, 1999) ; **V** = Veerecken et al (1989); **B** = Brakensiek et al. (taken from Hutson and Wagenet, 1992) ; **BSS** = British Soil Survey (taken from Hutson and Wagenet, 1992) ; **MJ** = Mayr and Jarvis (1999); **H** = Huston (Hutson and Cass, 1987); **HM** = Huston modified (Hutson and Cass, 1987); **J** = Jabro (Jabro, 1992) ; **MQJ** = Manrique and Jones (1991) ; **S**= Saxton et al. (1986); **JT**=Jaynes and Tyler (Jaynes and Tyler, 1984) ; **C** = Campbell (1985)

3. UTILITY TOOLS

Two tools are also included in the PTF component, one to convert across soil particle size distributions classification systems, and the second for fitting retention functions to pressure-soil water content data couples.

Different standards for the description of soil particle size distribution (PSD) are adopted in different countries, and some of these are also internationally adopted. The large part of PTF is based on the FAO [1999] standards, but also the ISSS is widespread, and in the UK, another standard is adopted. So, there is a need to convert across PSD classification schemes, also to use a specific PFT developed for another classification schema to avoid errors using an apparently similar one, as pointed out by Nemes and Rawls [2006]. The tool for particle size classification conversion is based on log-normal interpolation. Even if is the simpler and a low performing method [Nemes et al., 1999], it was chosen because it can be used also when data are based on only 3 textural classes, which makes it usable on most database; other methods require four textural classes in our knowledge.

The tool for fitting the retention functions to pressure-soil water content data couples (not available in the Java version) uses the multi-start simplex approach [Duan et al., 1992] to fit these highly non-linear functions. It does not require a set of coherent initial values as the traditionally used Marquardt's method. Because the direct method adopted (derivatives are not computed), standard errors of the functions coefficients and correlation among themselves are computed using a boot-strap-based approach [Shao and Tu, 1995].

4. SOFTWARE COMPONENT DESIGN

The software component PTF (PedoTransfer Functions) contains functions for the computation of different pedo-transfer methods. Each method is implemented as a class called "strategy". Data are provided as inputs using data-types called "domain classes". Domain classes implement a description of the domain being modelled via variables and a set of attributes associated to each variable (minimum, maximum, and default values; units; description; URL). Domain classes can be extended, and new strategies can be built implementing the interface exposed, thus allowing the extension of the component independently by third parties. Transparency and ease of maintenance are granted, also providing functionalities such as the test of input data versus their definition prior to using the PTF; same tests can be run on outputs. Such tests are run via a component called Preconditions, available both in .NET and Java, which allows sending the output to screen, a text file, and XML, or to a trace listener defined by the PTF clients (http://www.apesimulator.org/help/utilities/preconditions). The software architecture of this component further develops the one used in previously developed components and is fully described by Donatelli and Rizzoli [2008]. The component is freely available to scientists and institutions developing component-oriented models and applications in the agro-ecological field. There are two versions of the component, one is written in C# for .NET (extensions can be written in any .NET language), and the second is written in Java. The component is deployable and reusable in any application developed using the Microsoft .NET framework, or Java. The PTF software development kit includes sample projects (in VB.NET, C# and Java) which show how to use and extend the component. A proof of concept application is also made available in the Java software development kit (SDK). The .NET SDK includes also two proofs of concept projects to show how to use the component in a web application and to build web services. Code documentation is also provided and the online help file is available at: http://www.apesimulator.it/help/utilities/ptf.

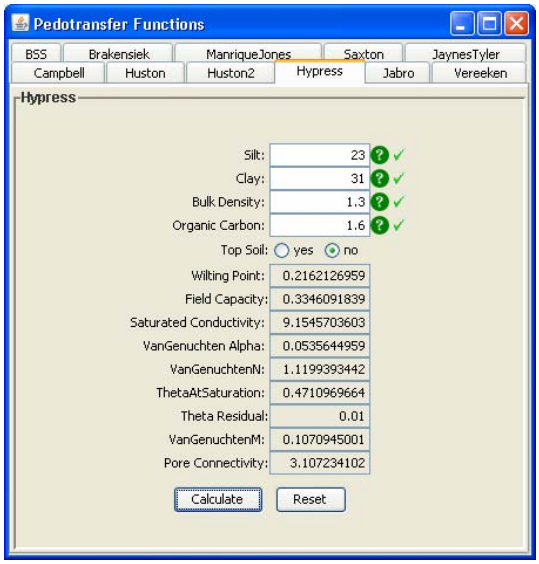


Figure 1 Screen shoot of a proof of concept application implemented in Java.

5. DISCUSSION

PTF are increasingly used to surrogate difficult and time consuming measures of soil hydrological parameters. However, because of the characteristics, behaviour and performance of different PTF are difficult to assess, there is no clear or unique path for their use and evaluation. An example of this difficulty is showed in figure 2, where some PFT offer good estimation in some case and not in other, and not in relation to soil textural classes.

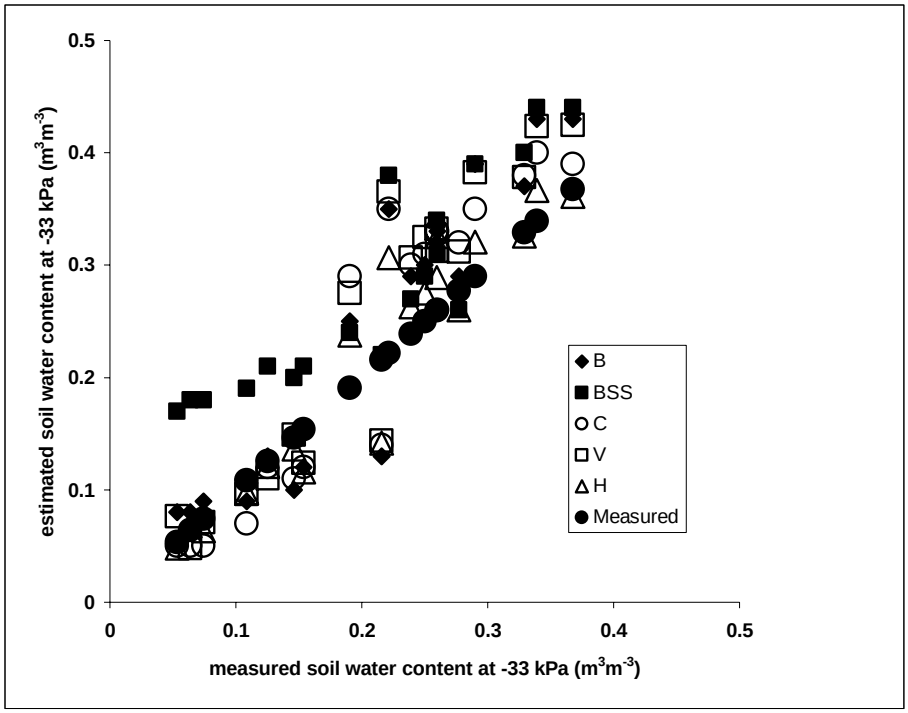


Figure 2 Examples of estimation of soil content at -33 kPa using different PTF function in several soil agains measured value. PTF: B=Brakensiek; BSS=British Soil Survey; C=Campbell; V=Veerecken; H=HYPRESS

Moreover, given that currently there is not a PTF approach that has shown a consistent superiority when compared to others, several researcher are working in developing new PTF methods. From the previous consideration, a software tool like the PTF component, which is developed targeting reusability and extendibility, can give an important contribution to both operational and scientific applications. Several agricultural and environmental models require soil data, and frequently these data are unavailable from direct measurement, so, the application of some PTF is needed, frequently re-implementing them; the PTF component offers an opportunity to provide such model packages of rich features with minimum effort. Also, the component PTF, that enables the use of several approaches, allows separating the effect of data estimation from the real ability of the model, and allows an easy functional evaluation of the PTF itself [Wösten et al, 2004].

Linking the PTF component with the IRENE (Integrated Resources for Evaluating Numerical Estimates), which shares the same architecture [Bellocchi et al., 2008], allows building composite indices for an articulated testing of PTF methods [Donatelli et al., 2004].

The first feedback received by the users of the prototype is strongly positive. The PTF SDK can be downloaded from http://www.apesimulator.org/public/downloads/ptf/ .

6. CONCLUSIONS

The component PTF makes available a number of Pedotransfer Functions in a discrete software unit. Its extensibility, also independently by third parties, allows easy maintenance and stimulates further development. The goal of PTF is to be a way to share knowledge among scientist and model users, via effective cost-benefit re-use and allowing an easier cross-testing of PTF.

ACKNOWLEDGEMENTS

This publication has been partially funded by the SEAMLESS integrated project, EU 6th Framework Programme for Research, Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract no. 010036-2).

REFERENCES

Acutis, M and M. Donatelli, SOILPAR 2.00: software to estimate soil hydrological parameters and functions. *Eur. J. Agron.* 18, 373–377, 2003.

Arya, L.M. and J.F Paris., A physicoempirical model to predict soil moisture characteristics from particle-size distribution and bulk density data. *Soil Sci. Soc. Am. J.* 45,,1023–1030, 1981.

Bouma, J., Using Soil Survey data for qualitative land evaluation. *Adv Soil Sci.* 9, 177-213, 1989.

Bouma, J., and H.A.J. van Lanen, Transfer functions and threshold values: from soiol charachteristics to land value. In: *Proc of the int. workshop on Quantified Land Evaluation Procedures, 27/04-2/05/1986, Washington DC, USA,* 106-110, 2007.

Campbell G.S, In: Soil Physics with BASIC: Transport Models for Soil–Plant System, *Elsevier, Amsterdam,* p. 150, 1985.

Donatelli M. and A. Rizzoli, A design for framework-independent model components of biophysical systems. *These Proceedings*

Donatelli, M, M. Acutis, A.Nemes and H. Wösten, J.H.M, Integrated indices for Pedotransfer function evaluation. In: *Pachepsky, Y., and Rawls W.J., Development of pedotransfer functions in soil hydrology. Elsevier, Amsterdam, The Netherlands.* 363-390, 2004.

Duan, Q., S. Sorooshian, and V.K. Gupta, Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models. *Water Resources Research* 28(4), 1015-1031, 1992.

FAO, Guidelines for soil description. *Third edition, FAO/ISICR, Rome, Italy,* 1990

Gijsman, A.J., S.S. Jagtap, J.W. Jones, Wading through a swamp of complete confusion: how to choose a method for estimating soil water retention parameters for crop models. *Eur J. Agron,* 18, 75-105, 2003.

Gonçalves, M.C., F.J. Leij and M.G. Schaap, Pedotransfer functions for solute transport parameters of Portuguese soils. *Eur. J. Soil Sci.*, 52, 263-574, 2001.

Gupta, S.C.and Larson. W.E. Estimating soil water retention characteristics from particle size distribution, organic matter content, and bulk density. *Water Resource Res*. 15, 1633-1635, 1979.

Hutson, J.L. and A. Cass, A retentivity function for use in soil water simulation models. *J. Soil Sci.* 38, 105–113, 1987.

Hutson, J.L., Wagenet, R.J., 1992. LEACHM, Leaching Estimation And Chemistry Model. *Department of Soil, Crop and Atmospheric Sciences Research series no.* 92.3. *New York, Cornell University*, p. 121.

Jabro, J.D., Estimation of saturated hydraulic conductivity of soils from particle size distribution and bulk density data. *Trans. ASAE* 35, 2, 557–560, 1992.

Jaynes, D.B. and E.J. Tyler, Using soil physical properties to estimate hydraulic conductivity. *Soil Sci.* 138, 298–305 1984.

Manrique, L.A. and C.A. Jones, Bulk density of soil in relation to physical and chemical properties. *Soil Sci. Soc. Am. J.* 55, 476-481, 1991.

Mayr, T. and N.J. Jarvis, Pedotransfer function to estimate soil water retention parameters for a modified Brooks–Corey type model. *Geoderma* 91, 1–9, 1999.

Minasny, B., and A. B. McBratney, The neuro-m method for fitting neural network parametric pedotransfer functions. *Soil Sci. Soc. Am. J.* 66, 352-361, 2002.

Mishra S., J.C. Parker, N. Singhal, Estimation of soil hydraulic properties and their uncertainty from particle-size distribution data. *J. Hydrol.* 108, 1-18, 1989

Nemes A., J.H.M. Wösten, A. Lilly and J.H Oude Voshaar. Evaluation of different procedures to interpolate particle-size distributions to achieve compatibility within soil databases. *Geoderma* 90, 187–202, 1999.

Nemes, A.and W.J. Rawls, Evaluation of different representations of the particle-size distribution to predict soil water retention. *Geoderma* 132, 1-2, 47-58, 2006.

Nielsen, D.R., M.Th. van Genuchten and J.W. Biggar, Water flow and solute transport processes in the unsaturated zone. *Water Resour. Res.* 22, 9, 89–108, 1986.

Pachepsky, Y., and Rawls W.J., Development of pedotransfer functions in soil hydrology. *Elsevier, Amsterdam, The Netherlands.* 2004.

Renard K.G., G.R. Foster, G.A. Weesies, D.K. McCool and D.C. Yoder (Coordinators), Predicting Soil Erosion by Water: A guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE). *Agricultural Handbook* No.703, 1997. *USDA-ARS, Washington D.C.*

Saxton, K.E., W.J. Rawls, J.S. Romberger and R.I. Papendick, Estimating generalized soil–water characteristics from texture. *Soil Sci. Soc. Am. J.* 50, 1031–1036, 1986.

Schaap, M.G., Leij, F.J. and Van Genuchten, M.Th., 2001. Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* 251, 163–176, 2001.

Shao, Jun, and Dongsheng Tu, The Jackknife and Bootstrap. *Chapter 8. In: Series in Statistics Springler,* 516 pp, 1996.

Suleiman, A.A. and J.T. Ritchie, Estimating saturated hydraulic conductivity from soil porosity. *Trans. ASAE* 44, 2, 235–239, 2001.

Vereecken H., J. Maes, J. Feyen and P. Darius, Estimating the soil moisture retention characteristics from texture, bulk density and carbon content. *Soil Sci.* 148, 389-403, 1989.

Wösten, J.H.M., Lilly, A., Nemes, A. and Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma 90, 169-185*, 1999.

Wösten, J.H.M., A Nemes and M. Acutis, 2004 Functional evaluation of Pedotransfer function. In: *Pachepsky, Y., and Rawls W.J., Development of pedotransfer functions in soil hydrology. Elsevier, Amsterdam, The Netherlands.* 390- 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

Community-based software tools to support participatory modelling: a vision

Alexey Voinov[a], David Arctur[b], Ilya Zaslavskiy[c], and Saleem Ali[d]

a – Chesapeake Research Consortium Community Modelling Program, Johns Hopkins University, USA, alexey.voinov@uvm.edu
b – OGC Interoperability Institute, Austin, Texas, USA
c – San Diego Supercomputer Center, CA, USA
d – Rubenstein School of Natural Resources, University of Vermont, USA

Abstract: Environmental management depends on analysis of complex dynamics and spatial relationships of ecological and socio-economic systems. Modelling, when used to conduct such analyses, is recognized as an effective decision support tool in environmental management. Modelling conducted in a participatory fashion, involving stakeholders in various stages of model building and data processing has evolved as an efficient method for conflict resolution and decision-making. However, successful participatory modelling efforts require specific software and computer tools that are not available or accessible for stakeholders. There is a clear need for specialized modelling and data processing infrastructure that would allow comprehensive environmental simulations, based on limited computer programming skills, computer power, and data availability. We are developing a software framework of model and data modules to enable various stakeholders to tap into the recent and ongoing advances in environmental modelling, and high-quality data available on the Internet. The proposed framework would allow managers and planners to run simulations of policy scenarios and utilize state-of-the-art algorithms to develop and evaluate policy alternatives.

The web-based modelling framework is based on the following components:

- A web-based domain-specific interface which facilitates the development, configuration, and execution of models applicable to region-specific watershed issues;
- A data-finder and transformer unique to the landscape modelling framework that lever-ages relevant Open GIS catalogue, RDF, and GRID resource discovery standards;
- A module composer that uses a module pool and guided composition of modules based on expert rules, which are either automatically acquired or input from human users, to guide the simulation-modelling process; and
- A semi-automatic model calibrator and verifier to deliver high quality simulation models.

The framework's core components, i.e. model composition, data finder, and geospatial data processing, serve the needs of a wide range of applications. The framework should be designed to be configurable for multiple domain specific user groups, enabling applications as diverse as agricultural forecast, real estate market analysis, transportation planning, etc.

Keywords: Modularity, interoperability, conflict resolution, environmental management

1. INTRODUCTION

Environmental management of natural resources has become an important responsibility of regional environmental agencies and planning departments around the country. Dealing with issues of natural resource management is quite challenging for five principal reasons:
(1) Public and private interests are in dispute, resulting in mutually exclusive alternatives;
(2) There is political pressure to make rapid and significant changes in public policy;
(3) Private and public stakes are high with substantial, often irreversible, costs and risks;
(4) The technical ecological and sociological facts are highly uncertain;
(5) Policy decisions have effects outside the scope of the problem.

Due to its complex, multi-disciplinary nature, and long-term effects on the integrity of both natural resources and socio-economic systems, environmental management has necessarily become a participatory exercise, requiring feedback from many stakeholders. This requires that all participants understand the complex interactions and processes in natural and social systems over space and time. Dispute resolution professionals have traditionally relied on the exchange of information at forums and tended to reduce the complexity of the problem to make it cognitively more suitable for resolution. The use of computational techniques was relegated to optimization algorithms that presented multiple criteria for maximizing benefits and minimizing costs in a matrix from which various options were ranked and presented to stakeholders. However, stakeholders themselves were not part of the epistemic process and hence were often quite sceptical of the decision matrices. The use of cyber-infrastructure in eliciting and processing information from dispute stakeholders can thus be both empowering as well as providing clarity through synthesis of complex data sets [Voinov, Costanza, 1999; van den Belt, 2005]. Environmental parameters that can be presented in spatial terms have also been shown to enhance the potential for dispute resolution [Kyem, 2005].

After several decades of intense development, environmental simulation modelling has established itself as a powerful paradigm for understanding and forecasting environmental processes. Computer simulation has become an invaluable method for environmental decision-support, and has spawned a wide variety of tools developed by EPA, US Army Corps of Engineers, USGS, universities, and private industry. Despite the advancement of such tools, resource managers still must overcome four major obstacles in order to routinely benefit from goal-driven environmental modelling:

- Each region requires unique spatial and temporal data. These data are often difficult to find and require much effort to acquire and pre-process, due to a wide variety of formats, different sampling procedures, instrumentation, semantic conventions, and varying quality.
- Extensive computer programming and environmental science expertise are required to select appropriate models, link them together, import necessary data, and execute scenarios. Regional agencies rarely possess the necessary computer programming skills.
- The complex behaviour of ecological and social systems is often sensitive to small perturbations, and requires large amounts of computing power to run several models in concert to simulate environmental dynamics in adequate detail.
- Different environmental stakeholders often employ conceptual schemas and simulation models which describe a narrow set of processes, are hardly compatible, while procedures for assimilating different modelling results in informed model-based consensus-building are either confusing or absent.

Instead of contracting difficult modelling tasks to external experts, we advocate direct involvement of stakeholders in the modelling exercise. To facilitate stakeholder participation there is a clear need for an advanced interactive simulation modelling system tuned to the needs

of environmental dispute resolution and providing resource managers with easy access to data and modelling tools, expert model composition rules, previous model runs, and visualizations of model outputs. This system can facilitate an approach that we call participatory environmental management. The target users of the system form a large user community, found in local and state governments, regional environmental agencies and planning offices. These users provide critical information to citizens and decision makers, and are positioned between the citizen and modelling 'experts'. They typically do not have experience in computer programming or model development, but do have experience and training in the environmental management domain.

So far there are no integrated modelling systems specifically developed to support the stakeholder participatory modelling process. Two types of modelling tools have been used in this context. General-purpose modelling systems such as Stella [2007], Microsoft Excel, Extend [2007], or Simile [2007] are simple enough for the stakeholders themselves to handle. More complex modelling tools, such as the LMF [2007] or Cormas [2007], have been used in a participatory context, but require teams of researchers to operate and support throughout the process. Neither of these software tools has been developed for multi-stakeholder decision-making and dispute resolution in a participatory setting, and lack essential functionality to integrate existing more complex models or access online environmental data repositories. In addition, most existing modelling environments are proprietary and are not designed for extensive customization needed to support participatory modelling.

There is a clear need to integrate simulation modelling and data processing tools to empower citizens and decision makers with the ability to evaluate management tradeoffs in an accurate, cost-effective, transparent, and publicly accountable manner. Any dispute can be treated as a clash of different models. Stakeholders contributing to a dispute resolution exercise come to the table with their different models, qualitative and quantitative, of the system at stake. The dispute evolves because of the inconsistencies and controversies between the different models. We hypothesize that by harmonizing the models for use in a common framework, much of the conflict can be resolved. In a way participatory modelling is a mechanism of joint fact finding and understanding when data and knowledge are shared among stakeholders in attempts to build a common model. When the participants mutually educate each other about the models they use, and arrive at a shared model of a system there remains less reason for conflict and dispute.

2. THE SYSTEM

We envision the architecture of the system as follows. The central components of the system are the "simulation modules" and the "composition metadata and rules" governing how the modules can be composed into comprehensive environmental models and workflows. The composition metadata consists of rules for composition, and quality indices for individual modules as well as rules to derive quality indices when modules are composed and used. Module metadata and the model composition rules are organized as a database that users can browse, query and annotate using a web browser. In addition, the system contains a registry of previous simulation model runs representing "best practices" of model applications for different scenarios. The main task of the "module composition wizard" is to guide users through the assembly of simulation models from the available modules. In a typical session, users will pick modules from the module registry, which best address the issue in dispute. If there is a history of previous use of these modules they will come already linked to necessary other supporting data and modules. When this information is missing, users are guided to compose modules into a modelling chain and, using model composition rules, define how data are exchanged between the modules and between online data sources and the modules. This compilation, saved as an XML document, is then interpreted by the model execution pre-processor responsible for retrieving data from online sources and converting them into model inputs. The next step is orchestrating a simulation run over the retrieved data, and calibrating model parameters. The

"Module building assistant" augments the wizard and contains routines that find and suggest new modules, data, and model composition rules that are not yet part of the system. This will guide future modelling exercises and provide tools to enhance the framework. In many cases this will require additional effort to "wrap" models and data making them compliant with the rest of the system. In this way, each module will be accompanied by a detailed and expanding use history to help stakeholders learn from the experience of others and share their own experience. A hierarchy of user access will provide several tiers of functionality for different groups of stakeholders, depending on their role in the project.

The "Optimized simulation execution engine" is responsible to run the models with the available data ("Locally archived data"). The model runs are optimized with this engine to facilitate faster response in support of interactive modelling sessions. Finally, the "Module calibrator and validator" provides necessary quality information about the simulation modules available in the system. Once the model is tuned to the available data, the calibrated parameter values are recorded and annotated, and model output is stored, analyzed and annotated using the online mapping, charting and annotation components of the modelling system.

One of the main reasons that existing modelling tools are hard to use is that data preparation is a hard and cumbersome task. The objective of the Locally Archived Data is to supply data to the simulation modules in the module pool. Data come from various archives in various formats. The Open GIS Consortium (OGC) standards on data interoperability are being implemented in many public agencies, driven by the Federal Geospatial One-stop initiative [OGC, 2007]. Relevant OGC standards for data and geospatial services should be adopted and taxonomic and ontology-driven discovery mechanisms should be used based on model requirements [OGC, 2007a].

2.1 Web-based User Interface (the Guru)

This component will serve as the entry point for users to interact with the system. System functions will be delivered to the user through this user interface driver, which will also provide essential tutoring and guidance. WEB technology will be used to create the user interface. The tasks that a user can perform from this interface are as follow:

- Search for useful modules and models;
- Compose simulation and data modules to form task-specific models;
- Test models;
- Run the model to obtain output data (delivered in different formats, plots or data).

The concept of workspace will be used to allow users to localize their own modules and models. A particular use of a module in this space is to construct a data module that actually takes data from the users' private holding. This will be especially useful for testing watershed questions that require more detailed or specific data inputs. The workspace will also allow users to upload their data to be used for their own simulation tasks. However in most cases users will be encouraged to make their data simply available over the Internet. This will be sufficient for the system to access them seamlessly.

The Guru will also serve the following maintenance tasks: (1) Compose simulation and data modules and insert them back to the Module Pool; and (2) Add relevant metadata.

Finally, results delivery is part of the web-based user interface. An important task is data visualization. These visualization modules are also in the module pool (see below) for the user to pick and choose.

2.2 Module Composition Wizard (the Composer)

The wizard first takes the user through a question-and-answer session, which automatically generates a draft model from existing modules. If the resulting model is not uniquely determined, the system will present various options from which the user must select. The user is guided through this with a series of explanations for the presented options, based on the module quality indices and composition rules.

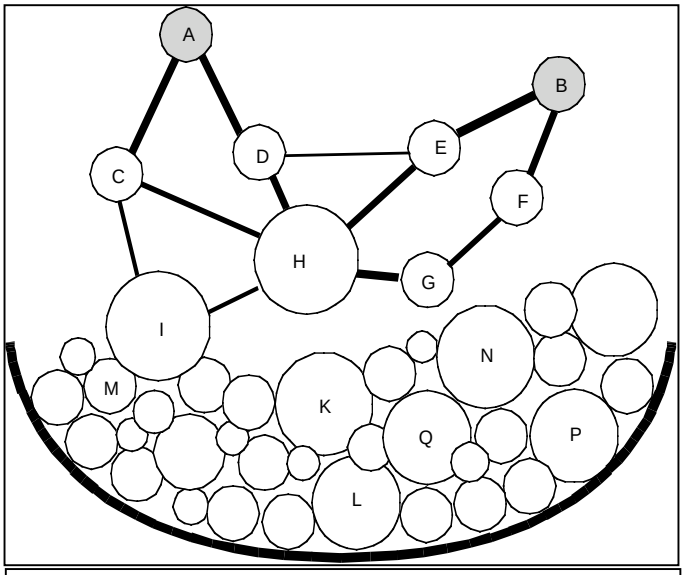


Figure 1: From the pool of available modules the user picks a few focal modules that respond to the issues at stake (say, A and B). These modules in turn are linked to other modules in the pool that provide input to the focal modules. The complexity of a module is indicated by the size of the circle, and the thickness of the link represents its reliability. There may be various minimal configurations that provide functionality to the model thus formulated (say, ADEB is the simplest model, whereas ADHEB is probably the most reliable one in terms of quality of links; ACHGFB may turn out to be the optimal one in terms of module quality). The users can choose the configuration that best suits their needs

The building block in our modelling approach is a module, which can be thought of as a function with inputs and outputs, and usually has a mathematical description. We allow different kinds of modules to co-exist in the module pool. For example, there might be modules described in XML-based Modular Modelling Language [Maxwell, Costanza, 1997], C++ code pieces, data sets, and Stella models. These simulation modules can be quite complex describing temporal or spatial dynamics of several variables, whereas others may be simple unit translators between output from one module and input to another. Some modules are "conduits" that "pipe" data sources from the Web, or in the local store, to other modules for a simulation task. Data modules are simple datasets that have only an output function. In short, the module pool contains simulation and data resources that are useful for watershed modelling, and can accept new modules as they become available.

In the simplest terms, a modelling task involves linking modules, connecting output from some modules to the input of others, and ultimately generating results that are useful for visualization and understanding of watershed processes. A general problem is how to describe the inputs and outputs of the modules so that only the appropriate connections are engaged. This, in general, is a very difficult task, akin to the code "reuse" problem in software engineering. However, we are not solving a general software engineering problem, since we are focused on a specific subset, that is, spatially explicit, coupled ecological, hydrologic, and socio-economic models. When we design a module for the module pool, we can determine which other modules will be connected (through input and/or output). A module design interface will be provided to help this connection task. This ensures that any future additions to the module pool are already linked to all the other possible sources/sinks of information and such that in future modelling tasks, the modules are used appropriately according to the domain knowledge inherent in such pre-built connections.

A conceptual view of the module pool is shown in Fig. 1. When a modelling task is started, the user identifies (through the composition wizard and ontology search tools) the few "focal" modules that generate the output essential to their task. Due to the pre-wiring, when these modules are "pulled" out of the pool, they "drag" out a whole vine of other modules that are connected to the focal modules. This pre-wiring defines all of the "possible" connections, including those that are critical for the focal module and those that are supplementary. It also determines the initial conditions, the forcing functions, and the boundaries of the system. Oftentimes, alternatives exist for many of the supplementary modules. For example, similar

data may be available from different sources (represented by different data modules). If real data is limited, a simulation module might be used to generate approximate data. To make the right choice for a particular model goal the "composition wizard" will guide the modelling process through the choices between alternative modules. To be successful, the wizard needs to know the tradeoffs between different choices and be clear which modules are essential for analysis determined by the focal module.

Several indices can be helpful:

- Simulation/data quality indices, assessing the quality of the modules for various applications;
- Composition quality index, rating the criticality of particular linkages between modules;
- Model quality index, assessing the overall quality of a complete model.

The users can specify a number of criteria that are important for them, including performance, reliability, and sensitivity. Based on these, the final configuration of modules is determined, and the model is generated. A specific challenge in creating the Composer is the heterogeneity of data and modules. A study of metadata, matching of modules, and conversion module generation will be needed.

2.3 Module Calibrator and Validator (the Tester)

Quality indices of the modules are particularly important for the system. Calibration and validation of modules requires both guided and automated processes. An important research question is how to measure the quality of a module or model based on its previous performance.

We can distinguish between model testing or evaluation (assessing the degree and significance of fit of the model with data or with other models) and calibration (the adjustment of model parameters to improve the fit). Unlike statistical models, there are no universally accepted methods for testing or calibrating complex dynamic, non-linear simulation models [Berk, et al., 2000]. For the initial calibration and validation we can apply a range of visualization techniques and quantitative statistical tests informed by the specific context of the problem [Berk et al., 2001]. The human brain is a powerful pattern processor, and if model output can be presented in appropriate formats, direct visual comparisons of models with data can yield significant insights about model performance [Kuhnert et al., 2005]. In addition, complex models must generally be tested against a broad range of data of varying types, quality, and coverage. For example, for some variables we may have only scattered field measurements, while for others we may have more complete time series data or even maps, while others may have no quantitative data at all, but only qualitative assessments. As discussed above, a system is therefore necessary for ranking or grading the relative quality of the data and the relative importance of the variables to be fit.

The difficulty of calibrating complex models by manually adjusting parameters can be overwhelming, and an optimal parameter set may not even exist [Beven, 1993]. It is therefore more appropriate to recognize the uncertain nature of the process, to focus on finding "good" parameter sets, and to create a better way of defining the quality of parameter sets. However, once we have done the original model testing, we can automate some of the recalibration/validation functions to the Tester that together with the Feeder will keep track of all the newly available information and do a background recalibration of the modules.

There are web sites with high resolution and continuous monitoring data that will be used as test cases. A well watched site will be used to assess the impact of various levels of data availability on the calibration process and on the model results. A research question is how to "incrementally" calibrate modules. That is, instead of running models "from scratch" when new data become available, we try to continuously test the modules, as in continuous queries [Babu, Widom, 2001]. Once statistics of module testing are collected, we can use data mining techniques to find ordinary cases (outliers), and general trends in the model output [Han, Kamber, 2000]. All this is an important part of the ultimate goal of these system components,

that is, to constantly evaluate and update the quality indices that are assigned to the modules and links in the module pool.

2.4 Optimized simulation execution engine (the Runner)

This part of the system is responsible to execute the simulation after the user has composed the model (with the help from the Composer and the Expert). In a participatory modelling environment, fast response is of paramount importance, as an execution of a model may be essential for further modification of the model. This is similar to OLAP systems (online analytical process systems) [Sarawagi, et al., 1998]. Since OLAP systems are mainly used in an exploratory fashion, fast response is the main concern. We plan to study techniques used in OLAP systems, such as pre-computation, for our purpose.

The second direction is to use model output from previous runs to accelerate the current model execution. This is realistic due to the fact that when models are constructed in an exploratory way (in a participatory environment), the previous model and the current model may share a lot of common modules. The intermediate results of the previous model execution may be used to expedite model executions. This is similar to using pre-computation to achieve faster response in database systems [Wiekum, 2002]. As in any pre-computation scheme, the critical issue is what computing results to store and how to find the useful results. In our simulation modelling system, to achieve such pre-computation, we can use a self-adaptive method to record computing results. That is, initially, when the storage space is available, we will store all pre-computation results of all intermediate computation (the modular approach of the models in our system makes this easy to implement). As the time goes by and we need to purge storage, we will retain those that have been used most recently while purging away the stored results that are not used recently. Another problem is how to search for available pre-computed results. We can associate the module metadata with pre-computation results. If the exact modules are used (including the same data modules), then the pre-computation result can be used.

4. DISCUSSION AND CONCLUSIONS

There are three important principles that the system is based on:

(1) The use of a module pool and guided composition of modules as a basis for simulation modelling. Modules come in a variety of time and space resolutions and scales. Matching these scales and ensuring consistency in the overall model is not a trivial process and may be difficult to fully automate. As one possible solution we can pre-link the existing modules making sure that they are compatible and can function in concert. Then the users will need to 'cut off' the links that are redundant, choosing the ones that suit their needs best of all. This can be done based on the reliability indices, or history of previous use, or personal preference.

(2) The use of expert rules, which are either automatically acquired or created by human users, to guide the simulation-modelling process. These are the various principles and rules that apply to module and data scale, reliability, compatibility, and relevance.

(3) The intelligent use-tracker, a system that registers previous instances of module uses and success or failure associated with that. This is basically the growing knowledge base of previous experience acquired from using different modules for various applications.

The success of a participatory modelling project depends upon the transparency and usefulness of the model that is created. The model and the modelling system used cannot be overly complex to bar users with low computer skills from accessing and using it. There should be several levels of access that would match the needs of various types of users, from the very general public type, that provides plug-and-play and scenario run functionality, to the sophisticated users who may need to extend the system with their own modules and expert rules.

Major steps in developing the system include the following:

- Develop a registry of commonly used and tested modules, and wrap the modules within a common model-coupling framework such as CCA [2007] or OpenMI [2007]. This will allow discovery and guided composition of dynamic, spatially explicit, models.
- In coordination with federal, state and local agency partners, develop or integrate available online data access and conversion services, to ensure that data retrieved from several common federal sources of environmental information are formatted as valid local model inputs (for observations data in particular).
- Create an online mapping interface providing browsing and query access to online metadata repository for the data typically required by the simulation models.
- Integrate the participatory modelling system in a portal environment, to support personalized storing, analyzing, documenting, and annotating models and simulation runs computed for specific locations and scenarios.

Most important, we need to build the community of users and developers who will collaborate on the project and contribute their modules, data sets, methods, applications, case histories, bugs and documentation. Building such a community requires openness and flexibility. One of the reasons that we start with this vision is because we invite input at this very early stage of project development and are prepared to modify the system based on this community input. Our hope is that if the community develops along with the project there are better chances that the product will suit the needs of the users.

REFERENCES

Babu, S. and Widom, J. 2001. Continuous queries over data streamsq. SIGMOD Record. 30(3): p. 109-120

Berk, R.A., Bickel, P., Campbell, K., Keller-McNulty, S., Kekky, E., and Sacks, J. (2000). Workshop on statistical approaches for the evaluation of complex computer models. Los Alamos National Laboratory. LA-UR-00-5034.

Berk, R.A., Fovell, R.G., Schoenberg, F., and R.E. Weiss (2001). The use of statistical tools for evaluating computer simulations - An editorial essay. Climatic Change. 51(119-130)

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. Advances in Water Resources. 16: p. 41-51

CCA (2007) - http://www.cca-forum.org/, www.llnl.gov/CASC/components/babel.html

Cormas (2007) - http://cormas.cirad.fr/indexeng.htm

Extend (2007). http://www.extendsim.com/index.html

Han, J. and Kamber, M. (2000). Data mining: Concepts and techniques. The Morgan Kaufmann Series in Data Management Systems, ed. Gray, Jim: Morgan Kaufmann Publishers. 500.

Kyem, Peter (2005). Of intractable conflicts and participatory GIS applications: The search for consenus amidst competing claims and institutional demands. Annals of the Association of American Geographers, 94:1:37-57.

Kuhnert, M., A. Voinov, R. Seppelt (2006). Comparing Raster Map Comparison Algorithms for Spatial Modeling and Analysis. Photogrammetric Engineering & Remote Sensing, Vol. 71, No. 8, August 2005, pp. 975–984.

LMF (2007) – Landscape Modeling Framework - http://www.uvm.edu/giee/IDEAS/lmf.html

Maxwell, T. and Costanza, R. 1997. A language for modular spatio-temporal simulation. Ecological Modelling. 103(2, 3): p. 105-113

OGC (2007a). Open Geospatial Consortium (OGC) Standards - http://www.opengeospatial.org/standards

OGC (2007). Open Geospatial Consortium Interoperability Program Policies and Procedures - http://www.opengeospatial.org/ogc/policies/ippp

OpenMI (2007) - http://www.openmi.org/openminew/

Sarawagi, S., Agrawal, R., and Megiddo, N. (1998). Discovery-driven exploration of OLAP data cubes. The 6th International Conference on Extending Database Technology: p. 168-182

Simile (2007) - http://www.simulistics.com/

Stella (2007). http://www.iseesystems.com/

van den Belt, Marjan (2005). Mediated Modeling: A systems dynamic approach to environmental consensus building. Washington DC; Island Press.

Voinov, A. and Costanza, R. (1999). Watershed management and the Web. Journal of Ecosystem Management. 56: p. 231-245

Wiekum, G., Monkeberg, A., Hasse, C., and Zabback, P. (2002). Self-tuning database technology and information services: From wishful thinking to viable engineering. Very Large Data Bases (VLDB) Conference. Hong Kong, China: p. 20-31

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

Sharing emergency information between Emergency Control Centres: the Project REACT

U. Delprato[a], M. Cristaldi[a], A. Gambardella[b]
[a]*IES Solutions Srl, Italy, www.i4es.it, www.react-ist.net, info@i4es.it*
[b] *National Corp of Firefighters, Italy, antonio.gambardella@vigilfuoco.it*

Abstract: Every year, thousands of emergency call-centres in the EU receive some 200 million calls from citizens in distress. In response to them, Public Safety Answering Points (PSAPs) dispatch ambulances, fire-fighter teams or police squads to help the callers. In Europe, response to the calls and interaction between different agencies are often uncoordinated. This holds true also between different departments of the same: information sharing between them is still typically based on faxes and phone calls. The European Project REACT (Reaction to Emergency Alerts using voice and Clustering Technologies) aims at creating a seamless way to allow Command and Control center of different agencies (or of the same agency in different locations) to share data in electronic format. Key technical drivers of the project are the CAP protocol, the TSO data dictionary, a distributed web service based architecture and the AtomPub protocol. REACT started circa one year ago and has currently completed the design phase.

***Keywords**:* Common alerting protocol (CAP); TSO; PSAP; E112; semantic clustering; GIS; voice recognition; Emergency management.

1. INTRODUCTION

Fire brigades, emergency medical services and police are confronted with an increasing number of natural disasters and incidents on the one hand, and the trend towards bigger Public safety answering points (PSAPs) serving larger areas or integrating different emergency operators (typically Fire fighters and ambulances) on the other hand. An increasing importance is given to systems that assist call takers and dispatchers in getting a fast overview on incidents, either answered locally or coming from a neighbour agency. The overall situation in Europe shows a large number of emergency services using different command and control systems and not sharing information in electronic format, yet. Figure 1 synthesizes the current situation, where different PSAP or Emergency Control Centers (ECC) get different "inputs" from citizens and use phone calls of faxes as interoperability means.
This produces negative effects both in the response time and in the possibility that the right recipient of the piece of information is actually addressed and informed about an event. As an example, during the murderous Sarno (Italy) mudflow, several notification calls of minor precursory flows were addressed from on-spot citizens to police stations or municipalities: however, such precious and precise information was not addressed to the right decision maker. Despite the inherent local geological conditions and extreme rainfall, there is evidence that the high number of casualties was partly due to a lack of a unified information repository, available to emergency crews as well as emergency "intelligence" managers.

1.1 The European project REACT

The project REACT (Reaction to Emergency Alerts using voice and Clustering Technologies), funded by the European Commission under the Sixth Framework

Programme, addresses these problems by designing and implementing a working prototype that would allow:

- A seamless share of information between different agencies in electronic format
- A reliable voice (and language) recognition for capturing more information from caller/call takers conversations
- An intelligent "clustering" of apparently not related incidents into a dynamic scenario, being, this way, a decision support tool based on a large(r) group of calls.

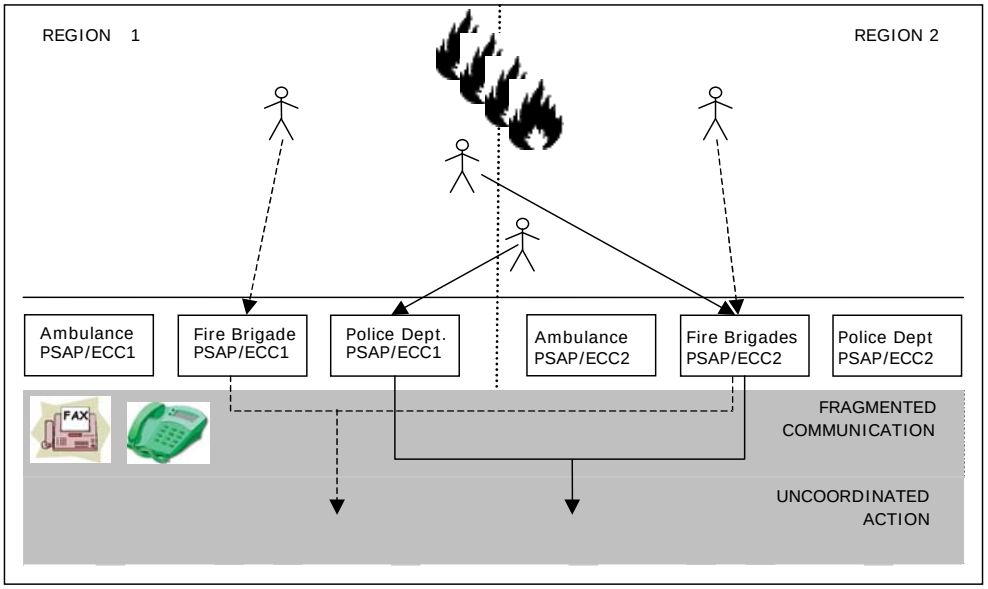


Figure 1. Communication between ECCs: today

The emergency services participating in the project have set the target for a system that will increase both efficiency and efficacy of work if, particularly when a lot of calls are received for the same event (potentially using different emergency numbers) and small incidents may go underestimated and escalate in scale. Figure 2 shows how the response to an event may be improved once REACT will be implemented and available.

The project partners are developing semantic technologies for call clustering and prioritisation, based on sensors (weather, air pollution, etc.), time and location of emergency calls, as well as keywords identified during the emergency call taking. Interoperability is implemented by using secure XML GIS-based data exchange between PSAPs. Beside the active participation of emergency services and experiences of other partners in this field, an active interaction with further emergency services all over Europe is actively sought to develop a system that would accommodate user needs, current processes, and infrastructures used for handling emergency calls at European level. It is worth to emphasize that REACT is designed as a system layer additional to existing computer aided dispatch and call taking systems, and it is not intending to replace them.

Being a shared effort of 11 partners spanning over two and half years, REACT is quite ambitious. This paper is focussed on the information sharing aspects, with the aim of describing how an event received and first managed by PSAP can be shared and displayed on a different PSAP system.

2. REACT ARCHITECTURE

The core architecture for REACT is based on a distributed model that supports efficient and performance optimised data exchange, coupled with platform independence. This type of architecture was chosen to satisfy requirements for deployment and assist integration with components from many different partners.

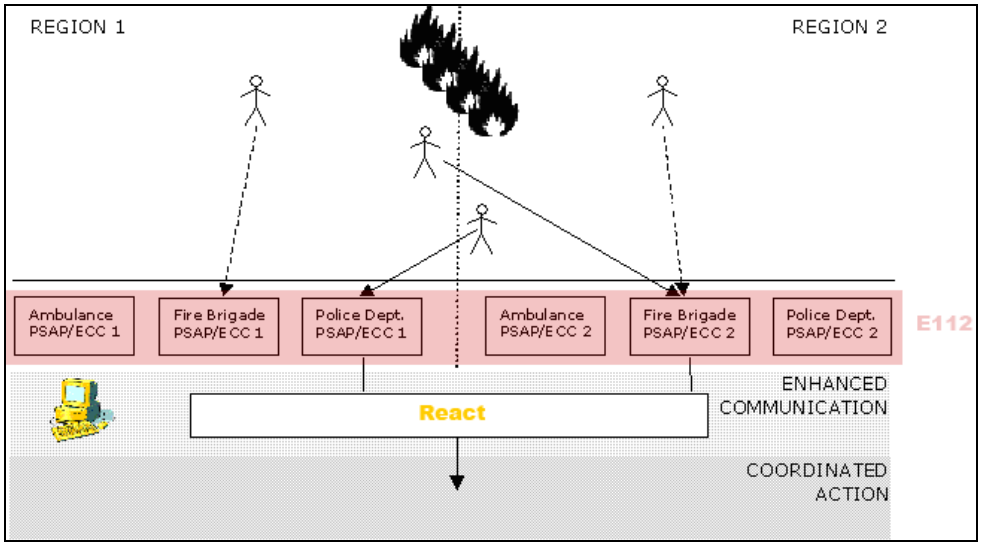


Figure 2. Communication between ECCs: with REACT

The communications backbone is provided by XML-based Web Services, to facilitate the interconnection of components using clearly defined APIs. Web Services were chosen because they are an industry standard way for disparate components to interoperate, independently from the used operating system or programming language.
Security is of paramount importance to protect the information transmitted between components or between REACT systems. Encryption is applied to the messages passed through Web Services for integrity and protection.
Because of the existing limitations in the modifications that can be brought to existing legacy infrastructure at each user site, a customised interface will be employed to plug their current Command and Control software into the REACT system.
Key elements of the REACT architecture are:

- Loose coupling between components
- Uses standard TCP/IP based protocols and links
  - XML
  - Web Services
  - Atom feeds
- Common Alerting Protocol as Incident Message format
- TSO as Incident Dictionary

Figure 3 shows the complete REACT architecture, with basic roles and responsibility of each components clearly depicted. The next sections of the paper will be focused on the Routing component.

2.1 Main Interfaces to REACT

Command and Control Legacy Interfaces (x3)
They represent the REACT interfaces to the Legacy Command and Control systems (i.e. the software currently used by REACT end users). The REACT consortium includes three end users, representing three typical examples of European Emergency Agencies:

(i) Italian National Corp of Firefighters (a single organization covering the entire Italian country)
(ii) Fire Brigades of Aachen, Germany (a brigade serving the municipality of Aachen in Germany, located close to the boundary with The Netherlands and Belgium)
(iii) Sussex Police Authority, United Kingdom (serving an English County)

The Interfaces are built on users specific needs and the REACT users will continue using their current existing Command and Control systems. To minimize the changes to existing working procedures, specific interfaces will be created, where the core functionality will be to create, update and delete events.

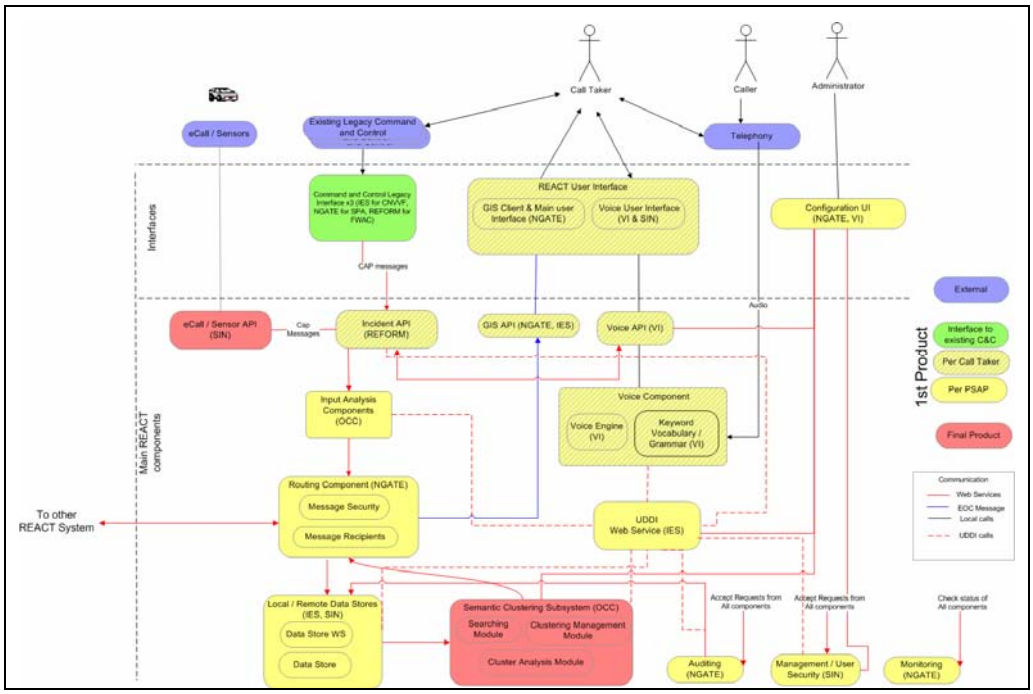


Figure 3. REACT Conceptual Architecture

REACT User Interface

The REACT User Interface includes:

- GIS Client and Main User Interface – displays incidents on a map. It connects to REACT using the GIS API.
- Voice User Interface – allows PSAP operators to check the incident data from the Voice Engine and make changes if required.

Configuration User Interface

A set of screens for configuring the REACT system on a per install basis, in order to fine tune the settings. Configuration settings include:

- Message Distribution Rules – for a given REACT system, defines what information to share with other connected REACT systems.
- Message Filters – defines which part of a message should be encrypted before sending, in order to reduce the viewing scope.
- Voice Configuration – settings related to voice integration with the telephony system.

2.2 Main REACT Components

eCall / Sensor API (Final product)

Gets emergency call information from eCall / Sensors and forwards it to the Incident API.

Incident API

Receives incident data from the legacy Command and Control systems and the eCall / Sensor API and forwards them to the Input Analysis component.

GIS API

Connects the GIS Clients to REACT.

Input Analysis

Receives incident details from the Incident API. Analyses each incoming incident to determine potentially interested recipients and forwards the analyzed data to the Routing component for address resolution and dispatch to other REACT systems.

Voice Component

Receives sampled speech from the Voice API and performs voice recognition analysis. Sends the results back to the Voice API.

Voice API

Gets speech from the emergency call (telephony) and sends it to the Voice Engine. Receives voice recognition results from the Voice Engine and transfers them to the Voice User Interface.

Routing Component

Receives requests from local or remote REACT components. Filters outgoing messages based on recipient, provides message logging, and dispatches messages to defined recipients.

UDDI Web Service

Web Service that stores and provides information about the location of other REACT components.

Local / Remote Data Stores

Stores incident data and other data received from the Routing Component, and makes the data available to other modules. Notifies interested modules when the data changes.

Semantic Clustering Subsystem (Final product)

Provides all the REACT searching and clustering functionality, notifying PSAPs of potential clusters and providing search results when requested to do so.

Auditing

Audits all the messages that pass through REACT, and all configuration changes.

Monitoring

Monitors the status of REACT components.

Management / User Security

Manages the REACT system. Applies filters, priorities and security settings.

3. REACT ARCHITECTURE COMPONENTS ALLOWING DATA SHARING

Figure 4 shows the flow of information in REACT for when a new incident is created. The diagram shows the flow of information both during the emergency call and after the call has finished.

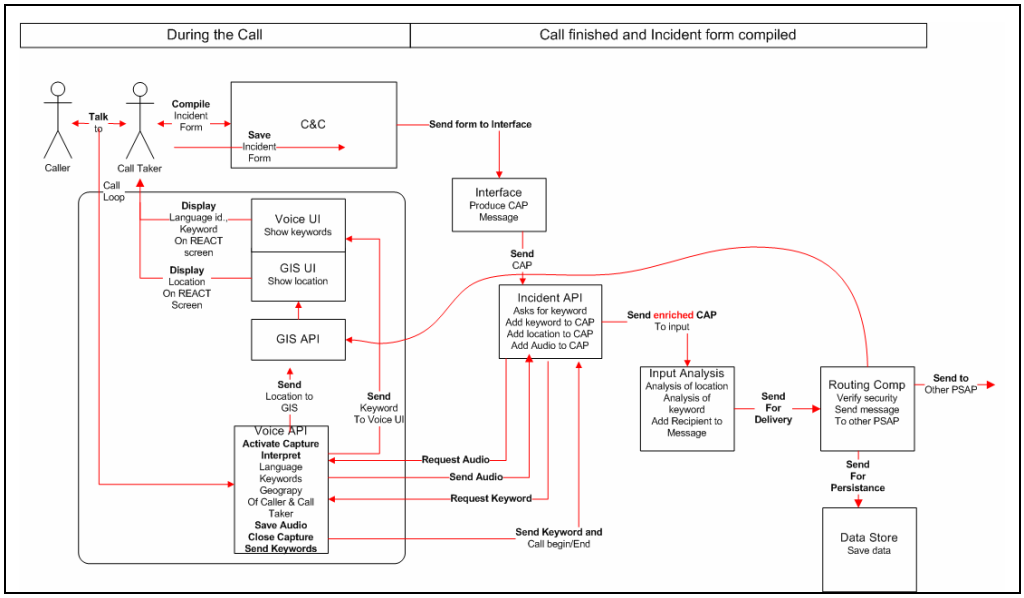


Figure 4. REACT Conceptual Sequence Diagram

During the Call:

- The emergency phone at the PSAP rings. The Call Taker and Caller talk to each other about the incident.
- The Call Taker enters data into an Incident form.
- The Voice API activates and captures language, keywords and geographical location of the caller and Call Taker.
- The Voice API sends keywords to the Voice UI, which displays LanguageId, keywords to the Call Taker on the REACT screen.
- The Voice API sends the location to the GIS API, which sends details to the GIS Client to display the location on the REACT screen

Call finished and Incident Form completed:

- A completed Incident form is sent to the C&C Legacy Interface, which converts it to CAP message format and sends it to the Incident API.
- The Incident API receives audio from the Voice API after requesting it.
- The Incident API receives keywords from the Voice API after requesting them.
- The Incident API sends an enriched CAP message to Input Analysis.
- Input Analysis analyses the location and keywords. It determines potentially interested recipients and forwards the analysed data to the Routing component.
- The Routing component verifies security, and sends the incident details to other PSAPs and the Data Store for persistence.
- The Data Store saves the received incident data.

Figure 5 shows the flow of information for incoming incident details from an external REACT system.

- The Routing component receives incident details from an external REACT system.
- The Routing component verifies security. It sends it to other PSAPs and the Data Store for persistence.
- The GIS API sends incident details to the GIS Client User Interface to be displayed on the Call Taker's and Dispatcher's REACT screen. The GIS API will be able to dispatch information to two different GIS User Interfaces: a thick client, provided by one of the Consortium partners and a thin client (a web browser based on Google Maps that will exploit KML and, potentially, GEORss data formats).
- The Data Store saves the received incident data.
- The Data Store sends the incident details to the Semantic Clustering component to update the cache. The Semantic Clustering component runs Geo and Semantic algorithms.

3.1 Focus on the Routing Component

From the REACT perspective, every system that is able to produce messages in CAP format can make use of REACT services. The real entry point for REACT is the IncidentAPI component: ECC legacy interfaces are offered to the final users as a set of API that can be used to convert legacy data into CAP format. This will allow a simple exploitation of REACT and will make the whole system attractive for other potential users.
The sequence diagrams above are reported to focus on the component that allow interoperability between two different REACT system: the Routing Component. It has been conceived as an Atom feed built upon Apache Abdera. The Atom Publishing Protocol, known also as the AtomPub protocol, is a way to publish and manage collections of resources using the basic HTTP GET, POST, PUT, and DELETE operations. While originally designed as a way to post new entries to weblog software, the Atompub protocol is well suited as a way to manage nearly any kind of Web-based content.

Web feeds allow software programs to check for updates published on a web site. To provide a web feed, a site owner may use specialized software (such as a content management system) that publishes a list (or "feed") of recent articles or content in a standardized, machine-readable format. The feed can then be downloaded by web sites that syndicate content from the feed, or by feed reader programs that allow Internet users to subscribe to feeds and view their content.A feed contains entries, which may be headlines, full-text articles, excerpts, summaries, and/or links to content on a web site, along with various metadata.

The development of Atom was motivated by the existence of many incompatible versions of the RSS syndication format, all of which had shortcomings, and the poor interoperability of XML-RPC-based publishing protocols. The Atom syndication format was published as an IETF "proposed standard" in RFC 4287, and the Atom Publishing Protocol was published as RFC 5023.

In REACT, the AtomPUB protocol is used within the Routing component to create and diffuse CAP messages. Other components create CAP messages into the Routing component and any other subscribed components will be acknowledged of the presence of a new CAP message. The routing component offers also security and message encryption.

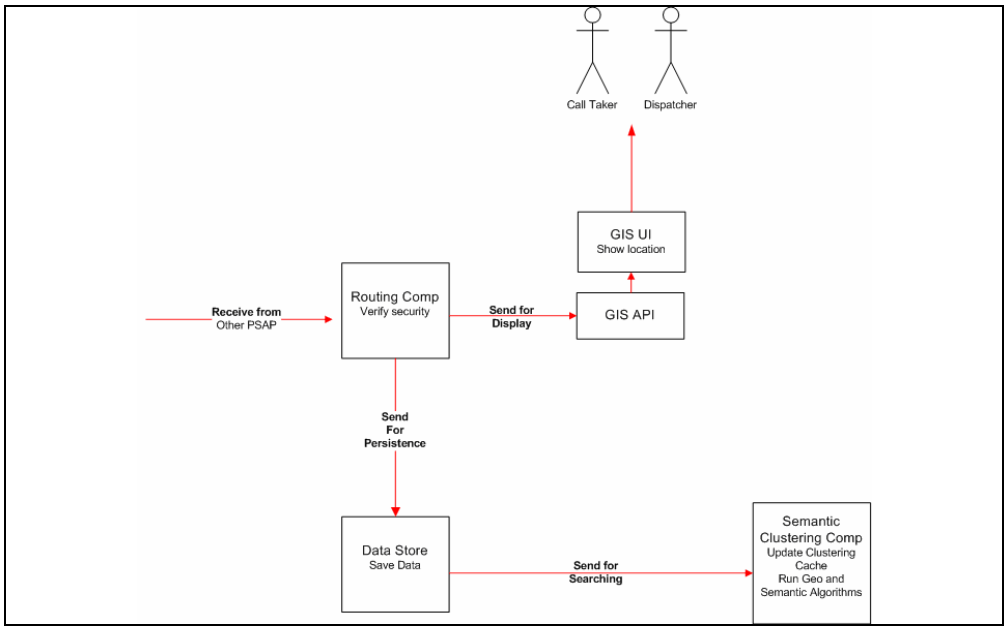


Figure 5. Flow diagram – Incoming incident

3.2 CAP protocol in REACT

The Common Alerting Protocol (CAP) is an XML data format used by many different organization in the to exchange information about a broad range of warnings and incidents. The CAP standard defines a type of document called an alert, which is used to exchange information about geological, meteorological, public health and safety, rescue, law-enforcement, environmental, transportation, infrastructure, and terrorist warnings and events. Such alerts can be generated either manually by incident responders or automatically by monitoring and sensing equipment, and they can be distributed using a variety of means.
In our case incident data are extracted from Command and control legacy system and converted into CAP format by the Command and Control legacy interfaces. They give the possibility to map each field in C&C system into a CAP field.

3.3 TSO in REACT

The Tactical Situation Object (TSO) has been adopted by the EU-OASIS project to increase the interoperability level between several agencies that are working jointly.

This object is one of the key means to reach a minimum level of interoperability between agencies during the disaster and emergency operations. This minimum level is defined for the purpose of the 1st version of the TSO, but in the future, the TSO could be extended progressively, allowing agencies to collaborate more efficiently during operations by sharing a timely and comprehensive common operating picture.
The TSO provides the capability to exchange pieces of information but it is not intended to provide all detailed information. There is a full flexibility in defining the mapping how each field in the ECC systems maps against the TSO data dictionary .

TSO in REACT has been selected ad a data dictionary to allow interoperability between agency. CAP will be used as the data format that will “carry” also TSO messages.

4. CONCLUSIONS

A key aspect of Emergency management is the communications between emergency services. Almost every emergency arising from everyday accidents requires the intervention of two or even three emergency services. Ambulances must help victims and police has to regulate traffic around the scene of the accident or must start investigations in

case of criminal acts. Fire-fighters are often required to liberate victims from wrecked cars or from debris. Of course they are the real protagonists in the case of fires – where the other emergency services are also required. All these interventions require communications between the emergency services involved. This becomes imperative in case of major incidents and disasters covering wide areas and necessitating the intervention of emergency services from different local or regional authorities or even (in the case of disasters with a European or international dimension) of multinational, multidiscipline teams.

Improvements are clearly needed for ensuring that the call to the 112 is "appropriately answered and handled". This involves many issues on the front-end of the PSAP (particularly multilingual support, caller location and verbalisation of the verbal communication), but also interoperability and storing/retrieval of information is a key point REACT aims at creating a favourable environment for an efficient management of local emergencies, where still a timely and focused intervention saves lives, but the communication and command chain is shorter and more direct. This requires anyway the setting up of an interoperable system, allowing an efficient exchange of data and a common representation of the needs of the person in distress.

With REACT in place PSAPs will be able to share information in a seamless way, only provided that their current Command and Control softwares allow an "incident-per-incident" extraction of incident data in CAP format. The project will hence demonstrate that an integration between emergency actors is possible and, will pave the way for triggering discussions at coordination, management and political levels.

In this article, the REACT architecture has been described, along with the motivation brought in by the finalk users in terms of usability, interoperability and functionalities needed for a successful future integration in their current legacy systems.

The key contributions of REACT to the current emergency communication space are:

- Integrated use of both CAP and TSO as enhanced protocol for representing and sharing information about events
- Integration of the emergency caller location featured by the E112 and eCall – reducing response times.
- Call clustering and prioritisation through semantic, time, location analysis allowing emergency services to identify and consolidate incoming data – leading to clearer decision making and support during cross-border incidents (4D GIS).
- Automatic identification of keywords from the PSAP operator speaking supporting the filling of a job description forms (including multilingualism)
- The adoption of the Atom Publishing Protocol for ensuring interoperability across call centers and emergency services.

ACKNOWLEDGMENTS

REACT is co-funded by the European Commission, Framework Program 6, IST priority, under the contract IST-2005-033607 "REACT - Reaction to Emergency Alerts using voice and Clustering Technologies". The project has started in September 2006 and will run for 30 months.

REFERENCES

Snell, J., An overview of the Atom 1.0 Syndication , *developerWorks*, 2005
Snell, J., The Getting to know the Atom Publishing Protocol series, *developerWorks*, 2006
The CAP Cookbook:
http://www.incident.com/cookbook/index.php/Welcome_to_the_CAP_Cookbook
The CAP V1.1 specification:
http://www.oasis-open.org/committees/download.php/15135/emergency-CAPv1.1-Corrected_DOM.pdf

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 a simulation-based study of grassland and animal diversity management in livestock farming systems

G. Martin[1,2,*], M. Duru[1], R. Martin-Clouaire[2], JP. Rellier[2], JP. Theau[1], O. Thérond[1], L. Hossard[1]

1: INRA, UMR 1248 AGIR, F-31326 Castanet Tolosan, France;

2: INRA, UR 875 UBIA, F-31326 Castanet Tolosan, France;

**: guillaume.martin@toulouse.inra.fr*

Abstract: In less-favored areas, livestock production involves the management of native vegetation as a resource base. A challenging issue in such systems regards the design of efficient and sustainable management strategies that make it possible to exploit the diversity of grasslands and herd batches over different time scales. This paper describes the conceptual approach that supports the representation and simulation of farm-scale models of grassland-based livestock systems in the SEDIVER project. It enables the coupling of species-rich grasslands and livestock models with sophisticated decision-making models. The latter embed a representation of management strategies that specify in a flexible and adaptive manner what activities are intended to be done. Once implemented, the SEDIVER model will provide guidance for a robust and improved exploitation of grassland and animal diversity at farm scale.

Keywords: Conceptual model; grassland/animal diversity; farm management.

1. INTRODUCTION

In less-favored areas e.g. semi-mountainous regions, agriculture plays a crucial role among other things for biodiversity conservation. In such areas, livestock production involves the management of perennial species-rich grasslands as a resource base (grazing or fodder) for animal production. Farmlands quite often include a diversity of perennial grasslands (regarding productivity, plant diversity and topographic/farming properties). It is now widely recognized that plant and/or grassland diversity has potentialities at field and farm scale (e.g. White et al., 2004) regarding a number of aspects e.g. flexibility gain in management, improved tolerance to extreme climatic events, complementarities between species... However, managing grassland diversity is a rather complex issue as it implies satisfying production and management objectives on different temporal scales and coping with farmland spatial heterogeneity which is characteristic of less-favored areas. Farmers need to satisfy animal feed requirements, which requires dealing with weather uncertainty. Taking into account long-term consequences, farmers have to manage carefully the technical operations on each field so as to ensure the sustainability of vegetation properties. For a given form of land use, depending on the grassland community characteristics, there is an optimal time window for using the grassland resource. Definition of this time window is based on biomass availability and/or herbage quality according to the herd feeding objectives. When organizing their production system, farmers also have to cope with the spatial heterogeneity of the farmland i.e. topographic and farming properties of individual fields (ease of access, suitability for mechanization…) as well as fields' potential production permitted by vegetation types. The combination of these two variables determines the single or various forms of land use a field can fulfill in the feeding system.

Herd batching that exploits animal diversity is of major interest to diminish herbage losses and to exploit all kinds of grassland communities by matching animal batch needs to herbage feeding value. Batch re-composition can also be used as a regulatory process to cope with yield variation. Consequently, as demanded by farmers and farm advisors, bio-technical and organizational guidance is needed to improve current management practices. This entails developing and integrating on a farm scale a coherent set of management practices for each grassland community type and herd batch over different time scales. Such integration can be used to identify particular farming system management strategies suited for specific contexts, and to characterize the consequences of these potential re-organizations on the land use, given the grassland community types present.
Forward looking strategies are needed in order to support adaptability and reactivity of the farming systems. Such management strategies might be obtained by harnessing organizational flexibility and biological buffering capabilities. We define organizational flexibility as freedom in the implementation and modification of a management strategy under a given set of topographic and farming constraints. We call biological buffering capabilities the farmer's ability to modify the target performances or the state of the plant and/or animal material; such changes of plant or animal target state have to remain within a specific feasibility range to benefit from the buffering capacity of these biological entities. Such buffering capabilities differ according to categories for these biological entities based on selected indicators, for instance functional plant ecology markers in this paper. When looking for such new management strategies, one should consider concomitantly those two dimensions, i.e. the exploitation of biological buffering capabilities and organizational flexibility, to devise management strategies that make good use of this leeway. In this context, farm scale simulation modeling is of major interest to study the consequences of management changes on simulated farming systems. For achieving the design of farming systems, a lot of attention has been given to improvement of biophysical simulation models. However, it is to a large extent the complexity of setting up coherent management options that make farming systems so complicated (Thornton & Herrero, 2001). This is why a number of authors (e.g. Garcia et al., 2005) call for more elaborate consideration of management aspects and for a better coupling between biophysical and decisional models to study farmer's management and its consequences on the biophysical system, such as in FARMSCAPE (Carberry et al., 2002). Research has produced several simulation-based farm models for beef and dairy farming systems such as SEPATOU (Cros et al., 2004) focused on rotational grazing management. To our knowledge, no example has focused on the efficiency of grassland and animal diversity exploitation through innovative management strategies centered on organizational flexibility and on the use of biological buffering capabilities. The present paper provides basic insight into a conceptual description of the SEDIVER approach, a farm-scale modeling and simulation approach designed to study the dynamics and interactions between biophysical and management processes in different weather scenarios. This conceptualization step is a necessary preliminary to the simulation-based investigation of the behavior of specific farm instances. At the core of this step lies the representation of the production management behavior and its linkages with the biophysical system, in particular grassland and animal diversity. Section 2 describes the background knowledge involved in the modeling process. In Section 3, we outline the necessary steps towards the design of a management strategy. Section 4 presents the main features of the SEDIVER conceptual model.

2. BACKGROUND KNOWLEDGE INVOLVED IN THE SEDIVER APPROACH

2.1. Simulation-based study of management strategies

The complex interaction between biophysical processes and those under human control is at the very heart of agricultural production. As a production manager, the farmer makes decisions about land use and the timing, combination and implementation of the associated technical operations (e.g. harvesting, feeding animals, etc.) in the hope of achieving his objectives. Clearly the important question of production management, dealing with risk control, change as new practices, products and techniques appear requires innovative

approaches that recognize and focus on the holistic, dynamic and human dimension of farming systems.

Management strategies are ways of structuring things one intend to do, or ways in which something is done. It deals with the organization of work activities across time and space. Studying management strategies implies a strong emphasis on what activities are deemed relevant for a production objective, what is the interdependence between them, what are the preconditions to their enactment and how they should be structured in time and space to tackle the constraints to the desired outcome. Prior to the execution of technical operations, management can be broken down into two steps: configuration and planning. Configuration consists mainly in sizing production units and herd batches concomitantly. Planning leads to determine the relevant technical activities and the key dates and conditions involved in their structuring in a production plan. Farmers implement on a field-scale management practices that they have devised by integrating considerations on a farm scale. To deal with the spatial dimension of farm management decision-making, Coleno and Duru (2005) introduced the concept of a production unit, i.e. the functional entity characterized by a single land use on which a farmer works out his management practices. A livestock farm is therefore considered as an aggregation of spatially organized production units.

Elaborate decision-making models rely on the decision-making context, for instance the farmer's material or physical constraints, his beliefs about what biophysical indicator is relevant to a decision or his personal know-how and preferences. In some cases, the complex issues of dynamic work scheduling and resource allocation may need to be considered. Thus a farm management strategy is perceived as the specification of a temporally-structured decision process, together with its context-dependent adaptations (Cros et al., 2004).

Computer-based simulation is one of the commonly approach used to aid in the design and evaluation of production management policies (FARMSCAPE: Carberry et al., 2002, SEPATOU: Cros et al., 2004). Simulation approaches have traditionally focused on isolated agronomic and technological aspects of the production processes, e.g. crop or animal responses to particular farming operations. Surprisingly, little attention has been paid to the realistic modeling and simulation of management strategies of farmers. Studying and supporting the design of management strategies by means of computational tools has rarely been addressed directly and systematically as an issue in its own right (Garcia et al. 2005). This is precisely what the SEDIVER project is intended to do in the case of grassland-based farming systems. Any SEDIVER model is an instantiation of a SEDIVER metamodel that offers abstract representation patterns enabling to capture the description of a particular farm configuration and management strategy set up to control the production process. The metamodel relies on an ontology dedicated to the domain of production management. An ontology (Smith, 2003) seeks to describe or posit the kinds and structures of the concepts, properties and relations in the domain of interest. This ontology takes the form of a set of classes that describe the structural components of the domain and a set of procedures that describe the inferential mechanisms applicable to instances of these components and representing how they might change over time. The ontology has been implemented in the modeling/simulation framework, called DIESE (Martin-Clouaire and Rellier, 2003). The central concept of the ontology is the notion of activity, the characterization of an operation to do on something. An activity has some attributes that specify, in particular, its relevance in a working context and the triggering conditions that control its execution status. A set of operators are used to represent various kinds of dependences among activities. These operators enable to define composed activities by specifying sequencing, concurrency, synchronization or delay enforcement. They also include a kind of programming construct able to specify for instance that an activity should be iterated or that an activity is optional.

Composed activities may describe nominal plans with the inherent and necessary specification of the sequence of activities that will eventually be executed. The vagueness of a plan is not a fault, but provides the flexibility needed to cope with the huge number of actual circumstances unfolding during its situated enactment, especially in agriculture where uncontrollable factors play a key driving role in many production processes. On the other hand, a nominal plan may encounter situations where it is beyond its bounds, notwithstanding its flexibility. To account for this need, our ontology is equipped with the concept of conditional adjustment.

2.2. Functional approaches to cope with the diversity of grasslands and animals

Perennial grasslands are complex agrosystems to model as they include a diversity of species. Recent simulation models describing herbage growth and/or digestibility developed for species-rich grasslands adopt a functional representation of grassland vegetation (e.g. Jouven et al., 2006). The functional composition is based on the definition and the measuring of value, range, and relative abundance of plant or animal functional traits (morphological, physiological and phenological) in response to availability of resources and perturbations. Thus, in these recent simulation models, vegetation-related inputs are provided in terms of grass functional groups distributed along a leaf dry matter content (LDMC) gradient. On-field grasses weighted mean LDMC ranking is well correlated with agronomic characteristics like organic matter herbage quality and timing of herbage growth pattern in relation to plant phenology and leaf lifespan (Al Haj Khaled et al., 2006). Vegetation agronomic characteristics are mainly determined by management i.e. practices regarding resources availability and land use (grazing/cutting). Based on such concepts, grasses weighted mean LDMC at plant community level provides a powerful descriptor of grassland vegetation for modeling the dynamics of growth and digestibility. In addition, it is especially suited for characterizing the time course of herbage production and consequently the latitude offered by each vegetation type to exploit its buffering capabilities through the time windows for grazing or cutting.
Functional classification of cattle categories has led to the identification of animal classes as proposed in dynamic models of herd demography. Based on this functional classification of animal categories and breeds, available feed evaluation and rationing system for protein and energy can be used to model animal feed intake and animal production.

3. DESIGNING A LIVESTOCK FARMING SYSTEM

The SEDIVER approach puts emphasis on decisions at the operational level at field and herd batch scales to evaluate management strategies at farm scale. Thus, the two dimensions of a management strategy, configuration of the system and initial planning of the activities, are model inputs that have to be designed prior to simulations and then incorporated in an instantiation of a SEDIVER meta-model. Configuration and planning must take into account uncontrollable aspects such as the topographic properties of the fields (e.g. area, slope) and induced farming constraints (e.g. suitability for mechanized harvest), agronomic characteristics of the fields (e.g. initial vegetation type) as well as general objectives related for instance to labor or profitability. Figure 1 outlines how the above aspects affect the design of a management strategy.

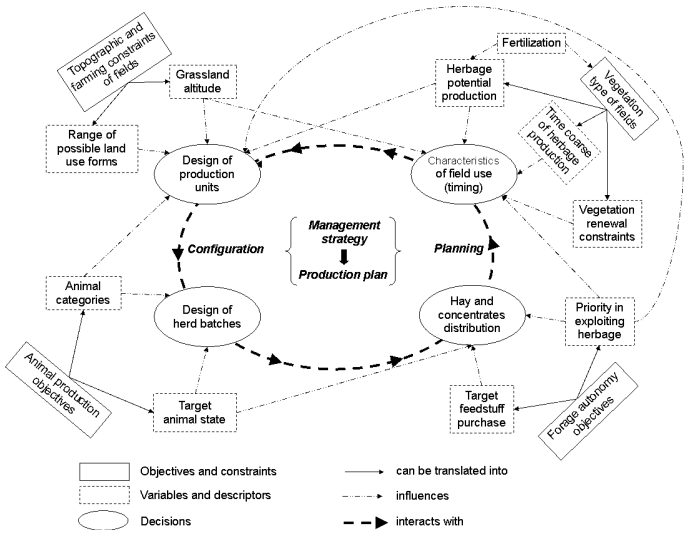


Figure 1: Factors involved in configuring and planning a livestock farming system to build the production plan used as inputs in a SEDIVER simulation run.

3.1. Base configuration of the farm system

When configuring a livestock farming system, considerations of topographic and farming constraints of the fields provide a range of possible forms of land use on these fields. One may set up groups of candidate fields likely to be allocated to the different production units. In addition, animal production objectives that can be translated into herd composition (animal categories and target animal state) affect the batching options. Batching decisions aim to optimize feeding conditions for the various animal categories to achieve their expected target states. Farmers may consider grassland community types to match herd feeding objectives to the potential herbage production permitted by these vegetation types, taking into account that fertilization is a leeway to modify this herbage yield as well as the encountered vegetation type in case of lasting change of fertilization practices. The farmer's priority in optimally exploiting herbage may also influence the configuration. For instance, if optimally exploiting herbage is not among a farmer's priorities, he may form production units with an above-average stocking rate and then feed bought-in hay and concentrates if necessary.

3.2. Planning and exploitation of organizational flexibility and biological buffering capabilities

The configuration decisions yield a set of candidate fields for each production unit. Planning restricts these sets by introducing timing considerations. To this end, grassland community types provide indications on the time course of herbage production and consequently on the thermal time windows for grassland use whereas field altitude can be used to map these thermal time windows to periods expressed in Julian dates. Thus farmers get an idea of the Julian period at which herbage will be available for given types of land use. They can arrange their grazing and cutting sequences by considering the successive time windows for using the candidate grassland fields identified in the configuration phase. Moreover, vegetation types and fertilization practices provide information on the attainable yield at any given moment. For a single field, successive activities can correspond to different forms of land use and therefore to different durations for herbage re-growth on the field. For instance, after the first cutting activity, a field can join a grazing sequence or it can be harvested a second time. This corresponds to re-growth durations ranging from one to twice. In this way, the farmer organizes the sequences of grazing and cutting activities over time and space. Depending on the priority he gives to exploiting herbage resources, he may decide to counterbalance a non-optimal timing for field use by an adjustment of fertilization practices or by the distribution of hay and concentrates. This may influence decisions about activities related to the distribution of hay and concentrates. In order to ensure the renewal of grassland vegetation, the timing of the grazing and cutting activities should be kept quite stable from one year to the next.

The result of the planning step is a plan of activities, a set of conditions that control their execution and, if needed, some state-dependent adjustments that enables plan or configuration adaptations in exceptional situations. The conditions may concern dates or restrictions for the execution of an operation, e.g. sward height of the next field to be grazed and phenological stages. It thereby includes information about the intensity of grazing and the range of intensities associated with a grazing activity performed on a field by a herd batch. In addition to the constraints imposed by the sequencing of activities, each activity is usually assigned opening and/or closing conditions such as a time window between two dates, a biophysical condition or a threshold (e.g. a degree day sum). For instance, opening-closing conditions can be used to differentiate between "normal grazing" and "topping" which actually are both grazing activities. Topping is a light grazing thereby corresponding to higher residual biomass when specifying closing conditions assigned to this particular grazing activity. More generally opening and/or closing conditions may involve any indicator, that is, any piece of information collected or synthesized for use in decision-making. Then the use of biological buffers lies in the specification of intervals of indicator values for activating production activities. These may be time windows determined according to the encountered vegetation types. Target animal state can also be specified as

an interval of body condition score to benefit from the biological buffer capacity of animals if needed.
Some fields are not fully committed at planning time. Their role is to enable the combined exploitation of organizational flexibility and biological buffering capabilities. Depending on the conditions, such uncommitted areas can be switched easily from one production unit to another. Such a switch has to remain in an acceptable range of frequency not to jeopardize buffering capabilities of herbage and to avoid irreversible side effects on grassland community characteristics. Organizational flexibility can also be exploited through batching in the course of the year. This may have consequences on stocking rates and subsequently on (i) the associated grazing practices thereby involving buffering capabilities of herbage, (ii) target performances of animals and their buffering capabilities. Suitability of fields for a range of land use forms, yield, quality and time course of the herbage production are considered in revising the base plan in the course of the year thanks to adaptability provided by such uncommitted areas. The latitude for revision of the production plan is largely dependent on organizational flexibility permitted by topographic and farming constraints.

4. OVERVIEW OF THE SEDIVER CONCEPTUAL MODEL

4.1. Biophysical system

Conceptually, the biophysical system is seen as composed of biophysical entities (fields, herd batches, etc.) that may have one or several descriptors (e.g. dry matter available, population size, etc.) and their own processes (e.g. plant development, animal intake and production). The processes are controlled by events such as a birth event which is programmed to occur every year and induces the firing of a process of creation of animal entities and a feeding process applied to these new animals. From a modeling point of view, the above entities, descriptors, processes and events can be defined as particularizations of ready-to-use DIESE classes and instantiated as needed for specific farm cases.
Fields are mainly described by a set of topographic and farming properties that determine the range of possible land use forms, a surface area, an altitude and a type of grasslands expressed through a descriptor of grassland vegetation (a grasses weighted mean LDMC). Then, the values of the descriptors (e.g. water stress index, leaf area index) involved in the processes governing the field dynamics are specific to each field. Daily plant growth and herbage quality dynamics are the main processes modeled at field scale using available models (e.g. Jouven et al., 2006). Linked to daily plant growth is sward height that make sense for farmers in grazing management. Thus, sward height is commonly used in the conditional adjustments involved in management strategies that deal with biological regulations.
Herd batches are mainly described by an average animal state and a size with target objectives for the modeled state variables (weight or milk production) over a given period that can be modified to account for the buffering capacity of cattle in the system, e.g. when the farmer decides to decrease animal state targets so as to save available herbage during drought. Simulated processes are daily intake, evolution of daily energy requirements, energy conversion and production (milk and/or fattening) for each animal batch using the INRA fill unit system. Moreover an available but unpublished demographic module will be used to simulate the ageing process of animal categories.
Hay stocks are described by the nature of the feedstuff, its available amount, herbage quality, feed and fill value using available models (Duru and Colombani, 1992). Modeling available amount of hay stocks has an interest for accurately representing management strategies in the model. When stocks become scarce, farmers are used to modify their plan, e.g. by advancing in time the date for turnout. Concentrate stocks are modeled with the same variables.

4.2. Decision system

One of the most attractive features of the DIESE modeling framework is its capacity to easily represent flexible production plans. Making a model of an executable plan is a matter of instantiating the abstract primitive activities provided as a library and articulating them using the composition operators. SEDIVER activities include:

- grazing, that applies to a herd batch and to a set of fields visited in turn;
- distributing hay and concentrates, that applies to a herd batch;
- cutting that applies to a set of fields.

Taking the example (simplified for the sake of clarity) of a farm with three production units, Figure 2 provides an example of a plan written in a similar way to the plan representation of DIESE. The operator "`and`" in the first line permits the aggregation of the primitive or composed activities specified in each subsequent line. For instance, lines 3-7 refer to a sequence of activities of topping and iteration of normal grazing by the H1 batch and the cutting of the field F1 topped by this herd batch. The "`meeting`" operator on line 3 has two arguments that are composed activities, the first one (lines 3-5) starting with an "`include`" operator, and the second one (lines 5-7) starting with an "`and`" operator. Semantically, "`meeting`" specifies that these two composed activities are contiguous in time: right after the end of the first one, the second one is activated. Similarly, the operator "`include`" has two arguments consisting of composed activities, the first one being a "`meeting`" sequence and the second one an iterative hay-distribution activity. The operator "`include`" forces the opening period of the second one to be included in the opening period of the first one. It is used to specify that the activities referring to this operator are concomitant at least over part of the time period over which the first one is open. Coming back to the grazing activity of line 3, the H1 batch should first be put to graze on field F1. Right after this (according to the semantics of the "`meeting`" operator) the same herd batch should be moved onto field F2 if the conditions specified with the operator "`optional`" are met. In the meantime (according to "`include`"), the farmer should distribute hay repeatedly. Some opening and closing conditions (not shown here) restrict the period and/or situation in which this should be done. Once these one or two fields have been grazed, the second argument of the "`meeting`" activity becomes "open". The herd batch should be moved successively to fields F3 and F4 and this same grazing sequence should be iterated (as specified by the operator "`iterate`") until a termination condition (not shown here) is met. In the meantime, field F1 should be cut by the farmer. The operator "`before`" in line 10 means that the H2 batch should first graze fields F6 and F7 and then the farmer should consider the state of field F8 to determine whether this uncommitted field is grazed or cut after fields F6 and F7. Finally, the operator "`equal`" on line 15 means that both the grazing sequence on the fields F10 and F11 and the distribution of concentrates should begin and should end together.

```
1  and(
2  //--1st production unit-----------------------------------------------
3      meeting (include (meeting (grazing (F1, H1)),
4                                 optional(grazing (F2, H1))),
5                       iterate (distributing_hay (H1)),
6               and (iterate (grazing ({F3, F4}, H1)),
7                     cutting (F1))),
8  //--2nd production unit-----------------------------------------------
9      meeting (grazing (F5, H2),
10             and (iterate (before (grazing ({F6, F7}, H2)),
11                                   or (grazing (F8, H2)),
12                                       cutting (F8)))),
13 //--3rd production unit-----------------------------------------------
14     meeting (grazing (F9, H3),
15             and (equal (iterate (grazing ({F10, F11}, H3)),
16                         iterate (distributing_concentrates (H3))))),
17     )
```

Figure 2: Example of a plan written in a similar way to the plan representation of DIESE. F and H refer to Field and Herd batch and are followed by the corresponding Field and Herd batch numbers. The indentation helps in identifying the arguments of the composition operators (`and`, `or`, `meeting`, `before`, `include`, `equal`, `iterate`, `optional`).

Decision-making modeling can consist in strictly applying the plan when the weather conditions are within the envelope of the average year. On those fields that were not fully committed at planning time, we considered the switch of land use form as the exploitation of the latitude permitted by the base plan (using the "`or`" operator, line 11). However, in

some situations, more profound adaptations should be implemented using conditional adjustments. For instance, batching could be used to organize the herd into two batches only, thereby implying a complete revision of the grazing sequences planned. None such conditional adjustment is presented in this example.

5. CONCLUDING REMARKS

In this article (see (Martin et al., 2008) for an extended version), we discussed the management issue of exploiting grassland and animal diversity in a livestock farming system, and how to address this issue using a farm simulation model.

At the time where research projects focused on climate change flourish, we believe that the prime need of farmers is to get support in dealing with climate variability and especially extreme conditions. Such support might be obtained by comparing alternative management options through computer simulations. To cope with weather uncertainty, managing the within-farm functional diversity of grassland community types is of major interest from the practical, economical and ecological points of view. Moreover, farmlands especially of less-favored areas are characterized by a strong heterogeneity in resource use in particular herbage, which results in high variability of yield and income in time and space. Common farm-scale modeling approaches are based on farm-encompassing indicators such as stocking rate or amount of fertilizer per hectare. We argued that the design of agri-environmental schemes should explicitly consider the within-farm diversity of grasslands and their practical use.

Most research approaches dealing with the design of farm management strategies are typically based on linear programming models that suffer from too unrealistic assumptions with respect to the importance of dynamics and uncertainty in farming. Keating and McCown (2001) already suggested that challenges for farming system modelers are “not to build more accurate or more comprehensive models, but to discover new ways of achieving relevance to real world decision making and management practice.” In this sense, the SEDIVER project is the result of consistent efforts to improve the representation of farm management strategies and get closer to the questions raised in practice. In their day-to-day management, farmers are coping with unpredictable events which call for flexible plans and conditional adjustments that yield different sequences of actions depending on the course of events met. Indeed there are some turning point events in the course of the year at which the farmer has to review part of his base plan.

The conceptual model of a management strategy presented in this paper is inherited from the ontology underlying the DIESE modeling framework. The level of detail of the knowledge that is dealt with by SEDIVER is totally in line with the expressive power of the modeling framework. Moreover, DIESE plays an important role in guiding the model design through ready-to-use representation patterns. For instance, compared to the widely used “`IF…THEN…ELSE…`” approach, the formal representation of management strategies exploited by the SEDIVER approach provides an intelligible and rigorous conceptual framework that can capture timing dependencies and concurrencies between farming activities.

The metamodel SEDIVER presented in this paper is still to be implemented in the DIESE modeling / simulation framework. This essentially amounts to create the entities, processes and events involved in the biophysical system and to particularize the necessary notions of activities and operations of the grassland-based livestock system domain. These models account for different biophysical processes and production management processes and can be used to perform *in silico* (i.e. computer simulated) experimental investigation of the merits and limits of different grassland and animal management strategies.

Once implemented, this model will be suitable for modeling real farm cases. Still, dealing with real cases implies to cope with the specificities of each case and important data gathering for being able to run the model. Instead of trying to reproduce very closely real cases, the project will focus on characterizing and designing virtual innovative farm types. In the field of agronomy, research has almost always used farm types based on structural criterion. In the course of this project, to enrich the range of simulated cases and to cope with the complexity of the management choices observed on field, a typology based on the farming style concept (van der Ploeg, 1994) will be elaborated in partnership with experts

to be crossed with a typology based on structural criterion. By providing a quite detailed overview of the possible trajectories of farms on a territory, this combination of typologies coupled with the SEDIVER approach will provide the necessary elements to deal with land use issues at regional scale.

ACKNOWLEDGEMENTS

This study is partly funded by the French ANR ADD (Agriculture et Développement Durable) programs, in the framework of the projects TRANS (TRANSformations de l'élevage et dynamiques des espaces - ANR-05-PADD-003) and VALIDATE (Vulnerability Assessment of LIvestock and grasslanDs to climAte change and exTreme Events).

REFERENCES

R. Al Haj Khaled, M. Duru, V. Decruyenaere, C. Jouany, and P. Cruz, Using leaf traits to rank native grasses according to their nutritive value, *Rangeland Ecology & Management*, 59, 548-654, 2006.

P.S. Carberry, Z. Hochman, R.L. McCown, N.P. Dalgliesh, M.A. Foale, P.L. Poulton, J.N.G. Hargreaves, D.M.G. Hargreaves, S. Cawthray, N. Hillcoat, and M.J. Robertson, The FARMSCAPE approach to decision support: farmers', advisers', researchers' monitoring, simulation, communication and performance evaluation, *Agricultural Systems*, 74, 141-177, 2002.

F.C. Coleno, and M. Duru, L'apport de la gestion de production aux sciences agronomiques. Le cas des ressources fourragères, *Nature Sciences Sociétés*, 13, 247-257, 2005.

M.J. Cros, M. Duru, F. Garcia, and R. Martin-Clouaire, Simulating management strategies: the rotational grazing example, *Agricultural Systems*, 80, 23-42, 2004.

M. Duru, and H. Colombani, Haymaking: risks and uncertainties in Central Pyrenees grasslands, *Agricultural Systems*, 38, 185-207, 1992.

F. Garcia, F. Guerrin, R. Martin-Clouaire, and J.P. Rellier, The human side of agricultural production management - the missing focus in simulation approaches. In: Proceedings of the MODSIM Conference, Melbourne, Australia, December 12-15, 2005.

M. Jouven, P. Carrere, and R. Baumont, Model predicting dynamics of biomass, structure and herbage quality of herbage in managed permanent pastures. 1. Model description, *Grass and Forage Science*, 61, 112-124, 2006.

B. Keating, and R.L. McCown, Advances in farming systems analysis and intervention, *Agricultural Systems*, 70, 555-579, 2001.

G. Martin, M. Duru, R. Martin-Clouaire, J.P. Rellier, J.P. Theau, O. Therond, and L. Hossard, Conceptualizing grassland and animal diversity management in livestock farming systems of less-favored areas – Overview of the SEDIVER approach. Submitted, 2008.

R. Martin-Clouaire, and J.P. Rellier, A conceptualization of farm management strategies. In: Proceedings of the EFITA Conference, Debrecen, Hungary, July 5-9, 2003.

B. Smith, Ontology, In: Floridi, L. (ed.), Blackwell Guide to the Philosophy of Computing and Information, pp 155-166, 2003.

P.K. Thornton, and M. Herrero, Integrated crop-livestock simulation models for scenario analysis and impact assessment, *Agricultural Systems*, 70, 581-602, 2001.

J.D. van der Ploeg, Styles of farming: An introductory note on concepts and methodology, In: J.D. van der Ploeg & A. Long (eds), Born from within: Practice and perspectives of endogenous rural development, pp 7-30, 1994.

T.A. White, D.J. Barker, and K.J. Moore, Vegetation diversity, growth, quality and decomposition in managed grasslands, *Agriculture Ecosystems & Environment*, 101, 73-84, 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

An Agent-Based Model for Exploring Land Market Mechanisms for Coastal Zone Management

Tatiana Filatova[a], Anne van der Veen[a,b], Alexey Voinov[c]
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
c – Chesapeake Research Consortium Community Modeling Program, Johns Hopkins University, USA, alexey.voinov@uvm.edu

Abstract: This paper presents an agent-based model of a land market (ALMA-C) to simulate the emergence of land prices and urban land patterns from bottom-up. Our model mimics individual decisions to buy and to sell land depending on economic, sociological and political factors as well as on the characteristics of the spatial environment. To this we add ecological and environmental considerations and focus on the question of how individual land use decisions can be affected to reduce the pressure on the coastal zone ecosystem functions. A series of model experiments helps visualize and explore how economic incentives at a land market can influence the spatial distribution of activities and land prices in a coastal zone. We demonstrate that economic incentives do affect urban form and pattern, land prices and welfare measures. However, they may not always be sufficient to reduce the pressure on coastal zone ecosystems.

Keywords: agent-based modelling; coastal zone ecosystems; land market mechanisms.

1. INTRODUCTION

Coastal zones (CZ) are important from both ecological and socio-economic points of view (Martínez, Intralawan et al. 2007). These are one of the most productive areas on our planet that provide many ecosystem services such as erosion control and sediment retention, habitat for species, food production, recreation and others (Costanza, d'Arge et al. 1997). Like other ecotones these areas are especially rich in biodiversity and have one of the highest values for ecosystem services per hectare of area. CZs require a delicate balance between human-dominated systems and ecosystem functions provided by interactions of land and sea. This makes protection of CZs in their pristine form an important component of environmental management. Unfortunately, CZs are also very lucrative area for development. CZs are also one of the most densely populated areas housing two thirds of world's population (Costanza,

Andrade et al. 1999). Particularly in the Netherlands 70% of the Gross National Product today is generated in the CZ (Veraart, Brinkman et al. 2007). Pressure on CZ induced by economic activities causes the disruption of coastal ecosystem functions including modification of a shoreline, reduction of habitat's carrying capacity and diminishing recreational value of coasts (Costanza, Andrade et al. 1999; Martínez, Intralawan et al. 2007). Waterfront properties are known to be several times more expensive than similar properties inland. People are willing to pay high prices for water view and water access. These areas have been also historically developed due to proximity to marine and river transportation. Further developments occur in the proximity to historic cities causing even more buildings constructed in areas vulnerable to flood or erosion. According to IPCC the damage from natural disasters has rapidly increased over the past decades due to the growth of capital in flood-prone areas (Nicholls, Wong et al. 2007). Thus, in addition to contributing to the deterioration of coastal ecosystem functions economic activities located in CZ are subject to flood risks. The potential damage from flooding/erosion depends on the economic value of land and concentration of economic activities or residential areas.

If we are to protect CZ ecosystem services we need to find ways to create disincentives for people allocation near the coastline. Using economic incentives to achieve environmental goals can be much more efficient than traditional command and control regulation, if the incentives can be put in place and enforced at relatively low cost (Costanza, Andrade et al. 1999). *What can be the economic incentives and market mechanisms that would reallocate population out of the CZs?*

In this paper we present a simple agent-based model that integrates heterogeneous landscape, explicit individual location choices via land market to explore the influence of economic incentives on the aggregated spatial pattern of the urban area. Since the area occupied by urban developments actually takes over the land that provides ecosystem functions in CZ, this aggregated pattern indirectly shows how much pressure is put on the CZ ecosystems.

2. THE MODEL

Our *Agent-based Land MArket for Coast* model (*ALMA-C*) simulates the emergence of urban land patterns and land prices as a result of micro-scale interactions between buyers and sellers of land with application to a coastal city. The main agents in the model are traders, spatial goods (i.e. land lots) and a market (see Figure 1).

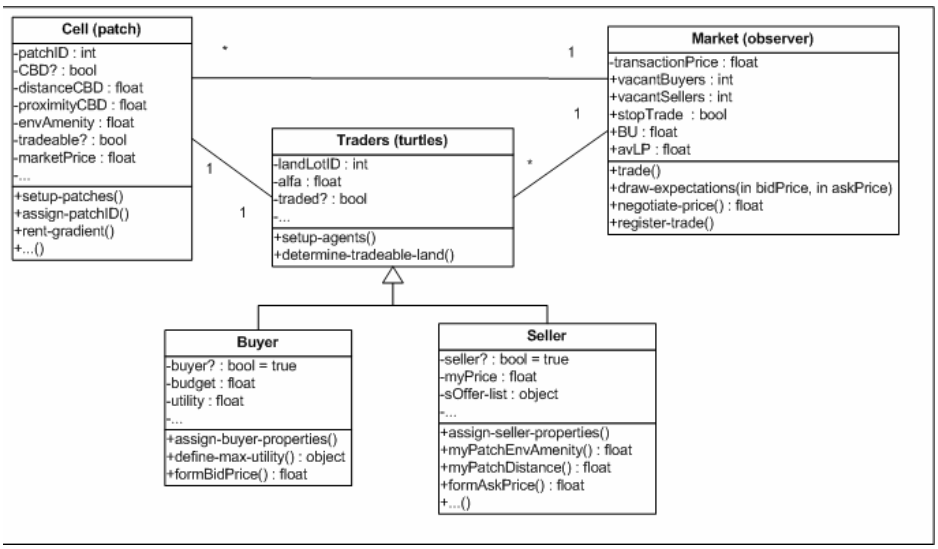


Figure 1. UML class diagram of the ALMA metamodel

The core of the model is presented by a land market (Filatova, Parker et al. 2007; Filatova, Parker et al. under submission). ALMA-C borrows much from the analytical monocentric urban model (Alonso 1964) and its application to a city with green amenities (Wu and Plantinga 2003). In line with the assumptions of the analytical model the ALMA-C model assumes that each spatial good is differentiated by distance (*D*) from the central business district (CBD) (or its inverse measure – proximity $P=D_{max}+1-D$) and level of environmental amenities (*A*) (estimated as a normalized distance to the coast). Other attributes of land can be easily added if needed. Buyers (i.e., households) search for a location that maximizes their utility $U = A^{\alpha} \cdot P^{\beta}$ (ά and β are individual preferences for green amenities and proximity correspondingly) and is affordable to their disposable budget for housing net of transport costs (*Y*). The rationality of agents is bounded by the fact that they do not search for the maximum throughout the whole landscape but rather search for the local maximum among N randomly chosen cells. We impose this assumption since the search for a house in reality is very costly (time-wise and money-wise), meaning that a global optimum is not likely to be located in real-world housing markets. After defining the spatial good that gives maximum utility a buyer forms the bid price. A bid price is a function of utility (*U*), individual income (*Y*) and prices of all other goods (influence of which is expressed by a constant *b*): $P_{bid} = \frac{Y \cdot U^n}{b^n + U^n}$. The justification and properties of this demand function are discussed in details in (Filatova, Parker et al. under submission). Sellers form their bid prices also depending on their utility, bit their ask price do not go below the price of an open space area (P_{nat}).

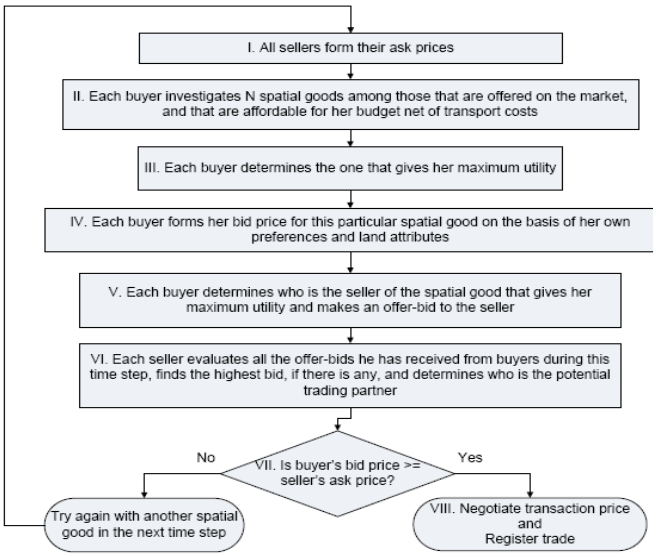


Figure 2. Conceptual algorithm of trade

Having found the optimal location, buyers submit their offer-bids to the sellers. Sellers choose the highest bid-offer and if it is above their ask price, then the transactions take place. If not then both buyer and seller participate in the land market in the next times step. The final transaction price is an arithmetic average of the ask price and the highest bid price. Figure 2 shows the logic of the trading mechanism, i.e. one time step in the model (more about event sequencing in (Filatova, Parker et al. under submission)).

Let us suppose that in attempt to protect the coastal ecosystem services we are introducing a tax to make coastal properties more expensive and less affordable. Suppose we calculate the amount of tax to be paid as $T = tr \cdot P_{nat} \cdot P_{amen}$, where *tr* is a tax rate in % and P_{amen} is proximity to the coast.

3. SIMULATION EXPERIMENTS

Running the model we can produce spatially explicit land price gradients and land patterns as simulation results. We are mainly interested in how the introduction of a land tax, aimed to protect CZ ecosystems, affects economic indicators and the spatial morphology of the city. In addition to graphical representations, we also present a set of metrics to analyze micro and

macro economic and spatial outcomes (listed in Table 1 below). Our major metric of success with respect to the ecosystem services preservation is the number of undeveloped cells within the prime coastal zone, which is defined as the stripe of land covered by 7 cells wide along the coast (all the zone that lies to the left of the CBD). We will call this zone *CZ buffer*.

All model experiments presented in this paper were performed on a 31x51 grid of cells with part of the landscape (5x cells) representing sea and a coastline. The CBD was set at the cell with coordinates (-4;0). There were 2652 buyers and sellers participating at the land market. The price of the natural area (P_{nat}) was set equal to 200. We run the model with agents having homogenous preferences for green amenities, i.e. for the view on the coastline, $\alpha = 0.7$. Although the code of the ALMA-C model allows simulation of a land market with heterogeneous agents, we did not use this possibility here to keep clarity. Transport costs per unit of distance were equal to 20 units, the constant b in (1) was equal to 70.

To understand the influence of such an economic incentive as tax on the aggregated outcomes we ran a series of 13 experiments. We changed only two parameters in the experiments: tax rate and individual income, which both affect households' willingness to pay for the house in the coastal city. Each experiment is associated with the value of tax and income, specifically experiment "3-900" means that we ran the model with a land tax equal to 3% and individual income equal to 900 units. We show the outcomes of 7 representative experiments in Table 1 and discuss the results of the others below.

Table 1. Economic and spatial metric outcomes of the ALMA-C experiments

Parameter	0-800	5-800	5- ND 800/20	0-900	5-900	0-1000	5-1000
Individual utility: Mean	73.17	71.86	72.68	70.26	70.59	67.79	68.15
St. dev.	11.13	10.82	11.41	12.14	12.18	12.98	12.87
Aggregate utility	36587.71	28672.4	34305.6	43494	40450.55	49082.77	48254.45
Urban transaction price: Mean	312.67	312.07	318.64	336.61	337.12	361.77	358.07
St. dev.	75.84	73.92	79.49	93.15	90.61	110.15	107.86
Total property value	156335.9	124514.1	147965.8	208363.2	193169.9	261924.4	253515.4
City size (urban population)	500	399	472	619	573	724	708
Undeveloped land in CZ buffer	86	175	117	42	74	6	12
Distance at which city border stops	20.88	18.11	19.23	23.77	22.36	25.71	25.7

We begin with the experiment that replicates the monocentric urban model with amenities such as coastline. The main difference between the simulation experiment and the analytical model is that the centralized equilibrium land price determination mechanism is replaced by a series of bilateral trades distributed in space and time. Individual households endowed with income equal to 800 try to buy a house that maximises their utility in the coastal city. There is no tax for land introduced in this city. The urban land price gradient is presented in Figure 3 and quantitative measures can be found in Table 1, column "0-800". The intensity of grey colour symbolizes the value of land: the darker the colour, the higher the land price. The urban land prices are the highest in the CBD. As in the benchmark case of a theoretical monocentric urban model, the land price gradient is decreasing with distance from the CBD. But because of the presence of the coast, which serves as an attractor for the households and increases value of the land, urban land laying to the left of the CBD has higher prices than the area to the right. The city expansion stops at the location when bid price of a buyer falls below the price of open space. The white area in Figure 3 shows the beginning of open space and symbolizes the city border.

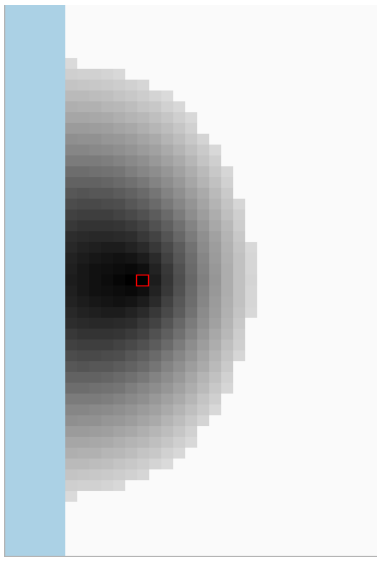


Figure 3. Land price gradients (households income = 800, no tax for land introduced). Red counter shows the CBD.

With the increase of individual income the city significantly expands and CZ buffer representing the amount of open space along the coastline decreases (compare experiments "0-800", "0-900" and "0-1000" in Table 1). So does an average individual utility from location showing that people enjoy less densely populated city, which provides more open space along the coast and less commuting in terms of time and money. In spite of decrease in average utility land prices still grow because the purchasing capacity, i.e. income, has increased.

We proceed with demonstration of changes in the location behaviour of households changes if an environmental tax on land is introduced. The idea is that now agents have to pay extra money for location closer to the coast, assuming all other factors are the same. If this happens then demand for these locations goes down, since it becomes harder to find buyers that could afford to buy there. The market mechanisms start to work. Since our model is not only demand-driven, meaning that land prices are determined via bilateral trades with sellers, the bid price that buyers offer is below the ask price (which also depends on the spatial characteristics of the land lot, but it could not be lower than the land price of the natural area). In this case the transaction simply does not occur and land lot is not occupied by an urban land use.

Let us set the tax equal to 5% of the land price of the natural area. The land price gradient is presented in Figure 4. First of all, the spatial form of the city has changed. The area closer to the coast became less attractive for households to pay more to cover the threshold of the P_{nat}. The households in this experiment basically "voted with their feet" in favour of living in

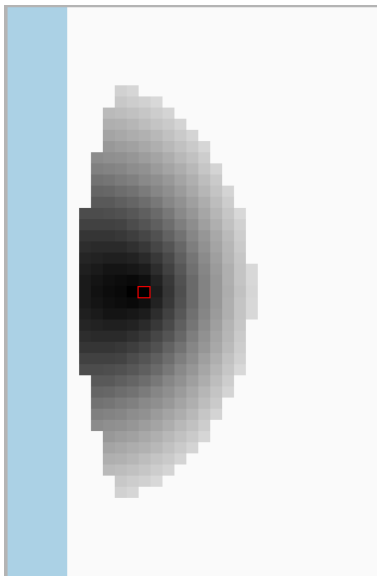


Figure 4. Land price gradients (households income = 800, and land tax = 5%). Red counter shows the CBD.

another town because living on the coast now bears additional costs. As a result a lot of land in the "primary" CZ was left unsold and remained in natural conditions. A positive sign is that the CZ buffer has significantly expanded: it became 103% larger than in the city without the land tax aimed to preserve CZ (compare columns "0-800" and "5-800" in Table 1).

The results of experiment "5-800" indicates that in the absence of real markets for the environmental services provided by the CZ economic incentives, such as taxes imposed by the government, may help to influence the land market to achieve more environmentally friendly land use patterns. However, the outcome can change quite dramatically if we open up the market and raise the limit on individual budgets. Effectively this means that we are inviting buyers that are more affluent to come from elsewhere and enter the market.

We repeat the second experiment but assuming that richer households enter this urban land market (we set tax equal to 5% and incomes equal to 1000 units). The results are presented in Table 1 (see "5-1000") and in Figure 5. Thus, in the presence of more affluent households the city not only expands over the borders of the city in the experiment "5-800" (Figure 4) but overflows the boundaries of the city without land tax (experiment "0-800", Figure 3). High-income agents can afford to pay additional costs in terms of the land tax and, thus, settle along the coastline. This fact can often be observed in reality: the wealthiest individuals buy houses with a view on the coast. As a result the CZ buffer has shrunk to 93% of what it was when agents' income was 800 (see "5-800").

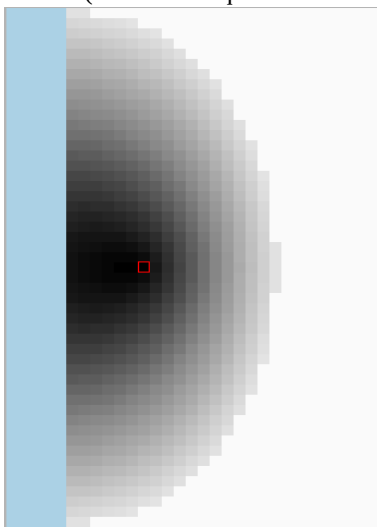


Figure 5. Land price gradients (households income = 1000, and land tax = 5%). Red counter shows the CBD.

On the other hand, in reality not all people have equally high incomes. There is a certain distribution of incomes in the city. Figure 6 shows the land price gradients in the case when we run the model for the population of agents with heterogeneous incomes and land tax rate of 5%. Individual incomes follows normal distribution with mean 800 and standard deviation of 20 (see experiment 5-ND 800/20 in Table 1). There are some households with income higher than average of 800 (as in experiment "5-800" and "0-800"), which can afford to buy houses in the CZ buffer and pay a land tax. However, market incentives still work and part of CZ buffer became vacant (117 instead of 86 in the case without land tax).

In fact, the last three experiments ("5-800", "5-1000" and "5-ND 800/20") showed that economic incentives work only if agents have fixed income. If taxes are introduced but the higher-income households enter the land market then the effect of the tax introduction is eliminated. The introduction of the tax and increase in household incomes drive spatial and land price dynamics in opposite directions. In this case we were interested to investigate the relationship between these two factors. We performed a series of additional experiments with tax rates changing from 0% to 7% and income changing from 800 to 1000. The results in terms of the CZ buffer area remained undeveloped are presented in Table 2.

Table 2. Number of undeveloped cells in the CZ buffer under different tax rates and income values

	Individual income		
Tax rate	800	900	1000
0%	86	42	6
3%	92	46	10
5%	175	74	12
7%	314	247	158

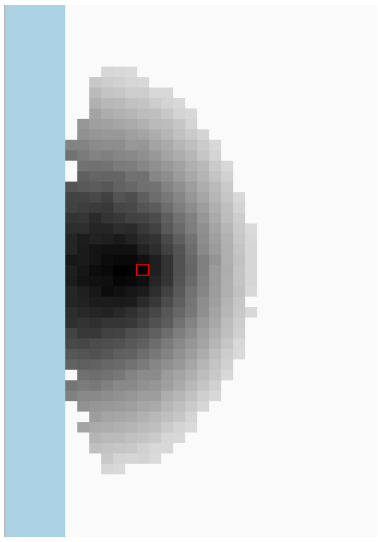


Figure 6. Land price gradients (households incomes follow normal ditribution and land tax = 5%). Red counter shows the CBD.

Visually the relationship is presented in Figure 7. One can see that in the zone of high income the increase in tax gives very small effect on the CZ buffer

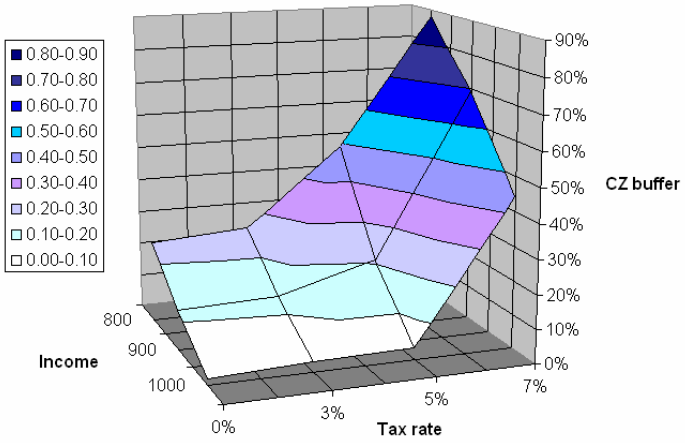


Figure 7. The percentage of CZ buffer not occupied by urban land use

4. DISCUSSION AND CONCLUSIONS

The paper aims at exploring the effects of economic incentives on CZ ecosystem services. At this stage we are interested in general qualitative trends of the ecological – economic system. We used artificial data for both the agents' behaviour and the spatial environment. However, the structural validation of the model was performed (Filatova et al. 2007). The main difficulty of applying ABMs to the real world cases and model validation is in acquiring data about people's preferences and perceptions. As a part of a more general survey on risk of flood perception carried out in March 2008, we created a questionnaire about individual preferences for locations, including preferences for coastal environmental amenities and individual risk perceptions. As the next stage of model development we plan to integrate survey data (when statistical analysis is available) into the agents' behaviour in our ABM.

The model has shown that economic incentives may help in managing the urban development in CZ but may not necessarily produce the desired outcome. For example, the environmental

tax instead of protecting the CZ by driving population away from the coast results in attracting more affluent residents – something clearly observed in reality. The CZ becomes populated by MacMansions, which still may have a positive effect on the environment assuring less residential density. The problem then is that environmental quality can potentially lead to social injustice and unrest. The environmental tax drives people away from their native habitats, which then get developed and shaped for the more affluent residents. The ABM helps quantify some of these impacts and analyze trends and spatial patterns as emerging outcomes of individuals interacting at the land market.

Different residential patterns may occur if we bring risk perception into consideration. While CZs and waterfront are attractive residential factors, they become also associated with higher risks of natural disaster (hurricanes, floods, etc.). These are likely to be more frequent with global warming and climate change. As a result, chances are that in addition to environmental taxes, insurance policies are likely to kick in, when insurance companies will deny insurance for properties located too close to the coast. In this case risk perception may become another major component in addition to affluence that will govern the patterns of allocation for residents. Incorporating these processes into our model is the next step in our research.

With a lack of markets for ecosystem services there should be other indirect economic mechanisms for CZ management, which would help account for these services in the decision making process. Developing tools for visualization and scenario analysis, such as the ALMA-C model, is an important prerequisite of environmentally sound land use planning.

ACKNOWLEDGEMENTS: Funding from NWO-ALW (LOICZ-NL) project 014.27.012 is gratefully acknowledged. The authors are also grateful to the staff of the Center for Social Complexity at George Mason University and to D.C. Parker in particular for their valuable input and advice.

REFERENCES

Alonso, W. (1964). *Location and Land Use*. Cambridge, MA: Harvard University Press.

Costanza, R., F. Andrade, et al. (1999). "Ecological Economics and Sustainable Governance of the Oceans." *Ecological Economics* 31: 171–187.

Costanza, R., R. d'Arge, et al. (1997). "The Value of the World's Ecosystem Services and Natural Capital." *Nature* 387: 253-260.

Filatova, T., D. C. Parker, et al. (2007). Agent-based land markets: heterogeneous agents, land prices and urban land use change. Proceedings of the 4th Conference of the European Social Simulation Association (ESSA'07), Toulouse, France.

Filatova, T., D. C. Parker, et al. (under submission). "Agent-Based Urban Land Markets: Agent's Pricing Behavior, Land Prices and Urban Land Use Change." *Journal of Artificial Societies and Social Simulation.*

Martínez, M. L., A. Intralawan, et al. (2007). "The Coasts of Our World: Ecological, Economic and Social Importance." *Ecological Economics* 63: 254-272.

Nicholls, R. J., P. P. Wong, et al. (2007). Coastal systems and low-lying areas. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the *Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. M. L. Parry, O. F. Canziani, J. P. Palutikof, v. d. L. P.J. and H. C.E., Cambridge University Press, Cambridge, UK: 315-356.

Veraart, J. A., S. Brinkman, et al., Eds. (2007). Climate Changes Spatial Planning. Introduction to the Dutch National Research Programme. CcSP-Brochure. Wageningen, the Netherlands.

Wu, J. J. and A. J. Plantinga (2003). "The influence of public open space on urban spatial structure." *Journal of Environmental Economics and Management* 46(2): 288-309.

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 activity based cellular automaton model to simulate land use changes

Jasper van Vliet[a] **and Hedwig van Delden**[a]

[a] *Research Institute for Knowledge Systems bv, Witmakersstraat 10, 6211 JB Maastricht, The Netherlands (jvanvliet@riks.nl, hvdelden@riks.nl)*

Abstract: This paper proposes a spatial explicit model for land use dynamics, based on activities, which can represent for example population or jobs. In land use models based on constrained cellular automata (CA), the total area per land use is defined exogenously, while the model computes the allocation based on transition rules. The activity based model is a CA model, but it is constrained by activities instead of areas. Each time step, activities are distributed over cells based on transition rules. The CA transition rules comprise the effects of the activity in the neighbourhood, the land use of a cell, externalities and a stochastic perturbation term. Land use then is computed based on the activity distribution. Hence land use and activities are mutually dependent and each cell has two values: a land use state and an amount of activity. The activity based model is applied to simulate population dynamics and land use changes in Spain. Simulation results show that the model can produce realistic land use dynamics. Moreover, we argue that the inclusion of activities closer resembles the process of real world land use dynamics and offers good opportunities for integrated modelling.

Keywords: Activity based modelling; cellular automata; land use change; population dynamics.

1 A WALK IN THE REAL WORLD

Imagine you are in the centre of a large urban area and you take a walk. Your starting point might be the central business district and you are surrounded by high rise office buildings, but the further you walk the more residential buildings you'll find. Eventually you will pass some suburbs to end in the agricultural area outside the town. Parts of the city you passed through each have some distinct features; some are densely populated while others are not, some are mainly for commercial use, while others are primarily residential but certainly not entirely. Although land uses can be classified, most areas actually have a mixed land use. When you try to find the boundary between commercial zones and residential areas, between the city and its surrounding, you will find that in reality these are not always clear. In fact geographical classifications as well as their spatial boundaries are often fuzzy (Fisher, 2000).

These fuzzy boundaries and mixed land uses are difficult to represent in computer models that simulate the real world. For example land use classes are usually discrete and dynamics are often Boolean processes. In the model proposed in this study, we aim to overcome some of these problems. It allows mixed land uses by allocating population and jobs separately from, and on top of, the predominant land use in a location. Moreover it simulates dynamics in population and jobs in an incremental way. Hence, land use changes are not sudden events, but the result of a more gradual process.

In the next part first a short overview is given of some methods that our model builds on. Section 2 describes in depth the computational scheme that is applied. Section 3 shows the

case study for which the model is applied and discusses some preliminary results to end with conclusions and directions for further research in section 4.

1.1 On agents, cells and population density

Over the last decades several methods are proposed and applied to simulate spatial explicit land use change. Overviews of different methods are among others available in Veldkamp and Lambin (2001) and Parker et al. (2003). Instead of giving another overview we would like to place the proposed model in the spectrum of modelling methods. For this we briefly consider three methods specifically: multi agent systems (MAS), cellular automata (CA) models and we will shortly point at some characteristics of economic land use models..
Both MAS and CA are dynamic approaches that can simulate land use changes in a way that approaches human decision making (White and Engelen, 2000). Moreover, both allow for heterogeneity among agents or land uses. These aspects are important since they differ from several other models on these points. In this discussion, agents in MAS are actors that can act and move independently over space. Hence CA are not considered MAS since there agents are the cells themselves, and their location is fixed. The advantage of MAS is that they can represent the behaviour of agents in a very straightforward way, since agents can interact directly with each other and with the environment. It is precisely these local interactions between agents and differences among them that generate the patterns observed on a global scale. However, since the agents are the basic unit of computation, MAS are computationally demanding. This is illustrated by an overview presented in Parker et al. (2003) where several case study applications are summarized. The applications that use individuals or households are all on a relative small scale using a limited number of agents.
Cellular automata, although sometimes considered agent based as well, differ from MAS in that sense that the basic unit for computation is a cell, not an agent. They have in common that both methods simulate changes from the bottom up, since the eventual global land use pattern is a result from interactions at the local scale. Together cells make up the lattice on which the CA exists, which makes them inherently spatial and therefore very suitable for the simulation of land use dynamics. Since cell sizes can be adjusted according to the scope of the simulation, models can keep a computational efficiency. Hence, CA can be applied to simulate land use changes on larger scales, from urban systems to regions or countries. This advantage comes at the cost of detail. Individual actors are not considered. Instead cells have a state, which generally represents the predominant land use, but the number of possible land use classes is limited. Moreover, cells have only one land use where in reality mixed land use is the rule rather than the exemption.
From the side of regional economics, land use models are usually considered in a non spatial way. Starting from Alonso (1964) these models describe population, utilities, jobs, or prizes as a function of the distance to the city centre. In this they differ from models like MAS and CA as described above (Irwin and Geoghegan, 2001). A quick look at any city tells that this simplification hardly holds, since urban areas are usually not symmetric but show local differences instead. This local variation can continue to grow and eventually results in polycentric developments (Wu, 1998). Moreover, the eventual results of these economic models are equilibrium situations and the way towards this equilibrium is usually not considered. Hence this approach neglects the incremental steps that play a role in the evolution of a city and the fact that most urban systems are still evolving.

In this paper we propose an activity based CA model to simulate land use dynamics. For this we introduce the notion of “activities” which is used to indicate the general concept of the model, in the case study described below it will be given a specific meaning. Activity is used here as a general term to denote people or jobs or any other feature that can be quantified and located. However, activities are considered a cell property, not entities by themselves. Hence a cell has a land use state, and a numerical value for each activity. Therefore, the cell remains the basic unit of computation, and it is possible to build on the CA computational framework. On the other hand, we do not consider his model a MAS for

the same reason; activities can represent individuals, but we do not simulate decisions from individual agents.
This activity based approach will generate more spatial information than original CA land use models while retaining their computational simplicity. Also it allows to explicitly model interactions among people and jobs, like externalities, without losing the explicit spatial extent as is the case in models from the economic background. This activity based approach overcomes some limitations of conventional CA models as mentioned above. First, activities can exist in locations that do not have the associated land use, like inhabitants living in agricultural areas. Therefore additional spatial information is available from model results and mixed and multifunctional land uses are possible. Second, the focus of applications needs not only to be on the predominant land use, but rather on the activities itself which allows for a higher level of detail. Moreover, it offers more direct coupling between the socio-economic and the bio-physical subsystems since impacts that origin from activities can be quantified more precisely now. The latter is an asset for integrated models that aim to simulate interactions between both sub-systems more elaborately.

2 THE ACTIVITY BASED CA MODEL

The activity based cellular automaton as proposed in this paper builds on the Metronamica constrained CA models as developed by White and Engelen (1993; 2000). Their model exists on a lattice of regular squares, which represents a land use map. Each cell can have one of a limited number of cell states, which represent the predominant land use. Each cell has a neighbourhood and there is a set of transition rules to compute state transitions based on this neighbourhood. All cell states are updated simultaneously within each time step. The “constrained” in constrained cellular automata refers to the notion that the demand per cell state is determined external to the CA (White, Engelen and Uljee, 1997). In land use terms, the total number of cells of, for example, residential land use in a certain time step is determined exogenously while the CA allocates these cells on the map. Hence an increase in population over time needs to be converted to a number of cells for the associated residential land use per time step.
To introduce activities to this model, we made a few additions to the constrained CA framework. By introducing activities, a cell has no longer one discrete cell state. Instead it has a land use state, as well as a quantity for each type of activity. Activities can represent for example people, jobs, or maybe more abstract terms like economic value. Moreover, more than one type of activity can be used at the same time. For example, when both jobs and population are modelled, a cell could have agricultural land use, contain 12 inhabitants (activity “population”) and have 4 people working in it (activity “jobs”). Instead of a demand for a number of cells per land use, the model is constrained by an amount of activity that needs to be allocated.
Each activity has one associated land use. For example, population and jobs are directly associated with residential and commercial land use respectively. For computation of activity dynamics and land use changes we assume that activities and land use are mutually influential. This is implemented as a two step process and each step is executed for all cells simultaneously. First activities are redistributed over the map based on the land use and activity distribution from the previous time step. When more than one activity is considered, this is done for all activities separately. Then, land uses are assigned based on the new distribution of activities. Now all cells have an activity, but only those cells whose activity exceeds a threshold value will get the associated land use.
However, not all land uses can be associated with activities, at least not in the sense that this activity can represent the eventual demand in terms of cells. For example the land use agriculture can be expressed by a number of jobs in agriculture, but that cannot directly be translated to an amount of hectares that need to be allocated on the map. Consider for example the difference between intensive and extensive agriculture. Therefore we need area constrained land uses as well. In fact the model uses four types of land use classes: activity constrained land uses, area constrained land uses, vacant states and features.

Activity constrained land uses are land uses that are directly associated with an activity as described above. Area constrained land uses do not have activities. Instead they have an area defined per time step that needs to be allocated. Vacant states are land uses that can only change as a result of other changes in the model. A typical example of a vacant land use is natural vegetation. Features finally are those land uses that are supposed to remain constant over time, like water bodies.

2.1 Distribution of activities

To allocate activity, the potential for that activity for each cell is computed according to the equation:

$$TP_{x,i} = c_{l(i),x} \cdot Z_{x,i} \cdot S_{x,i} \cdot A_{x,i} \cdot f\left(N_{x,i} + E_{x,i} + \varepsilon\right)$$

Where:

$TP_{x,i}$	is the total potential for activity *x* in cell *i*,
$c_{l(i),x}$	is the compatibility coefficient for land use *l* in cell *i* and function *x*,
$Z_{x,i}$	is the zoning for activity *x* in cell *i*,
$S_{x,i}$	is the suitability for activity *x* in cell *i*,
$A_{x,i}$	is the accessibility for activity *x* in cell *i*,
$f(p)$	is a transformation function to avoid negative potentials: $f(p) = \log_2\left(1 + 2^p\right)$,
$N_{x,i}$	is the neighbourhood effect for activity *x* in cell *i*,
$E_{x,i}$	is the externalities effect for activity *x* in cell *i*,
ε	is a stochastic variable drawn from a Cauchy (0, α) distribution, where α is an adjustable scale parameter. This variable is drawn independently for each function and each cell.

The compatibility coefficient is an adjustable parameter, which results from calibration. Suitability is a characteristic that is defined per cell a priori, as is zoning, but the zoning status can change over time. Both result from an overlay of GIS base layers. Accessibility is also a cell characteristic, which is computed from the distance from a cell to the nearest element of the infrastructure network as well as the importance of this element for the function at stake (exits of highways are often important for industry, while office building have a tendency to be located near train stations). The neighbourhood effect is a function of the current vector of activities and current land use in each cell of the neighbourhood of a cell. It is computed as the sum of the effects of all cells *j* at each distance *d* in the neighbourhood. This includes the current activities and land use of cell *i* itself:

$$N_{x,i} = \sum_{d} \sum_{j \in J_{d,i}} \sum_{f} w_{f,x}(d) \cdot X_{f,j}$$

Where:

$N_{x,i}$	is the neighbourhood effect for activity *x* in cell *i*,
$J_{d,i}$	is the set of cells at distance d from cell *i*

$w_{f,x}(d)$ is the weight function representing the attraction or repulsion from activity or land use *f* on activity *x* at a distance *d*,

$X_{f,j}$ is the level of activity for function *f* in cell *j*. For those land use functions that do not have an activity associated with them, the value is 1 when that land use is present and 0 otherwise.

Externalities represent the, mostly negative, effects of agglomeration. It is computed as a function of the amount of activity already present in an area:

$$E_{x,i} = \gamma_1 \cdot \left(N_{x,i} - \gamma_2 \right)^{\gamma_3}$$

Here, γ_1, γ_2 and γ_3 are parameters that need to be calibrated. Typically $-1 << \gamma_1 < 0$ and $\gamma 3 > 1$ to make sure that externalities are small initially and grow more than proportional with the present activity.

Once the total potential for each activity in each cell is known, all activity is (re-) distributed proportionally. When there is more than one activity type, allocation is independent for each type of activity. Hence activities only influence each other through the neighbourhood effect, which is based on the distribution in the previous time step:

$$X_{x,i} = \frac{TP_{x,i}}{\sum_j TP_{x,j}}$$

where

$X_{x,i}$ is the level of activity for activity type *x* in cell *i*,

$TP_{x,i}$ is the total potential for activity *x* in cell *i*,

$\sum_j TP_{x,j}$ is the sum of the potentials for activity *x* in all cells *j*, including *i*.

We assume that the basic unit of activity is 1. Therefore activities in all cells are rounded to whole numbers. This concurs with the assumption that people have to make a choice where to live and where to work. If two locations are equally attractive, one has to be selected in the end and based on this decision future development will take place.

2.2 Allocation of land uses

Land use is now assigned to cells per type of land use class. First features are allocated, then activity constrained land uses, then area constrained land uses, and finally vacant land uses. Since features cannot change over time, they are by definition allocated on the location they already hold. For the other land uses the order of assignment corresponds with the influence associated with them. In reality agents that claim residential or commercial land uses usually have more power than those for agricultural or natural land use. Activity based land uses represent these more powerful types of land uses, while the agriculture is represented by an area constrained class. Vacant land uses are those land uses that hardly have any economic power, like abandoned land or semi natural vegetation.
The basis for the land use assignment is the distribution of activities as well as a reference value for each activity. This reference value is taken as a percentile of the activities of all cells that have the associated land use in the initial land use map. Each time step, the activity per cell is compared relative to this reference value. A cell gets the land use function for which it has the highest relative activity as long as this is greater than one. For function land uses that are not constrained by an amount of activity but by a number of cells instead, the total potential for each cell is computed similar to the total potential for

activity. However land uses are not assigned proportional to this total potential, but instead to cells with the highest potential until the externally defined demands are met. Cells that already have a feature or activity constrained land use are excluded from this assignment. Finally, all cells that do not have a land use from either of the constrained land use types become vacant. Since the number of cells that is occupied by activity constrained land uses is not known a priori it is a computational necessity for the model to have at least one land use type that can occupy vacant cells, to be able to assign a land use to all cells.

3 A CASE STUDY ON SPAIN

The activity based CA model as described above was applied to simulate land use dynamics for the country of Spain. In this case study the model was defined on a regular grid of 912 by 1076 cells of 1 km2 each, and land use change was computed for yearly time steps starting from 2000. Population, with residential areas as the associated land use function was modelled as an activity constrained class. Forest, commercial and industrial, agriculture and recreational land uses are area constrained classes. Natural vegetation is a vacant land use and water, wetlands, airports and mining areas are feature land uses. Although some of the latter land uses can actually change over time, their dynamics are much less. Besides it is assumed that their behaviour depends on other influences than local interactions. Still they do influence other land use changes considerably.
A limitation to apply the activity based CA model on a real world case is its data requirements. For these applications we used land use data from the Corine Land Use database for the year 2000. Population data was not measured directly but derived from the Corine land use database (Gallego and Peedel, 2001). Moreover, since we could not find spatial explicit data on jobs, commercial land was not modelled using activities.
Eventually, the activity based land use change model will not be used as a stand alone model, but as a part of an integrated assessment model that is developed within the DeSurvey project. This integrated model aims to explore future land use scenarios in the context of land degradation and desertification. Since population is modelled as an activity, land abandonment can be simulated with this model. Moreover, information on the amount of population allows for more direct coupling between the socio-economic system and the physical effects of land degradation and it enables exploration of a wider range of scenarios.

3.1 Preliminary results and discussion

The aim of this case study is to assess whether the proposed model realistically simulates activity dynamics and land use change. Hence model results were assessed on realistic patterns rather than cell-to-cell accuracy of land use changes. A first way to assess model results is by visual interpretation. Although it is preferable to measure results objectively rather than interpret them subjectively, visual interpretation can still gives an important clue about the behaviour of the model. Visual interpretation of the population distribution indicates that the model retains local differences in density. This is an important aspect since it is exactly that what distinguishes a spatial explicit model from economic models. This is partly due to local characteristics that are incorporated in the model, and partly due to the internal feedback mechanisms in the model.
For clusters of residential land use, urban areas, scaling laws are found to apply (Roehner, 1995). To assess the outcomes of the model we choose one such measure to assess whether a realistic distribution for residential land use was obtained: the cluster size frequency distribution (White, 2006). Although specific urban areas can change their position on the ranking, the distribution over larger areas remains mostly constant over time. The cluster size frequency distribution confirms the visual interpretation that the model can produce realistic dynamics. Both large urban areas are growing and new clusters appear. Figure 1 shows the results of the cluster size frequency distribution, where cluster sizes are measured in cells. Only the largest clusters seem to grow disproportionally. This is due to

the input data. Cells around the major urban areas that are not residential initially already have a high population. Hence, in the first time step these cells change their land use to residential.

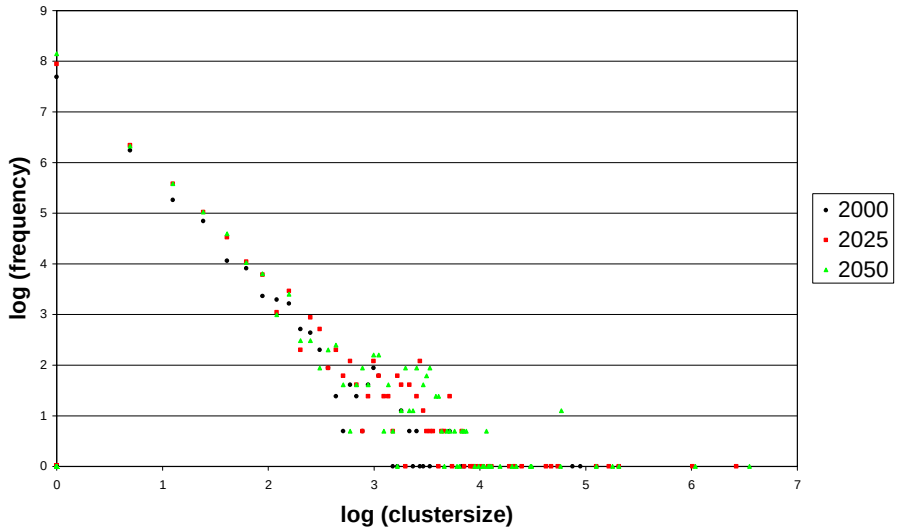


Figure 1: Cluster size frequency distribution for Spain. Blue points represent original data, while green and red are simulation results.

4 CONCLUSION AND DIRECTIONS FOR FURTHER RESEARCH

We believe that the important interactions among activities and from activities to land uses can be simulated using activities. In existing spatial explicit land use models the interface between a land use model and other models was mainly a land use map on the output side, and a multitude of information on the input side (see for example van Delden et al., 2007). However, many relations are not so much related or proportional to the predominant land use, but rather to the amount of activities on this location. The introduction of additional layers of information offers new opportunities for integrated modelling and more elaborate studies of land use dynamics. As an example Luck (2007) gives an overview of the relation between population density and its effect on biodiversity. He concludes that biodiversity and population density are significantly correlated and points at the need to focus on anthropogenic drivers of environmental changes. Still this case study only gives a first indication and the concept should be tested more extensively on other areas.
A major issue for the proposed model is data availability. Although synthetic applications can offer first insights in the mechanisms for land use change, real world data is required for meaningful studies. The model requires a land use map and a distribution for each activity. For proper calibration this data is even required for two points in time and independent validation would even require three such data sets. Due to developments in remote sensing, land use data is currently widely available. For population data this is certainly not the case, just as for jobs or other possible activities. Currently such data is at best available at more aggregate levels such as NUTS regions. The population data that was used in this research was derived from the land use map and is therefore biased. Alternatively, some remote sensing techniques are developed to estimate population from satellite imagery (Harvey, 2002), but this does not generate cellular values for population density yet. Finally there is the question how to assess the quality of the model results. Measures for land use maps as such are already topic of discussion (Hagen, 2003). Since the activity based model generates additional spatial results, more sophisticated measures are needed.

Acknowledgements

The activity based cellular automata model was developed within in the DeSurvey project, funded by the European Commission 6th Framework Program, Global Change & Ecosystems, Integrated Project Contract No. 003950.

References

Alonso, W. (1964) Location and land use. Harvard university press, Cambridge.

Anas, A., Arnott, R. & Small, K.A. (1998). *Urban spatial Structure.* Journal of Economic Literature 36. 1426 – 1464.

Delden, H. van, Luja, P., & Engelen, G. (2007) Integration of multi-scale dynamic spatial models of socio-economic and physical processes for river basin management. Environmental modeling & software 22, 223-238.

Fisher, P. (2000). *Sorites paradox and vague geographies.* Fuzzy sets and systems 113, 7 – 18.

Gallego J. & Peedell, S. (2001). *Using CORINE land cover to map population density.* In: European Environment Agency (2001). *Towards agri-environmental indicators; Integrating statistical and administrative data with land cover information.* EEA, Copenhagen, pp 92 - 104.

Hagen, A. (2003) *Fuzzy set approach to assessing similarity of categorical maps.* International journal of geographic information science 17 (3) 235 – 249.

Harvey, J.T. (2002) *Estimating census district populations from satellite imagery: some approaches and limitations.* International journal of remote sensing 23 (10), 2071 – 2095.

Irwin, E.G., & Geoghegan, J. (2001). *Theory, data, methods: developing spatial explicit economic models of land use change.* Agriculture Ecosystems and Environment 85, 7-23.

Luck, G.W. (2007). *A review of the relationship between human population density and biodiversity.* Biological review 82. 607 – 645.

Parker, D.C., Manson, S.M., Janssen, M.A., Hoffman, M.J. & Deadman, P. (2003) *Multi-agent systems for the simulation of land-use and land-cover change: A review.* Annals of the association of Americaan Geographers 83 (2). 314-337.

Roehner, B.M. (1995). *Evolution of urban systems in the Pareto plane.* Journal of regional science 35 (2), 277 – 300.

Veldkamp, A. & Lambin E.F. (2001) *Predicting land use change.* Agriculture, Ecosystems and Environment 85, 1 – 6.

White, R., & Engelen, G. (1993). *Cellular automata and fractal urban form: a cellular modeling approach to the evolution of urban land use patterns.* Environment and planning A 25, 1175 - 1199

White, R., Engelen, G. & Uljee, I. (1997). *The use of constrained cellular automata for high-resolution modeling of urban land use dynamics.* Environment and Planning B 24, 323 - 343.

White, R. & Engelen, G. (2000). *High resolution integrated modeling of the spatial dynamics of urban and regional systems.* Computers, Environment and Urban Systems, 24, 383-400.

White, R. (2006) *Pattern based map comparisons.* Journal of geographical systems 8, 145 – 164.

Wu, F. (1998). *An experiment on the generic polycentricity of urban growth in a cellular automatic city.* Environment and planning B 25, 731 – 752.

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 modeling of agricultural land use change in Romania: From retrospective causal analysis to future developments

T. Lakes[a] **and D. Müller**[a,b]

[a] *Geomatics Lab, Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany (Tobia.Lakes/D.Mueller@ geo.hu-berlin.de)*

[b] *Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO), Theodor-Lieser-Str. 2, 06120 Halle (Saale), Germany*

Abstract:
Agricultural land use changed rapidly in Central and Eastern Europe following the collapse of socialism. Cropland is arguably the most dynamic land use class, with cropland abandonment typically exceeding cropland expansion. Yet to date there has been little empirical evidence on the rates, patterns, and processes of cropland change for Central and Eastern Europe. To remedy this, we integrated socioeconomic and environmental data and employed exploratory statistics with predictive simulations to study the causes of past changes, as well as to forecast future cropland development in Argeş County, Romania. In a first step, we used spatially explicit logistic regressions to estimate the direction and strength of the influences underlying factors that led to a change in the extent of cropland. The regressions focused on the exogenous, underlying variables that foster land change, allowed us to rank the importance of factors, and was used to test causal hypotheses of land use change processes. In a second step, we calibrated artificial neural network models with the statistically significant variables to predict the likely spatial arrangement of future cropland change. Such pattern recognition techniques are computationally efficient tools for forecasting locations that are most likely to undergo future cropland changes, given a user-defined quantity of change. Both of the employed modeling approaches are commonly used in land change science and we show an empirical example of how they complement each other. The combination of exploratory and predictive findings are of particular importance for understanding and dealing with complex processes in regions such as post-socialist countries, where empirical evidence on the local driving factors and possible future developments is scarce,. The proposed multi-method modeling approach based on the spatial analysis of integrated human-environment data allows the generation of causal inferences that inform the development of land use change forecasts. We believe that combining both approaches generates insights that are greater than the sum of their parts.

Keywords: spatial analysis; cropland dynamics; logistic regression; neural networks; post-socialist.

1. Introduction

Large-scale land use changes were caused by the collapse of the socialist system in Central and Eastern European countries, as well as in the Commonwealth of Independent States (Peterson and Aunap, 1998; Bicik *et al.*, 2001; Lerman *et al.*, 2004). Since then, agricultural land use change has been dominated by large-scale cropland and pasture abandonment (Kuemmerle *et al.*, 2006; Palang *et al.*, 2006; Kuemmerle *et al.*, forthcoming; Müller and Munroe, forthcoming). This domination is attributed to a combination of several underlying causes, such as the collapse of state support for agricultural production, changing ownership structures, and the emergence of additional income opportunities, in particular a new geographic mobility which led to massive emigration from rural areas. While cropland abandonment has been the dominant land-use change, some areas have experienced cropland expansion; this was particularly so in EU-accession countries that

benefited from EU agricultural policies (Verburg *et al.*, 2006). Until now, neither the patterns of observed changes in agricultural land use, nor their impacts on the human-environmental system are well understood (MacDonald *et al.*, 2000).
We studied changes in Romanian cropland structure using an integrated dataset that contains remotely sensed land cover data, indicators derived from a socioeconomic census, geobiophysical variables and accessibility measures. The driving factors that led to changes in the extent of cropland were examined using logistic regression, and the statistically significant driving factors were employed to forecast future cropland changes using artificial neural network models. Combined, the modeling approach allowed us to draw causal inferences based on exploratory statistical insights, which helped calibrate the simulation models that yielded predictive evidence using statistical pattern recognition.

2. Material and methods

2.1 Study area

Cropland dynamics were investigated in Argeş County, Romania, which is located on the southern foothills of the Carpathian mountain range and which covers an area of 6,826km². The county is shaped by a wide range of environmental conditions and has a mean annual rainfall of 750mm and a mean annual temperature of seven degrees Celsius. Elevation and rainfall are both higher to the north, while temperature is lower. The mountainous northern part is characterized by rugged terrain and includes the foothills of the Carpathian mountain range, as well as Romania's highest mountain, the Moldoveanu Peak, 2,544m high. The hilly midlands contain the county's capital and major market center, Piteşti, which had 174,000 inhabitants in 2003 (NIS, 2004). Piteşti is linked to the country's capital Bucharest by Romania's first highway, built in 1960. The southern part of Argeş County consists of a plain between 100m and 150m above sea level. Most of the southern land in the Argeş region is used as cropland (for more details about the study area, see Kuemmerle *et al.*, forthcoming).

2.2 Data and data pre-processing

Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) images with a spatial resolution of 30m were classified into three land-cover classes: forest, cropland and grassland/shrubland. The data was then validated by means of field data and expert knowledge (see Kuemmerle *et al.*, forthcoming, for details of the classification). In this paper, we focused on the time period between 1995 and 2005 in order to exclude the immediate and rapid responses of land use agents to the fundamental changes in the political system following the collapse of socialism. The cropland class has a producer accuracy of 79% (71%) with a conditional Kappa of 0.78 (0.78) in 1995 (2005). We define cropland abandonment as one if a pixel was covered by cropland in 1995, but not in 2005. Cropland expansion takes the value of one for all cells that were not cropland in 1995, but had been converted into cropland by 2005. Settlements, roads, bodies of water and rivers were assumed to be constant over time and excluded from subsequent analyses. Topographic data was derived from the Shuttle Radar Topographic Mission (SRTM), with a spatial resolution of 90 meters. We calculated elevation, slope and terrain roughness (i.e., slope curvature) from this elevation model. The elevation data serves as a proxy for climate variations and for soil data, as both variables are strongly correlated to the north-south elevation gradient.
To capture the socioeconomic drivers of cropland change, we used census data provided by the National Institute of Statistics of Romania for 1996 and 2003, which we aggregated to the commune level. We merged this census data with the spatially-continuous raster data based on the administrative boundaries of the communes. Commune population density and the number of livestock in each commune were derived from this census. Further, we used village-level population figures from our own census, which we conducted in all rural communes of Argeş County. Spatially-explicit population density maps were created by interpolating the village population data using inverse distance weighting.
Several measures captured the accessibility of each location: We used Euclidean distances to four road categories, as well as the distance to each village center. The surrounding area's cropland density was included as an additional driver, and calculated in a 3x3 pixel

neighborhood as the sum of pixels covered by cropland in 1995. All data were referenced to UTM/WGS84 and re-sampled to a spatial resolution of 100 meters.

2.3 Methods

In the first step, we estimated spatially-explicit, reduced-form logistic regressions to extract the direction and strength of the influences of underlying factors that led to a change in cropland coverage. We estimated two logistic regressions, one for the abandonment of cropland and one for cropland expansion from 1995 to 2005 as the dependent variables. On the right-hand-side of the regressions, we used: topographic variables (elevation, slope, and roughness), distance measures (Euclidean distances to several road categories and to relatively built-up areas), the neighborhood structure of cropland and the interpolated population density as spatially-continuous covariates. At the commune level, we included population density, as well as the density of livestock, as covariates. All data was re-sampled to 100 x 100 meter grid cells and each cell represented one observation in the regressions.

Both regressions followed identical sampling procedures. In the first step, a regular spatial sample from the raster dataset was drawn using Besag's coding scheme (Besag, 1974) to avoid spatial autocorrelation in the dependent variable. We selected every fifth cell in both a north-south and east-west direction. This left us with 24,790 observations, from which we sampled all cells covered by cropland in 1995 for the abandonment equation, and all cells not covered by cropland for the cropland expansion equation. In this way, we incorporated the temporal dependence of the two processes on the state of cropland at the start of the change period (e.g. cells not covered by cropland cannot become abandoned). Following Maddala (1983), we then randomly selected 1,000 observations from the change class (i.e., abandonment or expansion) and from the pixels that did not undergo change. The logistic regression models were estimated based on these 2,000 observations. No multi-collinearity was detected in the resulting sample, as all the covariates' pair-wise correlation coefficients were below 0.7. The constant terms were adjusted to calculate the predicted probabilities due to the unequal sampling rates (Maddala, 1983).

The variables that emerged from the logistic regression with an explanatory power above an arbitrary significance threshold of 20% (p-values < 0.2) were fed into an artificial neural network model (ANN), the Land Transformation Model (LTM, Pijanowski *et al.*, 2005). Again, two models were calibrated as the single output layers, one for cropland abandonment and one for cropland expansion. The parameters estimated by the LTM served to predict the likely spatial arrangement of future cropland dynamics. We used a feed forward network with error backpropagation and a 3-layer structure of input, hidden, and output layer. There is ample flexibility in the choice of the layers in neural network estimations, yet we used 10 input layers for the abandonment model and 7 input layers for the expansion model, as determined by the variables that were above the significance threshold in the logistic regressions (see above). We used the same number of hidden and input layers as suggested by Pijanowski *et al*., (2005).

The neural network was trained on every second cell in the county. For each simulation, called a training cycle, we recorded the overall Kappa and the Percent Correct (PC), where Kappa assesses the accuracy of the location, and PC describes the percent of all cells correctly estimated as the proportion of the total number of cells. We stopped training the network at 250,000 cycles. The network file that performed best was then applied to the entire dataset to calculate a map of suitability values of change for each pixel (neural nets do not create probabilities *per se*, see e.g. Pijanowski *et al*., 2005). The suitability values were used to forecast changes in cropland, and high values were assumed to be more likely during transition. To forecast changes, we varied the number of cells that were expected to undergo change. To do this, we assumed that the yearly number of abandoned and expanded cells will remain stable over time. This resulted in spatial patterns of likely future changes.

3. Logistic regression and neural network results

Results from the spatial logistic regressions showed that cropland change since the collapse of socialism mirrored a distinct spatial variation that significantly contributed to reshaping post-socialist rural landscapes in Argeş County. Results for cropland abandonment are

shown in Table 1. For the subsequent analysis, significance levels are reported as p-values. Abandonment was more likely at higher elevations, steeper slopes, and in rougher terrain. Locations closer to European and national roads were less likely to undergo abandonment, as were locations further away from villages. A spatially more homogeneous cropland structure strongly reduced the likelihood of abandonment (large and highly significant coefficient). Moreover, a higher population density, coupled with a lower livestock density within a commune fostered abandonment.

Table 1. Significant underlying factors of cropland abandonment and expansion

	Cropland abandonment			Cropland expansion		
	Coef.	Std. Err.	p-value	Coef.	Std. Err.	p-value
Elevation (100m)	**0.440**	**0.066**	**0.000**	-0.017	0.033	0.609
Slope (degrees)	**0.179**	**0.030**	**0.000**	**-0.032**	**0.012**	**0.009**
Roughness	**0.030**	**0.011**	**0.005**	0.002	0.005	0.742
Distance to European road	0.000	0.001	0.726	0.000	0.001	0.917
Distance to main road	**0.005**	**0.002**	**0.001**	**-0.005**	**0.002**	**0.005**
Distance to county road	-0.010	0.008	0.211	**-0.010**	**0.005**	**0.036**
Distance to built-up area (100m)	**-0.029**	**0.009**	**0.002**	**-0.009**	**0.003**	**0.008**
Neighboring cropland (3x3 window)	**-0.296**	**0.037**	**0.000**	**0.521**	**0.044**	**0.000**
Population (interpolated), 1996	**-0.010**	**0.006**	**0.139**	0.006	0.005	0.209
Population density (pp/sqkm), 1996	**0.010**	**0.002**	**0.000**	**-0.006**	**0.002**	**0.004**
Livestock unit density per sqkm, 1996	**-0.016**	**0.007**	**0.017**	0.004	0.006	0.464
Constant	0.292	0.462	0.527	0.326	0.284	0.251

Notes: Coefficients (Coef.), standard errors (Std.err.), and p-values below the 20% significance level (p-values < 0.2) are reported in bold.

On the contrary, higher cropland expansion (Table 1, bottom) was expected at lower slopes, closer to European and national roads and closer to all roads (including unpaved communal roads). Cropland expansion was further associated with proximity to villages, high cropland density, and lower population densities.

We assessed the goodness-of-fit of the logistic regressions using three statistics: the PC, the area under the curve (AUC) of the receiver-operating characteristics (ROC), and Cohen's Kappa (Pontius, 2002). Given the small number of covariates in the model, the goodness-of-fit was moderate-to-good for abandonment (Table 2), with 71% of the pixels predicted correctly. Yet the logistic regressions showed moderate-to-low fit for cropland expansion. While the AUC was good, Kappa was especially unsatisfying for cropland expansion. Hence, we may not have included the driving factors that shape the spatial location of the expansion processes in our study area. We believe there are considerable stochastic elements involved, particularly in cropland expansion and, looking at the observed expansion map, we are unable to see spatial clusters. Therefore, we have low confidence in the underlying factors derived by the logistic regressions, and subsequently expect low accuracy results for the LTM predictions of cropland expansion. Nevertheless, we continued to account for both processes in order to be able to generate countywide forecasts of cropland changes.

Table 2. Accuracy assessment of the spatial logistic model and ANN results for abandonment and expansion (1995-2005).

	Model	**PC**	**Kappa**	**AUC**
Abandonment	Logistic	71.3%	0.43	0.882
	ANN	63.1%	0.56	-
Expansion	Logistic	59.2%	0.05	0.834
	ANN	34.8%	0.30	-

We trained the neural network for both change processes with 250,000 cycles and achieved the highest accuracy according to PC and Kappa for cropland expansion at cycle 250,000 and for cropland abandonment at cycle 100,000 (Table 2). For cropland

abandonment, both the PC and Kappa showed moderate fit, with 63% and 0.56%, respectively, predicted correctly. As expected, the simulation results were not satisfactory for cropland expansion with PC equaling 35% and a Kappa of 0.3. It also needs to be considered that all cells, including the ones used for training, have been included in calculating these assessment results.

The neural networks' output indicated how likely it was for a cell to transition out of cropland (abandonment) or to be converted to cropland (cropland expansion). Figure 2 shows the spatial patterns of likelihood for cropland abandonment on the left by using the 'best' cycle. The likelihood for abandonment was particularly high in the hilly areas of Argeş (middle area in Figure 2). A lower likelihood, but one which still affects a substantial number of pixels, appears in the lowland areas. For cropland expansion (Figure 2, right) the picture differed and the highest likelihood values for expansion were found around the valleys of the hilly area. There, large tracts of land are not yet used as cropland (as in the South) and are more suitable for cropland production than the remote areas in the northern mountains, where the likelihood for expansion is close to zero. Little expansion is expected in the southern part of the county, where most cells were already used as cropland and in the Northern parts.

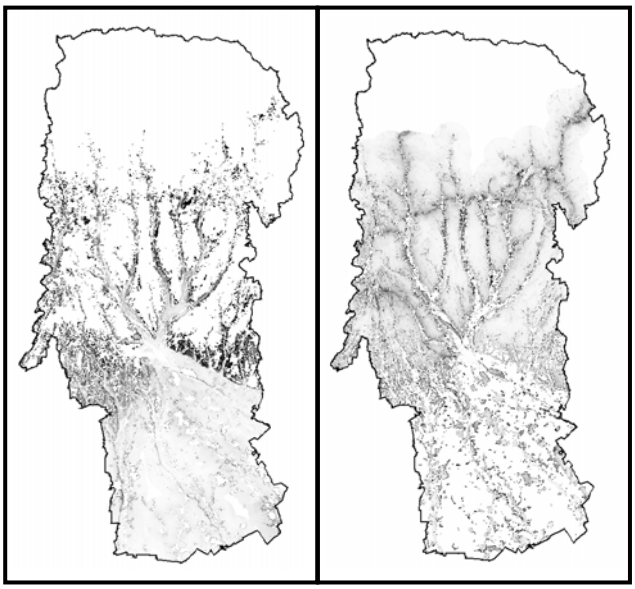


Figure 2. Spatial patterns of likelihood for cropland abandonment (left) and cropland expansion (right); darker colors indicate higher likelihoods.

Next, we used the neural network patterns from the cycle with the highest accuracy to forecast the likely spatial arrangement of future cropland change. The calculation of the forecasts was based on the observed change between 1995 and 2005, and this trend is expected to continue in the future. Forecasts were calculated based on past patterns and also on using twice (Forecast 1) and four times (Forecast 2), respectively, the number of cells that changed during the calibration period of 1995-2005. The forecasts of abandonment depicted the largest number of future change in the middle, hilly area of Arges County (see Figure 3 below). The southern plains were less likely to be abandoned, while cropland expansion was concentrated on the hilly areas (see Figure 3). Abandonment is expected to start in central Argeş County, and from there proceed to the south of the study area (see fig. 3). While in forecast 1, the areas with higher slopes will be abandoned, in forecast 2, the southern plains are also likely to be abandoned. Moreover, abandonment follows the major European roads that cross Argeş County from the southeast to the northwest and from the southwest to the northeast. Cropland expansion was concentrated in the hilly areas in both periods. Forecast 1 shows that non-cropland pixels which were adjacent to existing cropland are more likely to convert to cropland. In forecast 2, we see additional potential for cropland expansion in more remote areas of northern Argeş County.

4. Discussion: From retrospective causal analysis towards future developments

4.1 Advantages and disadvantages of the coupled modeling approach

The spatial logistic model focused on exogenous, underlying variables that determine land change. Such models are valuable for ranking the importance of factors and testing

hypotheses about the underlying processes of land change (Munroe and Müller, 2007), and thus provide a means for retrospective causal analysis. The knowledge gained from the exploratory spatial statistics helped us determine the calibration of the neural networks which relied on a similar statistical technique, i.e., logistic regression. Variables that turned out to be significant in the logistic regressions were therefore expected to have a bearing on the outcome of interest.

Yet the inclusion of insignificant variables should not affect the results of the neural network to a considerable extent. One important advantage of neural networks is the data-driven approach, based on machine learning. This approach enables modeling across space, time, and datasets; hence, the networks are a generalization tool that are not constrained by the requirement of considering incorrect sampling, multicollinearity between variables, spatial or temporal autocorrelation, or the insignificance of single variables (Bishop, 1995). This is an obvious advantage for integrated land use modeling of coupled human-environment systems, which necessitate a large number of different data, often from different sources and of varying quality.

However, neural networks also have some drawbacks, most of which are related to their data-driven approach. Neural networks do not require a priori knowledge about underlying processes; on the contrary, they are designed to recognize patterns that result from such processes. However, this requires a considerable amount of computing time, which is a limiting factor for analyses that require large amounts of data and myriads of potentially influencing factors, as is often the case in land change research. Apart from that, in almost every LUCC study, a priori top-down knowledge is prevalent; hence, the benefit of the neural networks' data-driven approach is less important. This is particularly so because it is the knowledge of the processes, not of the resulting patterns, which modelers and decision-makers are aiming for. This black-box approach assumes that the user does not know how the data is processed in detail, nor is the weighting of the included driving factors accessible. In this study, we address this issue by selecting the statistically most significant underlying factors by using the spatial logistic model. Furthermore, dynamics in processes and underlying factors in the coupled human-environment system are not explicitly integrated in the modeling approach. For example, in rapidly changing areas the neglected dynamics of population migration might result in misleading future development trajectories. Here, we tried to avoid processes dominated by the transformation phase from 1990 to 1995 by calibrating the models with the subsequent time period, which we expect to be more representative for likely future developments. A further extension, the artificial neural network, is not able to learn from evidence which contains previously un-introduced variables without a new network being created and trained, such as is the case with Bayesian networks (Aalders and Aitkenhead, 2006).

Hence, the coupling of the two modeling techniques – logistic regression and neural networks – addressed two issues of concern for land-change scientists: the exploration of underlying causes and the forecasting of likely future developments. Moreover, the selection of input variables for the neural networks was based on empirical grounds. This reduced the time required for training and testing the networks, which is a considerable advantage when dealing with large data sets.

4.2 Need for an integrated human-environmental system approach

Studying cropland dynamics requires an integrated approach that explicitly addresses the coupled human-environment system of land use (Turner II *et al.*, 2007). Changes in the extent of cropland are important for understanding the impacts of these changes on the multifunctional nature of agricultural production (OECD, 2001). Cropland dynamics also mirror large-scale socioeconomic alterations such as migration or changing economic opportunities that are reflected in the landscape (Müller and Munroe, forthcoming).

The two modeling techniques employed here allowed us to integrate various types of economical, social, and ecological data. The results of the spatial logistic regressions revealed retrospective causal influences on land-use decisions regarding cropland usage. The neural networks allowed us to use this information to predict likely future cropland changes for each location in the entire county. Taken together, these spatial insights may prove more useful for decision-makers than would a single-method approach (see Fig. 3).

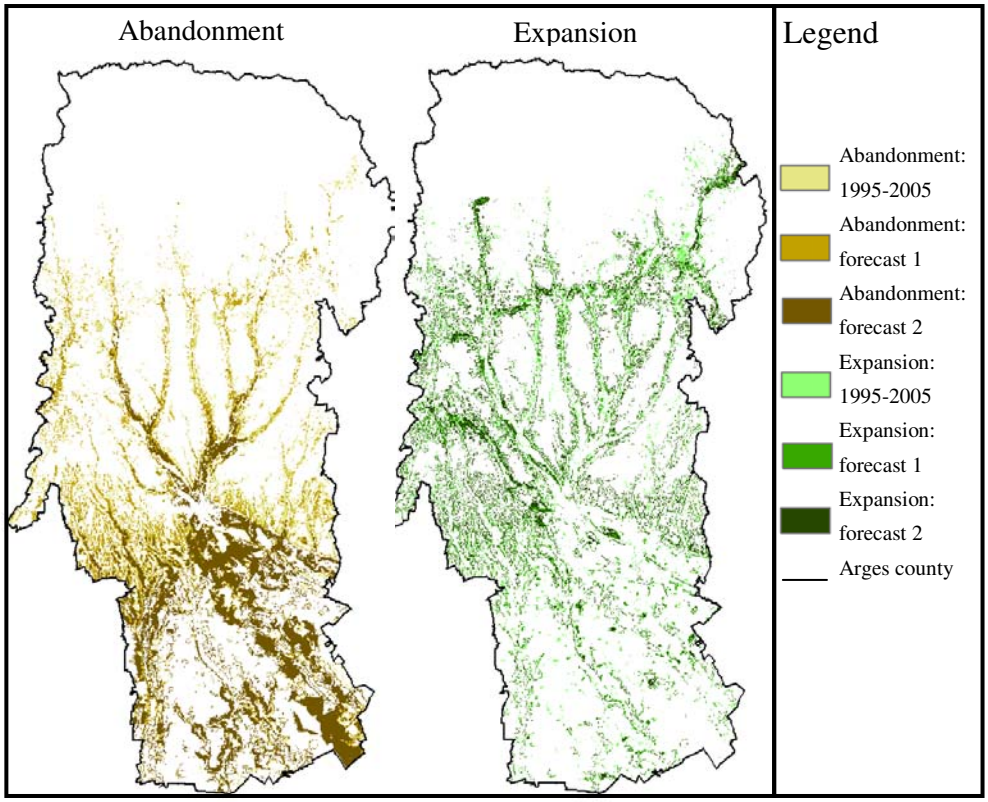


Figure 3. Observed and forecasted spatial patterns of cropland abandonment (left) and expansion (right)

However, there are several drawbacks that both models have in common when striving for an integrated model. One major limit is that they are not able to explicitly model the feedbacks between variables which should be an inherent characteristic of integrated modeling techniques of complex human-environment systems. The integrated approach of the complex human-environment land use system is only rarely pictured in the modeling techniques. Temporal changes of the influencing factors are not incorporated within both modeling approach. Yet, more holistic modeling approaches often require much more complex computational approaches and are more data-demanding. We believe our ad-hoc integration of models yielded exploratory and predictive findings that are important to understand and deal with processes in regions where empirical evidence on the local driving factors and possible future developments is scarce, such as in the case of most post-socialist countries. Several extensions are possible, such as a larger extent of the area to have a greater variety and exploit the computational advantages of the modeling techniques. One major benefit would the inclusion of external information about future land use demand was available to produce more realistic forecasts of future spatial patterns (Castella and Verburg, 2007).

5 Conclusions and outlook

Our approach of coupling two modeling techniques combined the benefits inherent in explorative statistics with the advantages of pattern recognition. This proposed multi-method modeling approach allowed us to generate causal inferences and to develop land use change scenarios that can offer new insights for decision-making support. The integrated spatial analysis of socioeconomic and environmental data may support spatial planning, environmental monitoring, and biodiversity conservation efforts to areas that are under the most eminent threat of land change. This is important because agricultural extensification and cropland abandonment are expected to be the largest land-use changes in the enlarged European Union, particularly in marginal areas (Verburg *et al.*, 2006).

ACKNOWLEDGMENTS

The authors wish to thank Tobias Kuemmerle for the remote sensing analysis of land cover changes supported by research assistants of the Geomatics Lab, as well as Carsten Krueger for support in various model runs. The authors are also grateful to the Deutsche Forschungsgemeinschaft (DFG) for having supported this research under the Emmy-Noether Programme.

REFERENCES

Aalders, I.H., Aitkenhead, M.J., 2006. Agricultural census data and land use modelling. Computers, Environment and Urban Systems 30, 799-814.

Besag, J.E., 1974. Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society B 36, 192-236.

Bicik, I., Jelecek, L., Stepanek, V., 2001. Land-use changes and their social driving forces in Czechia in the 19th and 20th centuries. Land Use Policy 18, 65-73.

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, Oxford.

Castella, J.-C., Verburg, P.H., 2007. Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. Ecological Modelling 202, 410-420.

Kuemmerle, T., Müller, D., Griffiths, P., Rusu, M., forthcoming. Land use change in Romania after the collapse of socialism. Regional Environmental Change.

Kuemmerle, T., Radeloff, V.C., Perzanowski, K., Hostert, P., 2006. Cross-border comparison of land cover and landscape pattern in Eastern Europe using a hybrid classification technique. Remote Sensing of Environment 103, 449-464.

Lerman, Z., Csaki, C., Feder, G., 2004. Evolving farm structures and land use patterns in former socialist countries. Quarterly Journal of International Agriculture 43, 309-335.

MacDonald, D., Crabtree, J.R., Wiesinger, G., Dax, T., Stamou, N., Fleury, P., Gutierrez Lazpita, J., Gibon, A., 2000. Agricultural abandonment in mountain areas of Europe: Environmental consequences and policy response. Journal of Environmental Management 59, 47-69.

Maddala, G.S., 1983. Limited-dependent variables in econometrics. Cambridge University Press, New York.

Müller, D., Munroe, D.K., forthcoming. Changing rural landscapes in Albania: Cropland abandonment and forest clearing in the postsocialist transition. Annals of the Association of American Geographers.

Munroe, D.K., Müller, D., 2007. Issues in spatially explicit statistical land-use/cover change (LUCC) models: Examples from western Honduras and the Central Highlands of Vietnam. Land Use Policy 24, 521-530.

NIS, 2004. Database of Localities. National Institute of Statistics, Bucharest.

OECD, 2001. Multifunctionality: Towards an Analytical Framework. Organisation for Economic Co-operation and Development (OECD), Paris.

Palang, H., Printsmann, A., Gyuró, É.K., Urbanc, M., Skowronek, E., Woloszyn, W., 2006. The Forgotten Rural Landscapes of Central and Eastern Europe. Landscape Ecology 21, 347-357.

Peterson, U., Aunap, R., 1998. Changes in agricultural land use in Estonia in the 1990s detected with multitemporal Landsat MSS imagery. Landscape and Urban Planning 41, 193-201.

Pijanowski, B.C., Pithadia, S., Shellito, B.A., Alexandridis, K., 2005. Calibrating a neural network-based urban change model for two metropolitan areas in the upper midwest of the united states. International Journal of Geographical Information Science 19, 197-215.

Pontius, R.G.J., 2002. Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogrammetric Engineering & Remote Sensing 68, 1041-1049.

Turner II, B.L., Lambin, E.F., Reenberg, A., 2007. The emergence of land change science for global environmental change and sustainability. Proceedings of the National Academy of Sciences 104, 20666-20671.

Verburg, P.H., Schulp, C.J.E., Witte, N., Veldkamp, A., 2006. Downscaling of land use change scenarios to assess the dynamics of European landscapes. Agriculture, Ecosystems & Environment 114, 39-56.

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

Go East: A Residential Land Use Model for the Periphery of Rome

G. L. Ciampaglia[a] **and B. Tirozzi**[b]

[a]MACS *Lab., University of Lugano, V.le Canavée,* CH-*6850 Mendrisio, Switzerland*
(ciampagg@lu.unisi.ch)

[b]*Dept. of Physics, University of Rome "La Sapienza", P.le Aldo Moro 8, 00100 Rome, Italy*
(b.tirozzi@libero.it)

Abstract: In this paper we present a model of urban growth and its preliminary application to a case study of the phenomena of residential development in the setting of the eastern periphery of Rome, Italy's capital city. The modeling approach we use synthesises the two typical paradigms widespread in the community of quantitative urban planning: the traditional one, based on cellular automata (CA), and the (relatively) new one, which is agent-based. In particular, our multi-agent system (MAS) is in-between a reactive MAS, with agents carrying out a two-staged decision process in a complex environment, and a model of statistical physics, since we use populations of agents in order to reduce the number of degrees of freedom of the system. While we explicitly model the consumption of agricultural and undeveloped land due to urban growth, our model may be easily integrated as a socio-economic part into a wider decision support system for environmental planning, e.g. our simulations can produce indicators of environmental impact of the growth of the city: electricity consumption, waste production, etc.

Keywords: Urban growth impact, Multi-agent system, Cellular Automaton, Rome, Model integration.

1 INTRODUCTION

It is now a matter of fact that the complex systems perspective has been fully accepted in urban and regional planning studies. Since their introduction, the two major bottom-up approaches, the one of multi-agent systems (see Batty and Jiang [1999]; Benenson [1998]) and cellular automata (see White and Engelen [2000]; Engelen et al. [2003]) have contended for the lion's share of the literature, with the current trend to synthesize the best from both approaches. The work we present in this paper is based on one recent attempt to provide such an interpolation (see Vancheri et al. [2008] for a primer to this methodology, and the references therein) and deals with the development of a major urban area characterized by sprawling phenomena and unregulated residential growth for many decades. From a methodological point of view, connecting MAS with CA is attractive because it turns out to be very easy to model a urban or geographical system with such a combination of approaches. The baseline is to identify "mobile" entities in the cognitive agents, such as the inhabitants of a city and to subdivide, on the other hand, the urban area in a regular lattice. One then models more or less faithfully the behaviour of the agents, usually at the cognitive level, while assigning to each cell a label describing the urban typology of the part of territory assigned to it. This procedure usually takes the form of the analysis, design and implementation of a software simulator of the complex interactions between agents and the cells, and the simulations produced with such a tool may provide valuable informations of the phenomena one wants to investigate. It is widely accepted that the introduction of multi-agent systems solved the problem of providing a meaningful description of the processes of land use change undergoing in

the system, without the need to describe these in terms of abstract interactions between the cells of a CA: it is much more convenient to describe in terms of the agents those socio-economical interactions that – at least one empirically assumes – drive the urban transformations of a city, rather than account for the existence of an ubiquitous process of "update", only indirectly based on the same interactions, and only explainable in terms of the local state of the neighbourhoods on the lattice of cells. However, this combined approach has disadvantages, no matter the degree of fidelity on can achieve with a well engineered software simulator. It is in fact difficult to calibrate and validate the simulations produced with it, and many times a proper sensitivity analysis is only able to identify critical parameters of the model, but not to truly assess what the uncertainty in the simulations' measurements is due, especially in presence of bifurcations or phase transitions of the original system. It seems thus important to provide a modelling framework that gives a proper mathematical definition of the entities one is going to model, and that at the same time contemplates the possibility to build a software simulator that cheaply outputs interesting scenarios about the future evolution of the city.

The work we present in this paper is an attempt to go in this direction. The structure of the paper is the following: in section 2 we introduce the main mathematical features of our model, and in particular in 2.1 we detail respectively how the collective decisions of the agents can produce the stochastic dynamics for the evolution of the cellular automaton, and in section 2.2 we look at how it is possible to synthesize a description of the city that enables each agent to take realistic decisions about urban events. In section 3 we discuss the preliminary results we got from running simulations under the assumptions of a stationary dynamics for the configuration of the system. Finally, in section 4 we discuss how the integrated approach to the production of computer tools for environmental assessment and decision making could benefit from our simple methodology.[1]

2 A SPATIAL MODEL OF URBAN GROWTH AND INTELLIGENT AGENTS

To a first approximation, with the CA we model the land uses of the urban system. As is usual for this kind of models, each cell of the CA is a 2-dimensional representation of a given piece of land belonging to the urban area. Our choice was to consider each cell corresponding to an administrative zone, as used in the master plan of the city; thus, we have an irregular lattice Γ, with adjacency relations between cells given by the actual geographical boundaries between zones. Figure 1 shows the cellular decomposition of the CA used in the case study of Rome.

We take a state space which is real-valued and multidimensional, with each cell $c \in \Gamma$ described by a vector of dynamical variables $\boldsymbol{v}(c,t) \in \mathbb{R}^p$, subject to the update rule of the automaton, and by some control variables that follow an exogenous dynamics, $\boldsymbol{w}(c,t) \in \mathbb{R}^q$. Dynamical variables are chosen to have a clear urban and geographical meaning. Examples include: number of households, total surface area of residential buildings, total surface area of empty land. In $\boldsymbol{w}$ there are instead variables that are not directly subject to the update rule of the automaton: distances from hospitals and services, access points to the transportation network, and other features related to infrastructure. The variables in $\boldsymbol{w}(c,t)$ can be used to define exogenously the settings for a simulative scenario. Taking $\boldsymbol{v}$ and $\boldsymbol{w}$ together for each cell c, the system is described by a n by d matrix $\boldsymbol{X}(t)$ where $n = |\Gamma|$ and $d = p + q$. The rule that updates the state of the cells of the automaton is not specified as a deterministic function, with fixed adjacency neighbourhood, as in the classical specification of CA. Instead, the updates of the cells' state are performed by decisor agents: the basic idea is that social entities interact locally with the city – that is, agents interact with cells

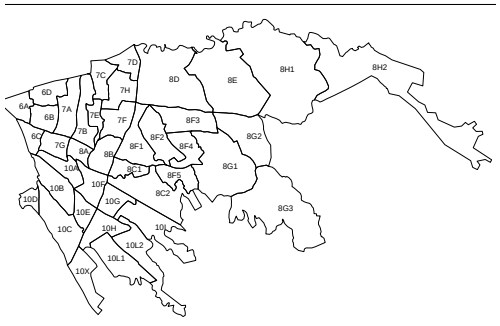


Figure 1: Cellular subdivision of the urban area in the case study of the eastern periphery of Rome, Italy.

[1] For additional material on this model, please see the following: http://www.inf.unisi.ch/phd/ciampaglia

Variable	Description	Update rule
N	Total number of free available flats in the cell	$N + \Delta n$
S_u	Total undeveloped surface in the cell	$S_u - (\Delta s_1 + \Delta s_2)$
S_b	Total built surface due to residential use	$S_b + \Delta s_1$
S_p	Total built surface due to parkings and paved use	$S_p + \Delta s_2$

Table 1: Update rule of the state of a cell due to the construction of an apartment building

– with possible interactions stereotyped by the elements of a set $\mathcal{A}$ of kinds of interactions.[2] in modeling the elements of $\mathcal{A}$, we have in mind classes of events with a clear urban meaning and scope, e.g.: construction of houses, malls and offices, changes in zoning regulations of free lots. It is thus possible to model very easily the effects of these simple interactions; see table 1 for an example of the update rule used to model the construction of an apartment building.

More formally, each $\alpha \in \mathcal{A}$ defines an incremental stochastic dynamic of the form:

$$v_k(c, t + \Delta t) = v_k(c, t) + \pi_k(\omega, c, t) \tag{1}$$

for $k \in V_\alpha \subseteq \{1, \ldots, p\}$, e.g. for some of the variables that make up the components of $\boldsymbol{v}(c, t)$. We take the standard assumptions so that the counting variables $N_{\alpha,B}(c, t)$ of the number of events of kind α, with increments $\pi \in B \subseteq \mathbb{R}^{n(\alpha)}$, and occurring in c during the time interval $[t, t + \Delta t)$, have law defined by the Poisson distribution with parameter $\lambda_{\alpha,B}(c, t)$. Thus we need to compute the intensity of a non-homogeneous Poisson process (see Kingman [1993]); the idea is to define it in terms of a density λ^α:

$$\lambda_{\alpha,B}(c, t) = \int_B \lambda^\alpha(c, \pi, t) d\pi \tag{2}$$

where $\lambda^\alpha(c, \pi, t)$ is the density of probability that in $[t, t + \Delta t)$, with Δt small enough, one interaction of kind α with increments in $[\pi, \pi + \Delta\pi)$ occurs in c. We then introduce the conditional probability $\beta^\alpha(c, \pi, t)$, that one event has increments in $[\pi, \pi + \Delta\pi)$, given it actually occurred. Then obviously $\lambda^\alpha(c, \pi, t) = \lambda^\alpha(c, t) \cdot \beta^\alpha(c, \pi, t)$, where

$$\lambda^\alpha(c, t) = \int_{\mathbb{R}^{n(\alpha)}} \lambda^\alpha(c, \pi, t) d\pi \tag{3}$$

This means that we are able to decouple the problem of modelling an interaction of kind α into two sub-models: the first accounts for how many interactions occur in a cell during a small time interval, the second allows us to generate the values that characterize the event that results from of an interaction between an agent and a cell. If we consider Δt to be the time unit, then $\lambda_{\alpha,B}(c, t)\Delta t$ is the average number of events per time unit. Of course, in the spirit of CA modeling, this quantity has to depend on the local configuration of the system, where the concept of 'locality', for any given α, plays the same role of that of neighbourhood in the classic definitions of CA. A reasonable assumption for this class of systems is that, for values of Δt small enough, the information about the events occurring in a cell during Δt does not change the configuration of the system, so that the intensities are constant during the time interval. This assumption let us to recover a time homogeneous process for each step of update of the system, which is of great aid in the simulation of the process.

2.1 Decision dynamics of the multi-agent system

We now briefly detail the first of the two sub-models, which enables us to compute $\lambda^\alpha(c, t)$. Each α has a population of decisor agents. Unlike those MAS where each agent is equipped with a complex agenda and models faithfully a given sociological class of individuals, we take the opposite approach and reduce the degree of freedoms of the system by describing populations of temporary and anonymous agents. An agent may be inactive, or may be active and locatable in a specific zone, e.g. resident in a cell. The decision process of an agent is composed of

[2] we say "social entity" since an agent may also be e.g. an enterprise looking for a venue where to open a new office.

four actions or steps: activation (A), diffusion (D), update (U) and leaving (L). After activation, an agent 'enters' into the automaton and is placed in a cell. From there it may either diffuse – that is, jump – to another cell, update the state of the cell by realizing an event of kind α (and subsequently become inactive and 'exit' the CA), or leave the decision process and become inactive, with the same consequences of the update step. We define four processes, at the global level of the whole automaton, for the above actions. Let these processes have each intensity Λ_i^α, for $i \in \{A, L, D, U\}$. This means that, as an example, Λ_A^α is the average number of agents belonging to the population of agents α that become active in any cell of the CA per time unit. The idea is to consider the global intensities Λ_i^α as parameters of the model and then, thanks to the property of composition of Poisson processes, distribute (or, generally speaking, assign) the overall rate of activations, diffusion, etc. among the cells. The following four formulas explain how this idea is put in practice:

$$\lambda_A^\alpha(c,t) = \Lambda_A^\alpha(t) \cdot \frac{F^\alpha(c,t)}{\sum_{c' \in \Gamma} F^\alpha(c',t)} \qquad \lambda_D^\alpha(c,t) = \Lambda_D^\alpha \cdot \frac{F^\alpha(c,t)}{\sum_{c' \in \Gamma} F^\alpha(c',t)} \tag{4}$$

$$\lambda_L^\alpha(c,t) = \Lambda_L^\alpha \qquad \lambda_U^\alpha(c,t) = \Lambda_U^\alpha \cdot G^\alpha(c,t) \tag{5}$$

The first formula says that the intensity for activations of agents in a cell c is distributed proportionally to a global 'attractivity' force $F^\alpha(c,t)$. Intuitively speaking, $F^\alpha(c,t)$ models how much the cell c is favorable to the occurrence of interactions of kind α due to the regional context in which c is: it integrates informations at a macroscopic or regional level, such as residential centrality in the case when α is the interaction of renting a flat; a similar definition has been used in the second formula for the intensity $\lambda_D^\alpha(c,t)$, of jumps that have cell c as a target, that is, regardless of the origin cell. The effect of these two definitions is to bias the exploration of the urban space, that the agents do during their life cycle, towards those areas of the city that exercise more force of attraction than others. The dependency of $\Lambda_A^\alpha(t)$ on time, as we shall see later, lead us to manipulate the number of events of kind α that actually occur during a simulation. The third formula defines the average time an agent spends in the active state to be $1/\Lambda_L^\alpha$. The intensity $\lambda_U^\alpha(c,t)$ is proportional to another attractivity force called $G^\alpha(c,t)$, which integrates information on the state of c, as usual with respect to α, at a more local and detailed level.

Before getting into the discussion on our approach to the modeling of the configuration of the system, that is, on the definition of F^α and G^α, we have to explain how, starting from the intensities of the local processes, the above definitions are brought together and give the overall dynamics of the MAS along a simulation's step $[t, t+\Delta t)$. In turn, this let us to compute the intensities $\lambda^\alpha(c,t)$ of the non-homogeneous processes of update events of the CA's state. If we analyse the number of agents entering and leaving a cell due to the four actions just introduced, we can derive a first order differential equation for the probability $P^\alpha(c,t)$ on the space of the states of an agent:

$$\begin{aligned} \frac{dP^\alpha}{dt} \;=\; & \frac{F^\alpha(c,t)}{T^\alpha(t)} \cdot \Big(P^\alpha(\overline{\Gamma},t) \cdot \Lambda_A^\alpha(t) + (1 - P^\alpha(\overline{\Gamma},t)) \cdot \Lambda_D^\alpha \Big) \;\; + \\ & - \;\; P^\alpha(c,t) \cdot \Big(\Lambda_D^\alpha + \Lambda_L^\alpha + \Lambda_U^\alpha \cdot G^\alpha(c,t) \Big) \end{aligned} \tag{6}$$

Where $T^\alpha(t) = \sum F^\alpha(c,t)$ is the normalizing factor, and $P^\alpha(\overline{\Gamma},t)$ is the probability that an agent is passive and thus not in any cell of Γ. It is then possible to find a solution, for the stationary case $dP^\alpha/dt = 0$:

$$P^\alpha(c,t) = \frac{F^\alpha(c,t) \cdot \Big(P^\alpha(\overline{\Gamma},t) \cdot \Lambda_A^\alpha(t) + \big(1 - P^\alpha(\overline{\Gamma},t) \cdot \Lambda_D^\alpha\big) \Big)}{T^\alpha(t) \cdot (\Lambda_D^\alpha + \Lambda_L^\alpha + \Lambda_U^\alpha \cdot G^\alpha(c,t))} \tag{7}$$

The assumption of stationarity is motivated by the fact that the configuration of the system is determined by variables that, with respect to P^α, have a slow dynamics. Now, since we are dealing with Poisson processes, $\lambda_U^\alpha(c,t)$ approximates the probability that one event occurs in a small time step of duration Δt; thus $\lambda_U^\alpha(c,t) \cdot P^\alpha(c,t)$ gives the probability that an active agent performs the update step in the cell – that is, that an event occurs – and if we multiply that by the number of agents in the population, we obtain the average number of events of kind α occurring in $[t, t+\Delta t)$. For this construction to work properly, it is important to ensure that the number of

events does not fluctuate too much around this expected value. This condition is satisfied, thanks to the law of large numbers, in the limit of the number of agents growing to infinity. In practical simulations, this condition is usually fulfilled with a number of active agents on the order of the thousands.

2.2 The urban configuration

As already stated, $G^\alpha(c,t)$ is the piece of information with which active agents decide whether to make an update in the current cell or not. In Vancheri et al. [2005], from which our work inherits the basic modeling framework, there is not a single and general model for such a force; instead, the authors use fuzzy decision theory and develop several models of G^α, one for each kind of event $\alpha \in A$, by refining and aggregating multiple indicators – by and large, demographic and geographic data – with the aid of fuzzy t-norms and t-conorms. This approach is very powerful when empirical models of the different α are available: in such cases fuzzy modeling is indeed a suitable tool for the translation in the formal language of mathematics.

The definition of the force of attraction $G^\alpha(c,t)$ we give is inspired by principal component analysis (PCA). PCA is a widely used technique for multivariate analysis (see Kent et al. [2006]), and can be also viewed as a simple form of unsupervised learning.[3] A common data analysis task that can be done with PCA is the identification, for each principal component, of a subset of the original variables that have highest correlation in absolute terms with that component, beyond a certain threshold (see Everitt and Dunn [2001]). Usually human experts are able to synthesize meaningful indices using the principal components. In our case, if we look at how a kind of interaction α is defined, there's already a subset of variables having a special status with respect to α. These variables have their dynamical behaviour influenced by interactions of kind α, that is, those with index $k \in V_\alpha$ in equation (1). One can compute the cell's scores, for these variables, in the space spanned by the principal components, and assess the cell's attractivity towards the kind of interactions α by looking at those scores. This is, conceptually, the opposite of the operation of principal components' identification stated above: we choose to model the force of attraction in terms of a fixed subset of variables, and then PCA automatically synthesizes an index that measures how much any cell is attractive, with respect to those variables. This technique has some disadvantages – which we'll discuss later – but it allows for a simple and general purpose model of the urban forces G^α.

Let $\boldsymbol{X} = \boldsymbol{X}(t)$ be the dynamic matrix of multivariate data we apply PCA to.[4] $\boldsymbol{X}$ is constituted of $n = |\Gamma|$ samples or data points, one for each cell c, and of $d = p + q$ observations of demographic and economic variables, that is, the concatenation of vectors $\boldsymbol{v}(c,t)$ and $\boldsymbol{w}(c,t)$. The principal components are standardized linear combinations of the urban variables, $\boldsymbol{y}_k = \boldsymbol{a}_k' \boldsymbol{X}$. Since the variables have different units of measure, we compute the coefficients $\boldsymbol{a}_k$ by diagonalization of the sample correlation matrix $\boldsymbol{R}$ of the data, $\boldsymbol{R} = \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}'$: the $\boldsymbol{a}_k$ are then the orthogonal columns of $\boldsymbol{A}$. Since the sample points are centered around their mean $\boldsymbol{\mu}$, for the i-th cell it is possible to compute the projection on the k-th component component as $y_{ik} = \boldsymbol{a}_k'(\boldsymbol{x}_i - \boldsymbol{\mu})$. In the terminology of PCA, $y_{i1}, \ldots, y_{id}$ are called the scores of the i-th sample. These 'raw' scores, however, are not suitable for computing an index. First, the original variable we take into account might correlate negatively with the component. Since our data are centered, and the transformation induced by the PCA is just a rotation of the space, values of that variable that are less than the mean have actually negative score on the component, while we want them to give a positive contribution to the total score of the cell. A symmetric argument holds if the correlation is positive. Moreover, it is desirable to take into account how much of the variance of the original variable the principal component is able to explain. Finally, since we are going to use this index

[3] A very good review is Roweis and Ghahramani [1999], in which PCA is presented as a learning problem in a linear model with latent variables, under the assumption that the hidden state is constant and constituted of independent normal variables and that the linear dependency of the observations on the state is affected by additive Gaussian noise with infinitesimal variation.

[4] Matrices are denoted by upper case bold letters, vectors are always column vectors and are denoted with bold lower case letters; thus $\boldsymbol{a}'\boldsymbol{b}$ is the scalar product of two vectors. When referencing matrices, i, j are the row/column indices for the original data, while h, k always refer to the principal components.

to define an intensity, which is positive definite, we want each score to give a positive contribution as well. Since it is possible to compute the correlation between the j-th variable and the k-th component as $l_{jk} = a_{jk}\sqrt{\lambda_k}$, then the above considerations lead us to define the contribution that each score gives as:

$$y^*_{ijk} = \frac{1}{\pi}\ \tan^{-1}\Big(\frac{l_{jk}\ y_{ik} - \mu_{jk}}{\sigma_{jk}}\Big) + \frac{1}{2} \tag{8}$$

so that $0 < y^*_{ijk} < 1$. μ_{jk} and σ_{jk} are the sample mean and standard deviation of the raw scores of the cells of the j-th variable with respect to the k-th principal component. Now let us consider the i-th cell c_i and the kind of interaction α. Let us denote with g^α the weighted sum of the scores (8) on the first $r \leq d$ principal components:

$$g^\alpha(c_i, t) = \sum_{j \in V_\alpha} \sum_{k=1}^{r} y^*_{ijk} \tag{9}$$

where, as stated, j is restricted to range in V_α, the subset of variables whose dynamics is affected by the update rule of α. r is chosen so that the first r components explain at least 50% of the variance of the original data. Finally, to define the dynamics of the force $G^\alpha(c_i, t)$, we introduce a temporal delay to smooth the changes of the term in (9):

$$G^\alpha(c_i, t + \Delta t) = (1 - \epsilon)\ G^\alpha(c_i, t) + \epsilon\ g^\alpha(c_i, t + \Delta t) \qquad \text{for } t > 0 \tag{10}$$

$$G^\alpha(c_i, 0) = g^\alpha(c_i, 0) \tag{11}$$

where $0 < \epsilon \leq 1$ acts as a learning rate.

The definition of $F^\alpha(c, t)$ 'averages' $G^\alpha(\cdot, t)$ over the regional context of c:

$$F^\alpha(c, t) = \sum_{c' \in \Gamma} i(c')\ h\big(d(c, c')\big)\ G^\alpha(c', t) \tag{12}$$

$$h(x) = h(x; m, n, h_0) = m\Big(1 - \frac{x^n}{x^n + h_0}\Big) \tag{13}$$

In (12) the regional context of a cell generalizes the concept of neighborhood of a CA by means of a simple gravitational model: the contribution of each cell c' is proportional to $G^\alpha(c', t)$ (e.g. its mass), and since (13) is a monotonically decreasing step function, decreases with a measurement of the distance between the cells. Such a measurement, which should be taken with respect to the transportation network of the city, is modeled by the term $i(c')h(d(c, c'))$. Usually, the network is explicitly modeled as a labeled digraph, and one takes a suitable graph-theoretic measure of integration of a node in a graph; however, the only data we had for our case study were the distances, from the center of mass of the developed areas of the cell, to the nearest access point to the transportation network (d_{net}), hospital (d_{hos}), university or school (d_{edu}) and major shopping or service center (d_{ser}). If we allow k to range in $\{net, hos, edu, ser\}$, our integration measure is:

$$i(c) = \prod_k h\big(d_k(c); 2, n, \mu_k\big) \tag{14}$$

The parameter $n = n_k$ can be set so that the k-th factor $\approx 3/2$ when $\mu_k - d_k(c) \approx \sigma_k$, which equals to reward those cells that are better integrated, and conversely to weight less those that are 'distant' from hospitals, schools, etc. Finally, note that, by using in (12) the euclidean distance $d(c, c')$ between the centers of the cells, we are taking the gross assumption of a homogeneous transportation network.

3 SIMULATION RESULTS

The model has been implemented for a case study on the eastern area of the city of Rome, Italy. The CA has $|\Gamma| = 40$ cells, each grossly corresponding to an administrative zone,[5] and each cell

[5] According to the Italian law, the zones we take into account are called "zone urbanistiche" and constitute a refinement of the subdivision of the urban area into municipalities. In our model, we decided to further subdivide some bigger zones into smaller parts, in order to keep them as homogeneous as possible. This subdivision has been made with the aid of planners; see Arcirdiacono and Bagnasco [2006].

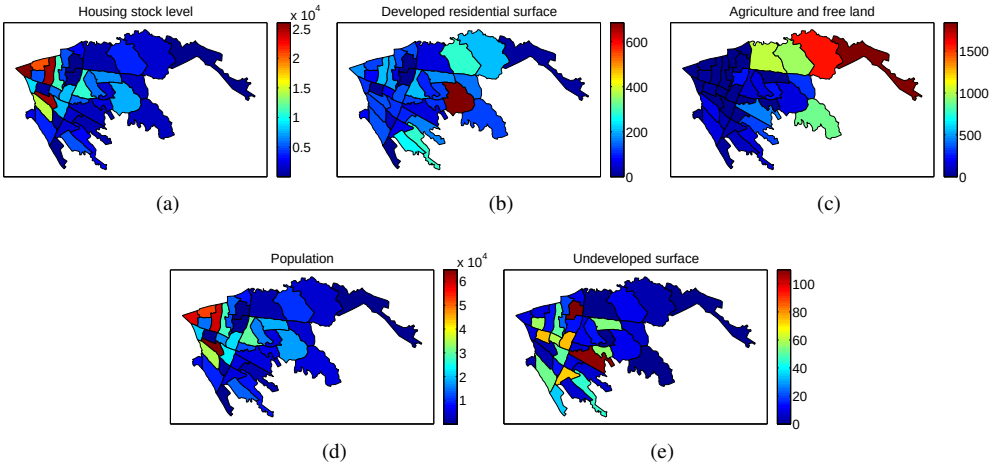
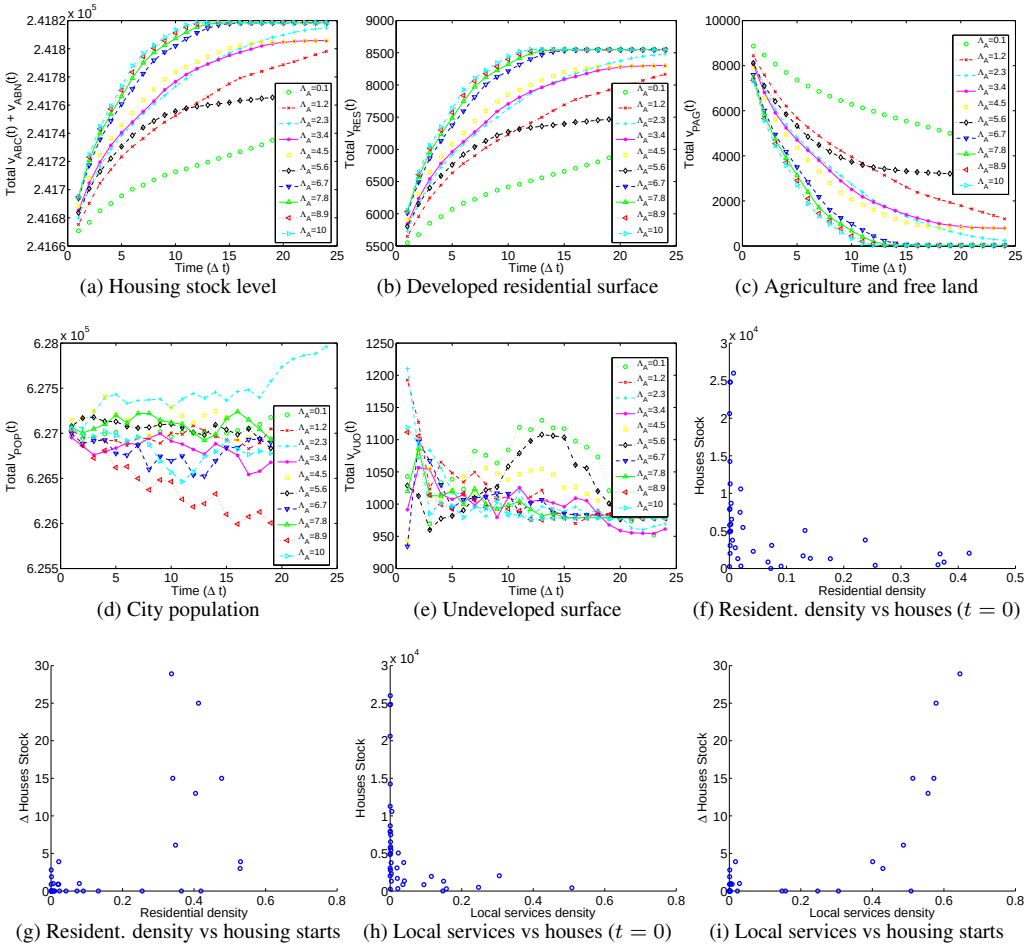


Figure 2: (a)–(e): Initial configuration of the CA. In each figure, the interval between the minimum measurement of the variable and the maximum is mapped on a colormap, such that the relative distribution of values on the cells is portrayed in terms of the difference between "warm" and "cool" colors. (a): Initial level of the housing stock. (b): Initial amount of residential surface. (c): Initial amount of surface due to agriculture or free land. (d): Initial resident population. (e): Initial amount of undeveloped or interstial surface.

is described by $d = 21 + 6$ demographic and economic variables. Figure 2 shows the initial configuration of the system, taking into account the subset of the dynamical variables in $\boldsymbol{v}$ related to the phenomena of residential growth. We performed multiple simulations to see if the model was able to show a plausible behaviour with respect to the phenomena of residential growth of the area under study. In these simulations we set $\Lambda_A^\alpha(t) = \Lambda_A^\alpha(0)$ for every t and for every $\alpha \in \mathcal{A}$, so we expect to see a stationary dynamics for the variables of the configuration of the system, a condition easily checkable by inspecting that the trajectory of $G^\alpha(c,t)$, after a transient growth, reaches a steady level. The parameters $\Lambda_L^\alpha, \Lambda_D^\alpha, \Lambda_U^\alpha$ are set to constant values so that agents from any population have, on the average, three jumps to explore the CA before passing to the inactive state. The overall rate of events occurring during a simulation is controlled by setting $\Lambda_A = \sum_\alpha \Lambda_A^\alpha$ to a constant value, and then taking fixed ratios to define the global intensity of activation for each population of agents: e.g. 20% of all activations are from the population of agents looking for a house, etc. We varied Λ_A from 0.1 to 10, and for each value we executed multiple simulations and averaged $\boldsymbol{v}(c,t)$ over the simulations. We deem this strategy for the determination of the activation rates to be reasonable, since the city is far from being in a period of expansion, and thus the balance between the rates of activity of the different processes is unlikely to change significantly. On the other hand, by varying Λ_A we can explore the parameter space in a consistent way and test the dynamics of the systems at different levels of activity. Figure 3 shows some of these results. The quantities in 3(a)–(e) are aggregated over all cells of the city. These plots clearly show, with the exception for the population, phenomena of saturated growth or consumption of the plotted variables. Moreover, as Λ_A grows it is possible to see a clear convergence to a stable trajectory. For (a), (b) and (c) we computed the confidence intervals at 95% probability for the parameters of a logistic growth / consumption model: while for the parameter corresponding to the saturation threshold we get $< 1\%$ error for $\Lambda_A \geq 3.4$, the best we can do to estimate the intercept of the logistic is an error of 12% for $\Lambda_A = 6.7$. This doesn't surprise us, since the city is already at a late state of growth and thus we cannot hope to give a good estimate for a parameter that is meaningful for an earlier period. The scatter plots (f)–(i) instead give substantial informations about the patterns of residential expansions occurred. They refer to $\Lambda_A = 10$, and each point is a cell. It seems clear that new houses are built in zones with a high residential density or in zones with a substantial presence of local infrastructures. Since we deal with Poissonian diffusion processes, agents explore a portion of the urban space, on average, only 3 jumps deep, starting from the most attractive cells. Thus the patterns expressed in (f)–(i)



(a) Housing stock level (b) Developed residential surface (c) Agriculture and free land

(d) City population (e) Undeveloped surface (f) Resident. density vs houses ($t = 0$)

(g) Resident. density vs housing starts (h) Local services vs houses ($t = 0$) (i) Local services vs housing starts

Figure 3: (a)–(e): comparisons of time series of global variables; different curves in each plot correspond to different values of the parameter Λ_A, the global intensity of activation of agents. (f)–(i): scatter plots of land use densities for residential building and local infrastructures versus initial stock level and new housing starts; in these graphs Λ_A is taken to be equal to 10.

give good evidence that PCA, on which the urban and regional forces of attraction are defined, provides a good model of residential centrality.

4 INTEGRATION FOR ENVIRONMENTAL IMPACT ASSESSMENT

It is possible to simulate the dynamical behaviour of the MAS with an asynchronous algorithm that takes advantage of some properties of Poisson processes, namely that the time between two jumps is exponentially distributed, and that the sum of independent processes is still a Poisson process (a sketch of the algorithm is given in Vancheri et al. [2008]). Since the assumption of independence holds only for a time step of length Δt, we recover an evolution schema familiar to that of a CA. This means that the model could be integrated very easily in an environmental decision support system (EDSS), even as a simple routine call. At each time step, either the configuration of the system may be updated by running the PCA, and thus producing new values of the force of attraction F and G, or the value of the dynamical state $\boldsymbol{v}(c, t)$ of each cell c can be retrieved as an output, (or both). Moreover, interaction issues with other subsystems – such as the biophysical component of an EDSS – arising from a different time scale of simulation, can be reduced by decoupling those two operations. Integration at the data level can be done in a straightforward manner if the model has to produce only output values to be fed into other components. In this case it is worth to note that we explicitly model the dynamics of the population of the system (see figure 3(d)), as well as the number of workers in local facilities of each cell (not shown in the

figures). This just to name a few examples. These can become the inputs to compute measurable indicators, and thus evaluate the significance of the growth of the city on the environment. A more pondered approach - instead - should be taken if data integration contemplates the possibility to feed variables describing the status of the environment into our model. A very simple operation would be to change (9) to allow a sum on more variables than those in V_α, and thus it would be only needed to apply PCA to an augmented data matrix $\boldsymbol{X}(t)$ with a new column for each environmental indicator one would like to consider. As an example, for the event of buying a house, agents could then take into account also informations about air and noise pollution in a cell.

5 CONCLUSIONS

The results we have discussed in this paper stem from a calibration that makes several highly idealized assumptions, namely the choice of the values for $\Lambda_D^\alpha, \Lambda_L^\alpha, \Lambda_U^\alpha$, which result in homogeneous decision process of the agents with respect to the kinds of interaction α, and the dynamics of the global rate of activation Λ_A, which lead to a stationary dynamics of the system, which is motivated by the fact that the city in our case study is already in a late stage of development. Nonetheless, our model was able to show a meaningful urban dynamics of growth, with key factors such as saturation due to depletion of resources (see figures 3(a), (b), (c) and (e)), and realistic decision processes of urban agents. We thus believe that this model has the potential to produce quantitatively precise scenarios of growth – given enough data to perform a proper calibration. A shortcoming of our approach is the fact that PCA does not define a probability density model, and thus doesn't allow one to compute the likelihood of data, which can be problematic for the calibration of the forces of attraction F^α and G^α. We will have to reflect about a proper calibration technique for this part of the model. Some benefits of our approach are: *(i)* the dynamics of our MAS are mathematically well defined so that one could study effects due to bifurcations or phase transitions analytically, and not only through simulation, *(ii)* the asynchronous algorithm we use can simulate many thousands of agents at once, *(iii)* this approach doesn't have to create coherent identities for various social classes of interest, *(iv)* update rules of the CA have a clear urban meaning and are built up from simple events (e.g. building a house), *(v)* we generalize a CA to have a real valued multidimensional state space, and this allows a very easy integration in an EDSS and *(vi)* the cells of the CA correspond to real administrative zones which would make it easier to policy makers and other stakeholders to assess the model's outputs.

ACKNOWLEDGEMENTS

We would like to thank Carlo Bagnasco and Ilia Arcidiacono for having provided us the data from their study on Rome's eastern area (see Arcirdiacono and Bagnasco [2006], in Italian) and for many insightful discussions.

REFERENCES

Arcirdiacono, I. and C. Bagnasco. *L'analisi statistica. Aspetti teorici e applicazioni alla lettura del territorio*. Aracne, 2006.

Batty, M. and B. Jiang. Multi-agent simulation: New approaches to exploring space-time dynamics within GIS. *CASA working paper*, 1999.

Benenson, I. Multi-agent simulations of residential dynamics in the city. *Computers, Environment and Urban Systems*, 22(1):25–42, 1998.

Engelen, G., R. White, and T. Nijs. Environment explorer: Spatial support system for the integrated assessment of socio-economic and environmental policies in the netherlands. *Integrated Assessment*, 4(2):97–105, June 2003.

Everitt, B. S. and G. Dunn. *Applied Multivariate Data Analysis*. Hodder Arnold, February 2001.

Kent, J. T., J. M. Bibby, and K. V. Mardia. *Multivariate Analysis (Probability and Mathematical Statistics)*. Elsevier Limited, 2006.

Kingman, J. F. C. *Poisson Processes (Oxford Studies in Probability)*. Oxford University Press, USA, 1993.

Roweis, S. T. and Z. Ghahramani. A unifying review of linear gaussian models. *Neural Computation*, 11(2):305–345, 1999.

Vancheri, A., P. Giordano, D. Andrey, and S. Albeverio. Continuous valued cellular automata and decision process of agents for urban dynamics. In *Computers in Urban Planning and Urban Management (CUPUM 2005)*, 2005.

Vancheri, A., P. Giordano, D. Andrey, and S. Albeverio. A model for urban growth processes with continuous state cellular automata, multi-agents and related differential equations. Part 1: Theory. *Environment and Planning B: Planning and Design*, 2008. (Submitted).

White, R. and G. Engelen. High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Computers, Environment and Urban Systems*, 24:383–400, 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

Xplorah, A multi-scale integrated land use model

H. van Delden[a], E.R. Gutiérrez[b], J. van Vliet[a] and J. Hurkens[a]

[a] Research Institute for Knowledge Systems, P.O. Box 463, 6200 AL Maastricht, The Netherlands (hvdelden/jvanvliet/jhurkens@riks.nl)
[b] Graduate School of Planning, University of Puerto Rico, P.O. Box 23354, San Juan, Puerto Rico (eliasgutierrez@yahoo.com)

Abstract: Processes of land use change and their drivers take place at different spatial and temporal scales. The fact that these processes are very often interacting with each other throughout these scales provides major challenges to modellers. Xplorah is a tailor made Spatial Decision Support System for the three major islands of Puerto Rico. It aims to assess the impact of different scenarios on the development of the island. The system encompasses an integrated set of dynamically linked models working at different scales and incorporating knowledge from numerous disciplines. This paper describes the integrated model as well as its individual components. It focuses on the integration between the socio-economic and bio-physical processes and models, and discusses how the system can be used to assess the impact of different policies relevant to and defined by the user of the system, Puerto Rico's Planning Board.

Keywords: Spatial Decision Support System; Impact Assessment; Land use modelling, Model integration, Integrated spatial planning.

1. INTRODUCTION

Puerto Rico is a very special island because of its vibrant cultural life, the architecture of its towns, the remnants of its past, and its natural marvels: coral reefs, beaches, caves, and tropical forests. With a population just below 4 million living on an area less than 10,000 km^2, it is also a densely populated place (Engelen, 2003).
Puerto Rico has many of the generic problems that islands, and small islands in particular, experience: an open and relatively small economy, a rugged landscape with a concentration of activities in the coastal zone, fragmentation and loss of high quality land, pressure on the coastal wetlands, deforestation, flooding, pollution, scarcity of drinking water, etc. Moreover, the island is located on the path of many tropical storms entering the Caribbean seas. Some of these gain hurricane force and cause tremendous damage. If global warming causes an increase in both the intensity and the frequency of tropical storms, Puerto Rico should be prepared and implement integrated spatial planning practices to minimise the potential damages (Engelen, 2003).

The Xplorah Spatial Decision Support System (SDSS) has been developed with the aim to support integrated decision making on the island of Puerto Rico. It allows the user to explore the impact of a wide range of scenarios –consisting of a combination of external factors and policy options– on policy-relevant indicators by simulating future developments in the region over a time span of 20 to 30 years. It comprises an integrated model in which the processes at stake on the island are incorporated.
The main aim of Xplorah is to assist policy makers in (1) understanding the important processes in the region and their interaction, (2) indicate current or future problems on the island, (3) assess the impact of possible policy measures, (4) evaluate the different alternatives and (5) stimulate discussion and improve communication between the different actors involved in the decision-making process.

Rather than focusing on all elements that are important in the initialisation, design, development and implementation process of the Xplorah SDSS, we limit ourselves in this paper to describing the different elements of the system, their interaction and the way it can be used for impact assessment studies. More insight in the development process of integrated spatial decision support systems as well as their practical applications and use can be obtained from Hurkens et al. (2008), Van Delden et al. (2007), Rutledge et al. (2007) and Van Delden and Engelen (2006).

2. AN INTEGRATED MODEL

Throughout the design and development of the Xplorah SDSS integration has always played a crucial role: integration between disciplines, integration between scales and integration between (scientific) knowledge, information technology and policy-making.

Where sectoral models often focus on the detailed representation of the processes at stake, Xplorah aims to create dynamic feedback loops between the different sectoral models included. For this reason it uses, to the extent possible, existing disciplinary models. For the integration Xplorah makes use of the GEONAMICA software environment that is specifically designed for the development of spatial decision support systems that integrate a number of non-spatial and spatially explicit models (Hurkens et al., 2008).

Due to the non-linear character of the model and the high degree of linkage among the variables, the set of equations cannot be solved analytically; rather their simultaneous solution is computed numerically in discrete time steps. The temporal resolution of the system is a year, its temporal horizon 2030.

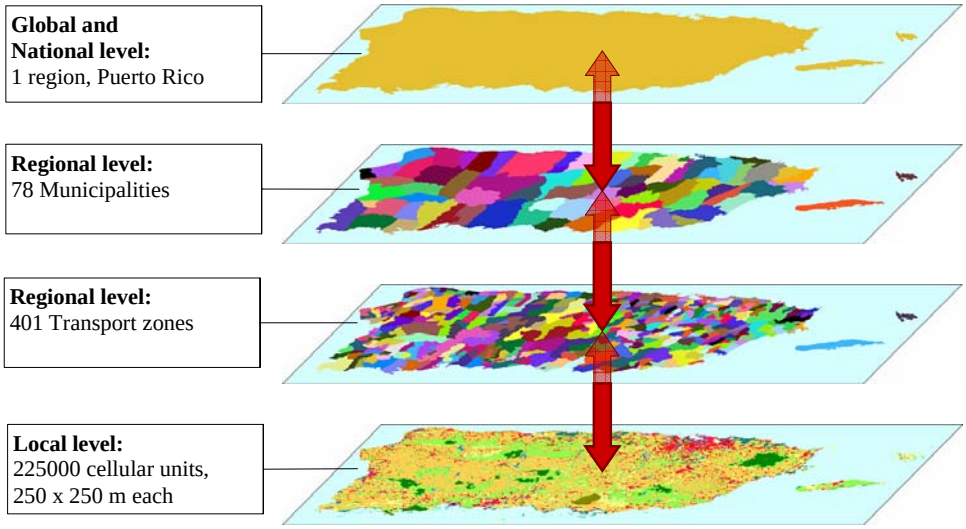


Figure 1. Different interlinked spatial scales in the Xplorah SDSS.

The models that are incorporated in Xplorah simulate activities that take place at four spatial scales: global, national, regional and local. At the global level, climate change has an important impact on the national economy by influencing tourism, agriculture and the demolishment and reconstruction of buildings as a consequence of hurricanes and changes in rainfall and temperature.

A national macro-economic model is tied with an age-cohort model that simulates structural demographic changes and population levels. This model incorporates immigration patterns and provides the labour force supply. Economic conditions, in turn, have an impact on migration and mortality rates.

At the regional level, socio-economic changes take place based on the relative attractivity of regions and the costs required to travel from one region to another. These costs are provided by the transport model that uses information from the regional and local models to

generate trips. This provides the basis for the distribution of national growth as well as migration of jobs and people over regions.
Furthermore, on the local level, land use demands from the regional model are allocated in cells based on several elements including local accessibility, physical suitability, zoning regulations and the attraction and repulsion between different land use functions. The local bio-physical and socio-economic characteristics, finally, feed back into the attractivity at the regional level.
The different components and their interactions are schematised in the system diagram of the Xplorah SDSS in Figure 2. The processes modelled in the components are described in the paragraphs below.

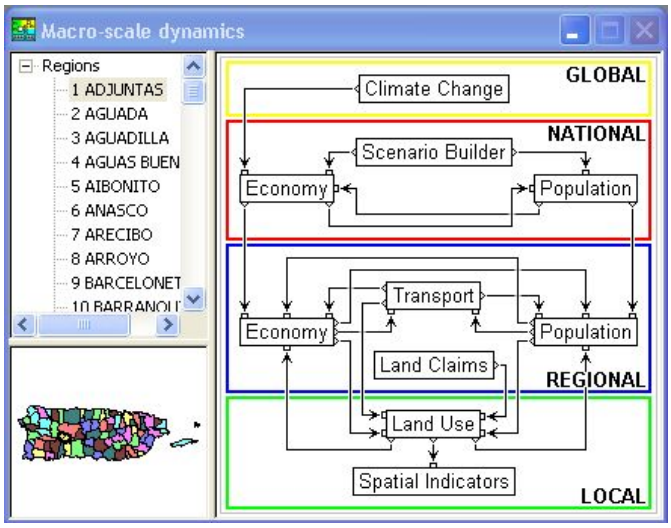


Figure 2. System diagram of the Xplorah SDSS.

2. 1 Models at global and national level

At the highest levels of the model there are components representing climate change, macro-economics and demographics. We have incorporated the impact of (global) climate change as a set of relations that impact different sectors of the economy on the island. At national level models to describe the macro-economics and the demographics are closely linked through mutual feedback loops. Part of the output of these models is used as a driver for the regional models: national population and jobs in main economic sectors.

The climate sub-system consists of a set of linked relations expressing the change in time of temperature and its effects on precipitation, storm frequency and, secondarily, external demands for services and products from Puerto Rico. This component is not a true model, rather a set of linked hypotheses representing current knowledge and assumptions, which the user is free to change. Clearly, this component is a very strong simplification of reality. However, and in the absence of more elaborate models, this representation has the great advantage of enabling to test hypotheses in a relatively straightforward manner.

The economy of Puerto Rico is modelled by means of a macro-economic model (Gutiérrez, 2007) linked to an input-output model, forecasting all final demand components and employment by sectors. In combination these models describe the economic situation on the island over time and also show the impact of changes on the main sectors of the economy –agriculture, construction & mining, manufacturing, services and government– in various levels of detail. Growth in the economic sectors is the result of changes in the final demands for goods of these sectors. On a yearly basis, the final demands of the input-output model change in response to information generated by the macro-economic model: aggregate demand changes in response to an increased demand for consumption driven by changes in population figures, the population's buying power, public and private investment, and exports. Puerto Rico is a very open economy. Thus, changes do take place as a result of external and policy drivers captured in scenarios. Exports change in response

to changes in the US economy and other markets. Furthermore, output from the climate sub-system can influence certain sectors, such as tourism, agriculture and construction.
In turn, the outputs of the macro-economic system interact with demographic trends. Jointly, these forces influence changes in sectoral production. The latter affects the regional activity (jobs in the main economic sectors and population) as well as the demands for land at the local, cellular level.

An age-cohort model generates population figures at yearly intervals by sex and age, on the basis of births, deaths and net migration. Fertility, mortality, and net migration rates are each assumed to be characterized by trends. In addition, it includes features that capture the effects of changing economic conditions (as calculated by the macro-economic system). While the structural component represents the long-term underlying demographic trends, the economic component captures changes that may weigh in on demography due to changes in the wealth of the population. The economic component is particularly important in the case of migration, since the flow of migrants generally responds to changes in relative economic conditions. Initial birth, mortality, and migration rates are taken from US census statistics; number of births and deaths are provided by the Commonwealth Health Department. Migration rates are inferred from census data. All of these rates can be explored in their consequences on land uses, through scenarios entered by the user of the model.

2. 2 Models at regional level

At regional level three model components are incorporated: a spatial interaction model that distributes national totals for population and jobs over the municipalities, a transport model that represents the travel behaviour and calculates the main transport flows and distances between the different regions and a land use demand component that converts the regional levels of activity in different socio-economic sectors, expressed in number of jobs or people, to land use demands for the local model.
The first component of the regional model in Xplorah uses the 78 municipalities of the island as regions and provides information on population as well as the five main economic sectors. The migration of activity between regions uses the concept of a gravity or spatial interaction model (Fotheringham & O'Kelly, 1989). From the levels of activity in each region and the migration flows, we can derive the demand for activity in each region in the next time step. Moreover, the demand for activity also incorporates other factors, such as the national growth of activity, which is calculated by the national demographic and economics models. If we incorporate these three factors, the demand for activity can be derived from the inert activity in that region, the activity that migrates to that region and the national growth that ends up in that region.
The measure of distance that we use in the spatial interaction model is very important. In fact, it is the determining factor of this model for the accuracy with which migration flows can be modelled. We know from experience and literature –see for example Fotheringham, et al. (2000)– that Euclidian distance, or distance measures derived from this, are a relatively inappropriate measure of distance for modelling population migration. Modelled travel times provide much better results. For this reason we have incorporated a transport model that can calculate travel times based on activities, travel behaviour and the road network.

The transport model incorporated in Xplorah is based on a classical four step approach: production-attraction, distribution, modal split and assignment. Like the spatial-interaction model it works at regional level, although its regions are transportation zones. They are generally smaller than municipalities and their sizes depend on the activity that can be found within the zone. Since urban and rural areas have different characteristics and behaviour regarding transport, a classification of the urbanization level of the different transport zones is used throughout the model.
In its first step the model calculates the production of trips from each zone to each other zone based on the activity levels in each zone and the travel behaviour of different groups in society. This generation of trips is not uniform in space but varies with the degree of urbanisation. Over time the number of trips per unit of activity can also change, as the general level of mobility changes. Furthermore, the model accounts for car sharing and

cargo transport, where the first decreases the impact of trips on the network and the latter increases it since trucks use more space than regular cars.
In the distribution step, production and attraction levels in different zones will be linked together. In other words, it is decided which pairs of origins and destinations form trips. Actors select their destinations and their mode of transport as a function of the associated generalised costs. It takes time however for actors to change their behaviour and preferences. Therefore a major factor determining the selection of destinations and modes is the existing transport pattern. It is here where we have made a modification to the classical equilibrium based transport model in favour of a more dynamic approach. The result of the distribution step is an origin-destination matrix (OD-matrix) specifying trips for each transport motive.
In the modal split it is decided which trips are made by car or public transport based on the generalised costs of each decision. These generalised costs include travel distance, travel time, parking costs, toll costs and the aversion for one or the other mode. The latter includes the preference a lot a people have to go by car because of the freedom that it gives them. The travel time component is heavily determined by the congestion on the network and therefore this step requires a strong interaction with the assignment step.
In the assignment step, actors make decisions based on the same generalised costs as in the modal split step. This time however they aim to find the cheapest route from A to B and therefore need information from the transport network. Their behaviour and route selection has a direct impact on the intensity and congestion of the road network and might influence others to take a different route. The result of the assignment step is the allocation of all trips over the road network and the calculation of the intensity, congestion and travel speed on each road element.

We have stated above that the levels of activity form a restriction on the land use demands of the local model. To be more precise, the demand for activity is converted to a number of cells that needs to be allocated in a region. Cell-productivity expresses the amount of activity in a sector that is located in one cell. The average cell-productivity in each region is modelled differently for economic and population sectors and for natural sectors. In the latter case, the activity is defined in terms of surface area, so the productivity is equal to the surface area of one cell. In the former case, the activity is expressed in terms of the number of jobs or people. For these sectors the average cell-productivity in a region will rise as the total level of activity in that sector rises, even if there is enough space for the sector to expand. If the amount of space available is also a critical factor, the average cell-productivity will rise even more. We term this the crowding effect.
We can distinguish two growth factors that can cause a rise in the average cell-productivity in a sector and region, complemented by the crowding effect and a growth coefficient. These five factors are:

- The growth in the level of activity. This is defined as the ratio of the demand for activity and the current level of activity.
- Changing conditions at the local level related to the physical suitability of locations, the zoning regulations, the accessibility and the neighbourhood effect.

The number of cells that needs to be allocated to each sector in each region is equal to the ratio of the demand for activity and the average cell-productivity.

2. 3 Models at local level

At the local level, the main island of Puerto Rico is subdivided in 225,000 cellular parcels of 6.25 ha each (250 x 250 m cells). Each cell represents the dominant land use of the cell. In Xplorah, 19 land use categories are incorporated of which residential, industry, trade & services, agriculture, forest and natural are the most important.
A cellular automaton based land use model is used to determine the state of a cell within the overall growth for each of the 78 municipalities calculated by the regional model (White and Engelen, 1993, 1997). Changes in land use at the local level are driven by four important factors that determine the potential for each location for each actor (see also Figure 3):

1. Physical suitability, represented by one map per land use function modelled. The term suitability is used here to describe the degree to which a cell is fit to support a particular land use function and its associated economic or residential activity.

2. Zoning or institutional suitability, represented by one map per land use function modelled. For different planning periods the map specifies which cells can and cannot be taken in by the particular land use.
3. The accessibility, represented by one map per land use function modelled. Accessibility is an expression of the ease with which an activity can fulfil its needs for transportation and mobility in a particular cell, based on the transportation system.
4. Dynamic impact of land uses in the area immediately surrounding a location. For each land use function, a set of spatial interaction rules determines the degree to which it is attracted to, or repelled by, the other functions present in its surroundings; a 196 cell neighbourhood.

If the potential is high enough, the function will try to occupy the location, if not, it will look for more attractive places. New activities and land uses invading a neighbourhood over time will thus change its attractiveness for activities already present and others searching for space. This process constitutes the highly non-linear character of this model.

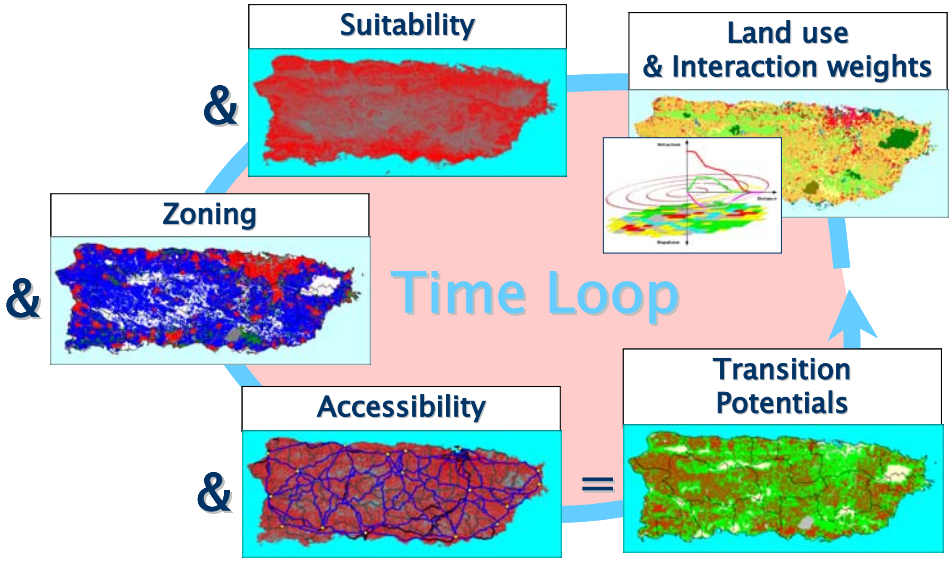


Figure 3. Schematic representation of the model at local level.

3. PROVIDING SUPPORT TO IMPACT ASSESSMENT

For any system to provide support to impact assessment it has to provide the possibility to enter the crucial driving forces on the input side and show the impacts on relevant indicators on the output side. The Xplorah SDSS is designed in such a way that scenarios can be set up that include a combination of external factors and policy options, which can be entered or changed by the user. External factors included in Xplorah are amongst others climate change; macro-economic drivers like interest rates, oil prices, growth rates in the US and money transfers from the US to Puerto Rico; and demographic drivers like fertility, migration and mortality rates. Xplorah also includes a wide range of policy options such as government consumption, public investment, development of public housing units, road prices, zoning regulations and construction of infrastructure.

To provide support to the decision making process the system is equipped with several policy relevant indicators. An indicator in this context is a measure to make a particular phenomenon perceptible that is not –or at least not immediately– detectable. On the other hand, indicators can also be set up to verify legislative guidelines or policy goals. Indicators are in fact small sub-models that generally simplify model results in order to make complex phenomena quantifiable in such a manner that communication is either enabled or promoted. Indicators in Xplorah are provided as (non spatial) numerical values and as maps with a resolution similar to that of the local model (250 x 250 m). Examples of numerical

indicators are GDP in main economic sectors, proportion of population employed, total population, number of jobs per municipality and housing stock. Examples of spatial indicators are urban clusters, distance from residential locations to work locations, flood damage and habitat fragmentation.

By running the model under different scenarios, future developments of the island can be explored and evaluated using a number of selected indicators. Different alternatives can either be compared against each other, or against a 'baseline scenario' that assumes similar behaviour in the future as in the past. Furthermore also the temporal developments within a scenario can be investigated and compared to initial conditions or any other selected point in time.

4. CONCLUSIONS AND RECOMMENDATIONS

Real world problems and processes are not limited to a specific sector or discipline, but are interacting with each other throughout different domains. Models attempting to represent the real-world system should therefore incorporate the different disciplines and their dynamic feedback loops. To be able to represent real-world phenomena in their true complexity, the Xplorah SDSS has been equipped with a fully dynamic model that integrates climate, economy, demography, transport and land use.

Spatial Decision Support Systems that comprise an integrated model are able to assess the impact of (sectoral) policy options on a wide range of indicators. Xplorah is equipped with a number of policy options and its strong feedback loops between the different sectors allow the user to assess the impact of a policy option not only on his or her own discipline, but also on the other disciplines incorporated. Understanding those impacts can prevent the occurrence of unexpected and unwanted side-effects after the implementation of new policies.

In the development of an SDSS it is important to create links between user-relevant drivers and model inputs as well as between model outputs and user-relevant indicators. Moreover, the user should have the possibility to set up a scenario that consists of (a combination of) external factors and policy options, since this facilitates the creation of different scenario alternatives and the assessment of their impacts.

Xplorah has been developed in an iterative manner. Initially it started with the local land use model and over time a range of components has been integrated into the system. During such a development process it is very important to evaluate the system against its required capabilities for impact assessment. At present, the main points for attention are:

- The climate component incorporated in the system consists of simple and direct relations between a few variables and is as such a very strong simplification of reality. In more elaborate versions of the model, the relations used could be replaced by dedicated sub-models describing the underlying processes at deeper levels.
  Moreover, the natural sub-system is represented in a very limited manner as a measure of the physical suitability of the land for receiving the activities included in the model. The climate sub-system could be complemented with a natural sub-system including models describing the dynamics of natural systems such as forests, coral reefs, mangroves, river systems, beaches, etc.
- The input-output model describes the economy of the island as a set of linear equations. This approach has the advantage of ensuring that the output of the economic model is internally consistent. In addition, it captures the interdependencies between economic sectors and thus represents well the multiplier effects by which changes in one sector propagate through the entire economy. On the other hand, the method is based on an assumption of constant technical coefficients, so that changes in the underlying structure of the economy, due to, for example, technological innovations, import substitution, or factor substitutions in response to changes in relative prices, are not represented. This is not a problem for short run situations, but becomes much more of a concern when the modelling period runs to several decades (which is the case in the present

model). Ideally, in the context of the present model, the technical coefficients would evolve in response to substitution and changes in productivity generated by the local model.

At present, the Xplorah system is being implemented in Puerto Rico's Planning Board. A group of around 50 staff members, ranging from technicians and middle managers to executives, are following an extensive training course to become expert users in the system. During this implementation phase the system is also evaluated on its capability to provide support to actual planning problems on the island. In this process it will be a challenge to create awareness not only on the capabilities of the system, but also on its limitations. The system helps to understand how different processes are related and explores the impact of different external factors and policy options. Xplorah, like any model, is however a simplification of reality and will not give exact predictions of the future. Its strength is in *supporting* learning, analysis and communication, not in *making decisions*. The responsibility of the latter lies with the decision-makers.
We expect that the current phase will provide some important recommendations for the further development of the system to improve its usefulness as well as its usability.

REFERENCES

Engelen, G., Xplorah, The Spatial Decision Support System for Puerto Rico. Brochure, 2003.

Fotheringham, A.S. and M.E. O'Kelly, Spatial interaction models. Dordrecht, The Netherlands, Kluwer, 1989.

Fotheringham, A.S., T. Champion, C. Wymer, M. Coombes, Measuring Destination Attractivity: A Migration Example. International Journal of Population Geography 6: 391-421, John Wiley & Sons, 2000.

Gutiérrez, E.R., Long-term View of the Commonwealth of Puerto Rico's Socio-Economic Environment: Planning Context for Corporate Decisions. Draft for discussion, 2007.

Hurkens, J., B.M. Hahn and H. van Delden (this conference). Using the Geonamica® software environment for integrated dynamic spatial modelling. In: M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.) Proceedings of the iEMSs Fourth Biennial Meeting: "Integrating Sciences and Information Technology for Environmental Assessment and Decision Making". International Environmental Modelling and Software Society, Barcelona, Spain, 2008.

Rutledge, D., McDonald, G., Cameron, M., McBride, G., Poot, J., Scrimgeour, F., Price, R., Phyn, D., van Delden, H., Huser, B., Small, B., Wedderburn, L. & Fenton, T., Development of Spatial Decision Support Systems to Support Long-term, Integrated Planning. In: Oxley, L. and Kulasiri, D. (eds) MODSIM 2007 International Congress on Modelling and Simulation, volume December 2007, p.308-314, Modelling and Simulation Society of Australia and New Zealand, ISBN 978-0-9758400-4-7

Van Delden, H., P. Luja and G. Engelen, Integration of multi-scale dynamic spatial models of socio-economic and physical processes for river basin management. Environmental Modelling and Software, 22(2), 223-238, 2007.

Van Delden, H. and G. Engelen, Combining participatory approaches and modelling: lessons from two practical cases of policy support. In: Voinov, A., Jakeman, A.J., Rizzoli, A.E. (eds). Proceedings of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software". International Environmental Modelling and Software Society, Burlington, USA, July 2006.

White, R. and G. Engelen, Cellular Automata and Fractal Urban Form: a Cellular Modelling Approach to the Evolution of Urban Land use Patterns. Environment and Planning A, 25, 1175-1199, 1993.

White R. and G. Engelen, Cellular Automata as the Basis of Integrated Dynamic Regional Modelling. Environment and Planning B, 24, 235-246, 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

Multi-Agent Models for Teaching, Extension and Collaborative Learning

Thomas Berger and Pepijn Schreinemachers

Universität Hohenheim (490d),Stuttgart, Germany, (i490d@uni-hohenheim.de)

Abstract: This paper describes an agent-based land use modeling approach, called MP-MAS, developed at Hohenheim University. The main focus of this approach is the integration of economic decision-models with biophysical models of water supply and soil fertility at a fine spatial resolution. Short-term production and consumption decisions of agents are represented as mathematical programming problems, whereas longer-term decisions, for example investment and migration, are represented using heuristics. Here, we position the approach in relation to alternative agent-based models of land use and water management and describe empirical applications to Chile, Uganda, Ghana, and Thailand. Based on these practical experiences, we discuss the use of MP-MAS as a tool for collaborative learning and participatory research.

Keywords: *Natural Resource Use; Participatory Simulation; Mathematical Programming*

1. INTRODUCTION

Multi-Agent Systems (MAS) are increasingly used as a tool to disentangle and explore the complex relationships between land use including water resources, policy interventions and human adaptation. The development and application of these tools has been made possible by the rapid increase in computational power available at modest cost. The strength of agent-based land use models lies in their ability to combine spatial modeling techniques, such as cellular automata or geographical information systems (GIS), with biophysical and socioeconomic models at a fine resolution.

MAS are flexible in their representation of human land use decisions and therefore appeal to scholars from diverse backgrounds, such as sociology, geography, and economics [Schreinemachers and Berger, 2006]. The behavior of individual actors can be modeled one-to-one with computational agents which allows for direct observation and interpretation of simulation results. Large part of their fascination—especially to scholars who are otherwise skeptical of any attempt to quantify and model human behavior—rests on this intuitive and potentially interactive feature. Scholars combine MAS with role-playing games in which a group of resource users, typically farmers using some common-pool resource, specify the decision rules of computational agents and observe how these rules might affect both people's well-being and their natural resource base [Bousquet et al., 2001; D'Aquino et al., 2003; Becu et al., 2003].

In this paper we reflect on the interactive use of multi-agent models not only for participatory simulation of land-use change but also for teaching, extension and collaborative learning in general. At Hohenheim University, we use our MP-MAS software for teaching at M.Sc. and Ph.D. levels, teach training courses for water resource managers in Chile and parameterized the MP-MAS model for empirical applications in Thailand, Uganda, Chile and Ghana [Berger, 2001; Berger et al., 2006; Schreinemachers 2006; Berger et al., 2007]. MP-MAS distinguishes itself most clearly from other agent-based land use models in its use of a constrained optimization routine, based on mathematical

programming, for simulating agent decision-making. Apart from describing the rationale behind this modeling approach, this paper reports on various case study applications, and the use of the model for collaborative learning and research.

2. MULTI-AGENT SYSTEMS FOR LAND USE MODELING

Multi-Agent Systems for land use modeling couple a cellular component that represents a landscape with an agent-based component that represents human decision-making [Parker et al., 2002]. Models of this type have been applied in a wide range of settings (for overviews see Janssen 2002, Parker et al., 2003] yet have in common that agents are autonomous decision-makers who interact and communicate and make decisions that can alter the environment. Most multi-agent systems applications have been implemented with software packages such as Cormas, NetLogo, RePast, and Swarm [Railsback et al., 2006].

The philosophy of agent-based modeling has always been to replicate the complexity of human behavior with relatively simple rules of action and interaction. In empirical applications to the complexity of land use changes, the question arises how simple these rules need to be? Most applications have used relatively simple heuristics to represent the economic decision-making of agents. Schreinemachers and Berger [2006] argued that agents in such applications might have too limited heterogeneity and adaptive capacity, and henceforth preferred implementing agents with goal-driven behavior based on mathematical programming.

The use of mathematical programming has a long tradition in agricultural economics [Hazell and Norton, 1986], and the precursors of today's agent-based models – so-called adaptive macro and micro systems – were implemented with mathematical programming [Day and Singh, 1975]. Examples for agent-based land use models using mathematical programming are Balmann [1997] and Happe et al. [2006] who analyzed structural change in German agriculture with software called AgriPoliS. In applications to Chile and Uganda, Berger [2001] and Schreinemachers et al. [2007] developed software called Mathematical Programming-based Multi-agent Systems (MP-MAS), which we will present in the following.

3. THE MP-MAS APPROACH

3.1 Model features

In Berger [2001], Schreinemachers et al. [2007] and Berger et al. [2007] we described in detail the model components, parameters and equations of MP-MAS. Our approach shares many characteristics with bio-economic farm household models (see for example, Ruben et al., 2000]. There are, however, three important additional features that distinguish MP-MAS from the independent, representative farm modeling approach:

- *Number of farm models*: Each real-world farm household is individually represented by a single agent in the model; that is, there is a one-to-one correspondence between real-world households and modeled agents. Monte Carlo techniques have been developed to generate alternative agent populations from random sample surveys [Berger and Schreinemachers, 2006].
- *Spatial dimension*: The MP-MAS model is spatially explicit and employs a cell-based data representation where each grid cell corresponds to one farm plot held by a single landowner. Sub-models of water run-off and crop growth are linked to this cell-based spatial framework.
- *Direct interactions*: Several types of interactions among agents and their environment are explicitly implemented in MP-MAS such as the communication of information, the exchange of land rights and water resources on markets, the return flows of irrigation water, the irrigation of crops, soil nutrient management and crop growth.

This one-to-one MAS representation is able to capture biophysical and socio-economic constraints and interactions at a very fine spatial resolution. Including this heterogeneity of constraints and interactions of farm agents and their biophysical environment broadens the scope of land-use modeling significantly. Phenomena that conventional models cannot easily address—such as local resource degradation, technology diffusion, heterogeneous policy responses and land-use adaptations—can now explicitly be modeled.

MP-MAS is freeware software developed at Hohenheim University and can be downloaded from http://www.uni-hohenheim.de/igm/. A detailed user manual is available from the same website. MP-MAS was written in C++ and is available for both Unix and Windows operating systems. MP-MAS works with a set of input files that are organized in Microsoft Excel workbooks.

3.2 Applications

MP-MAS has been applied to a variety of case studies in Chile [Berger, 2001], Uganda [Schreinemachers et al., 2007], Ghana, and Thailand (Table 1). Other applications to Vietnam and Germany are in the pipeline. Applications to Uganda and Thailand have been small-scale applications at village or micro-catchment level including relatively few agents, while applications to Chile and Ghana have been large-scale applications at the level of watersheds and including thousands of agents.

Table 1. Case studies using MP-MAS

	Application	No. of agents	Spatial dimension		Temporal dimension		Type of agriculture
			extent [km^2]	resolution [m]	duration [years]	time step [days]	
1	Chile, Maule basin	3,592	5,300	100	20	30	Market-oriented and commercial
2	Ghana, White Volta basin	34,691	3,779	100	15	30-365 *	Semi-subsistence
3	Uganda, southeastern	520	12	71	16	365	Semi-subsistence; maize, cassava, bean and plantain
4	Thailand, northern uplands	1,229	140	40	15	30-365 *	Commercial vegetable and fruit production

Note: * Components of the model have different time steps. The decision-making follows an annual sequence while crop water requirements, irrigation water supply, and rainfall are specified on a monthly base.

In all case studies, research questions related to the interaction between the economic and biophysical sub-systems at the farm household level (Table 2). For example, the objective of the Uganda application was to assess the effect of high-yielding maize varieties on soil nutrient dynamics and economic well-being. The agent-based approach gave a detailed assessment of distributional consequences and let to the conclusion that although poverty could be substantially reduced, the incidence ratio of households below the poverty line would still be 20 percent.

3.3 Empirical parameterization

Robinson et al. [2007] compared five empirical methods for building agent-based models in land use science: sample surveys, participant observation, field and laboratory experiments,

companion modeling, and GIS and remotely sensed spatial data. The empirical base of MP-MAS is mostly random sample surveys of farm households and GIS data, both of these are used to define the initial conditions of the model [Berger and Schreinemachers, 2006]. Additional parameters, mostly related to the dynamics of the model, are based on secondary data, qualitative data from field observation, and feedback from stakeholders. For instance, fertility and mortality levels are obtained from statistical agencies, crop yield response from field experiments, while agent interactions can be based on qualitative field observation.

Table 2. Model features of each application

Application	Objective	Economic component	Biophysical component
1. Chile	Provide information to water resource managers (small and large-scale infrastructure projects)	Detailed production functions especially on irrigation methods (agent decision model with 1119 activities, 224 constraints).	Crop growth under water deficits. Spatial distribution of surface water flows.
2. Ghana	Land and water use mostly under rainfed conditions. Test the profitability of irrigated agriculture	Agent decision model contains 752 activities and 250 constraints. Includes a detailed expenditure system	Model simulates the water supply and water distribution with a feedback to crop yields
3. Uganda	To disentangle the relationship between technology adoption, soil nutrients, and poverty levels.	Detailed production functions; 2350 activities, 560 constraints. Includes a detailed expenditure system.	Availability of soil nutrients and organic matter is endogenous and affects crop yields.
4. Thailand	The ex-ante assessment of technology adoption and sustainability strategies.	Based on gross-margin analysis; 53 activities and 60 constraints.	Model simulates the water supply and water distribution with a feedback to crop yields.

4. COLLABORATIVE RESEARCH AND LEARNING

As argued above, one of the key advantages of agent-based modeling is the one-to-one correspondence of real-world and computational agents, which facilitates participatory simulation and model-enhanced learning [e.g. Becu et al., 2007]. Applying agent-based land use models effectively—so that model users receive early warnings, share their system understanding and improve the outcomes of their land-use decisions [Hazell et al., 2001]—poses a number of challenges that have not been fully resolved yet (see for example von Paassen [2004] who reports mixed success for applications of mathematical programming models in developing countries). Based on our first practical experiences of applying multi-agent simulation with stakeholders, we reflect on the following critical issues: (i) participatory techniques for model validation; (ii) building trust in model results; (iii) using MP-MAS for agricultural extension; and (iv) development of teaching and training programs. More formal results will become available soon from our monitoring and evaluation system that we implemented in Ghana and Chile.

4.1 Participatory techniques for model validation

According to our recent experiences in the *CGIAR Challenge Program on Water and Food* (see project website http://www.igm.uni-hohenheim.de), MP-MAS has a clear advantage

over other integrated modeling approaches we have applied before, for example, aggregate regional land use models. Single-agent models for representative farm households can be constructed and validated jointly with stakeholders in interactive model validation rounds. First, we collected farm-specific data on factor endowments such as labor, land and water, and processed these data for a standalone version of MP-MAS. Then we calibrated each single-agent model to replicate current land use decisions and performed sensitivity tests together with stakeholders. Through what-if scenarios, for example, how do you adjust your land use if you receive less irrigation water, we could elucidate additional constraints the farmers actually faced and that were originally not included in the model. The single-agent models were then gradually improved until sufficient model fit for each of the representative farm households was reached. In a second step, the full agent model for the study area was calibrated and validated, using the Monte Carlo approach as described in Berger and Schreinemachers [2006]. In our experience, the use of standardized questionnaires is an efficient way of collecting basic agent data on agricultural land-use. The alternative of stakeholder group interviews, as used by other scholars [see Robinson et al., 2007] is much more time-consuming.

4.2 Building trust in model results

The interactive modeling rounds for parameter testing and model validation can help building trust in the simulation results of MP-MAS. Since farmers and water managers are directly involved in compiling the model database and performing the sensitivity analyses, they become familiar with the model and its interfaces. Results from special model computations, for example of individual water shadow prices, can be compared with local data and experience and create confidence in the model if the results are plausible. Typically, testing and calibrating of MP-MAS requires more than one modeling round and might demand additional time if unforeseen behavioral constraints need to be included. Our impression from applying multi-agent simulation with many feedback rounds is that stakeholders and potential model users are prone to losing interest if these rounds consume much of their time. The interactive modeling rounds should therefore generate information that is perceived as immediately useful by stakeholders. In case of market-oriented farm household such information typically involves estimates of crop yields, farm profitability, and household income; in case of water managers it involves minimum river flows, average water uptake and water use efficiency per irrigation section.

4.3 Using MP-MAS for agricultural extension

Mathematical programming is part of planning methods taught in farm management schools and is used in agricultural extension. Standard farm decision problems such as partial budgeting, investment and income analysis can be directly addressed by the tools incorporated in MP-MAS, making use of the database that has to built up for the model application. Our experience is that workers in farm extension programs can therefore be convinced with relative ease of using the single-farm features of MP-MAS. The practical challenge, however, is the maintenance and adaptation of the MP-MAS input files, which requires some minimum knowledge in database management and MP. To address this challenge, we use the ubiquitous software MS-Excel for input/output operations and have formed a group of advanced model users that are trained in using MP-MAS.

4.4 Development of teaching and training programs

MP-MAS requires, as all other software, teaching and training. We started developing specific programs targeted at various potential user groups, ranging from introductory demonstrations of a few days to a series of workshop sessions held over one year. At Hohenheim University, we offer consecutive courses on *Farm-Level Modeling* and *Land-Use Economics* at MSc level and *Advanced Techniques for Land-Use Modeling* at PhD level. The inclusion of agent-based modeling in the curriculum of the Master Study

Programs *Agricultural Economics* and *Agricultural Sciences in the Tropics and Subtropics* in Hohenheim has added young scientists to the model developer group and increased the number of empirical research applications as part of dissertation projects. Currently, we are planning to develop on-line resources to be inserted in distance learning programs.

5. CONCLUSIONS

MP-MAS is a software for agent-based modeling that through the use of mathematical programming represents goal-driven behavior of farm agents. Biophysical models simulating soil fertility dynamics, water supply, or crop yields have been spatially integrated with agent decision-making through the use of GIS layers. The method is suitable for research questions related to the interaction of economic and biophysical sub-systems and to assess distributional consequences of policy and environmental change. MP-MAS has been applied to case studies in Chile, Uganda, Ghana, and Thailand and valuable experiences have been gained about using MP-MAS in participatory settings. Research is ongoing; the evaluation of the effectiveness of the MP-MAS approach in improving land-use decisions as envisaged in the *CGIAR Challenge Program on Water and Food* is still not completed. Our conclusion is that initial results form using MP-MAS in interactive settings are promising but more methodological research is needed to fine-tune and insert MP-MAS as an effective tool into land-use planning and farm extension programs.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of the *Deutsche Forschungsgemeinschaft (German Research Foundation)* under SFB-564 and the *CGIAR Challenge Program on Water and Food.*

REFERENCES

Allen, R.G., Pereira, L.S., Raes, D., Smith, M. (1998): Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. Available online at: http://www.fao.org/docrep/X0490E/X0490E00.htm

Balmann, A. (1997) Farm-based Modelling of Regional Structural Change: A Cellular Automata Approach. European Review of Agricultural Economics 24: 85-108.

Becu, N., Neef, A., Schreinemachers, P., Sangkapitux, C. (2007) Participatory computer simulation to support collective decision-making: Potential and limits of stakeholder involvement, Land Use Policy (in press).

Becu, N., Perez, P., Walker, B., Barreteau, O., Le Page, C. (2003): Agent-based simulation of a small catchment water management in northern Thailand: Description of the CATCHSCAPE model. Ecological Modelling, 170: 319-331.

Berger, T. (2001) Agent-based models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agricultural Economics, 25: 245-260.

Berger, T., Birner, R., Díaz, J., McCarthy, N., Wittmer, H. (2007): Capturing the Complexity of Water Uses and Water Users within a Multi-Agent Framework. Water Resources Management 21 (1): 129-148.

Berger, T., Schreinemachers, P. (2006) Creating agents and landscapes for multi-agent systems from random samples. Ecology and Society, 11(2), p. Art.19. Available online at: http://www.ecologyandsociety.org/viewissue.php?sf=25.

Berger, T., Schreinemachers, P., Woelcke, J. (2006): Multi-Agent Simulation for Development of Less-Favored Areas. Agricultural Systems, 88: 28–43

Bousquet, F., Lifran, R., Tidball, M., Thoyer, S., Antona. M. (2001): Agent-based modelling, game theory and natural resource management issues. Journal of Artificial Societies and Social Simulation 4(2). Available online at: http://www.soc.surrey.ac.uk/JASSS/4/2/0.html

D'Aquino, P., Le Page, C., Bousquet, F., Bah, A. (2003): Using Self-Designed Role-Playing Games and a Multi-Agent System to Empower a Local Decision-Making Process for Land Use Management: The SelfCormas Experiment in Senegal. In: Journal of Artificial Societies and Social Simulation 6 (3). Available online at: http://jasss.soc.surrey.ac.uk/6/3/5.html

Day, R.H., Singh, I.J. (1975): A Dynamic Model of Regional Agricultural Development. International Regional Science Review 1 (reprint from journal).

Happe, K., Kellermann, K., Balmann, A. (2006) Agent-based analysis of agricultural policies: an Illustration of the Agricultural Policy Simulator AgriPoliS, its adaptation and behavior. Ecology and Society, 11(1), p. 49.

Hazell, P.B.R., Chakravorty, U., Dixon, J., Celis, R. (2001): Monitoring Systems for Managing Natural Resources: Economics, Indicators and Environmental Externalities in a Costa Rican Watershed. EPTD Discussion Paper No. 73. International Food Policy Research Institute. Available online at: http://www.ifpri.org/divs/eptd/dp/papers/eptdp73.pdf

Hazell, PB.R, Norton, R., 1986. Mathematical programming for economic analysis in agriculture. Available online at: http://www.ifpri.org/pubs/otherpubs/mathprog.htm

Janssen, M.A., (Ed., 2002). Complexity and ecosystem management: The theory and practice of multi-agent systems. Cheltenham, U.K. and Northampton, Mass.: Edward Elgar Publishers

Parker, D.C., Berger, T., Manson, S.M. (Eds., 2002): Agent-Based Models of Land Use / Land Cover Change. LUCC Report Series No. 6, Louvain-la-Neuve, 79-88. Available online at: http://www.globallandproject.org/Documents/LUCC_No_6.pdf

Parker, D.C., Manson, S.M., Janssen, M.A., Hoffmann, M.J., Deadman, P. (2003) Multi-agent systems for the simulation of land-use and land-cover change: a review. Annals of the Association of American Geographers, 93(2), pp. 314-337.

Railsback, S., Lytinen, S., Jackson, S. (2006) Agent-based simulation platforms: review and development recommendations. Simulation, 82(9), pp. 609-623.

Robinson, D.T., Brown, D.G., Parker, D.C., Schreinemachers, P., Janssen, M., Huigen, M., Wittmer, H., Gotts, N., Promburom, P., Irwin, E., Berger, T., Gatzweiler, F., Barnaud, C. (2007) Comparison of empirical methods for building agent-based models in land use science. Journal of Land Use Science, 2(1), pp. 31-55.

Ruben, R., Kuyvenhoven, A., Kruseman, G. (2000): Bioeconomic Models and Ecoregional Development: Policy Instruments for Sustainable Intensification. In: D.R. Lee and C.B. Barrett: Tradeoffs or Synergies? Agricultural Intensification, Economic Development and the Environment, CABI Publishing, Wallingford, 2000, 115-133.

Schreinemachers, P. (2006) The (Ir)relavance of the crop yield gap to food security in developing countries. With an application of multi-agent modeling to farming systems in Uganda. Cuvillier Verlag, Göttingen. Available online at: http://hss.ulb.uni-bonn.de/diss_online/landw_fak/2006/schreinemachers_pepijn/.

Schreinemachers, P., Berger, T. (2006) Land-use decisions in developing countries and their representation in multi-agent systems. Journal of Land Use Science, 1(1), pp. 29-44.

Schreinemachers, P., Berger, T., Aune, J.B., 2007. Simulating soil fertility and poverty dynamics in Uganda: A bio-economic multi-agent systems approach, Ecological Economics 64(2): 387-401.

van Paassen, J. M. (2004): Bridging the gap: computer model enhanced learning about natural resource management in Burkina Faso. Dissertation, Wageningen. Available online at: http://www.gcw.nl/dissertations/3530/dis3530.pdf

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

Addressing stakeholder concerns using the Integrative River Rehabilitation Model (IRRM)

M. E. Borsuk[a], S. Schweizer[b], and P. Reichert[c]

[a] *Thayer School of Engineering, Dartmouth College*
Hanover, NH 03755 USA (mark.borsuk@dartmouth.edu)

[b] *Kraftwerke Oberhasli AG (KWO), P.O. Box 63,*
3862 Innertkirchen, Switzerland (sste@kwo.ch)

[c] *Swiss Federal Institute of Aquatic Science and Technology (Eawag),*
P.O. Box 611, 8600 Dübendorf, Switzerland (reichert@eawag.ch)

Abstract: We have previously reported on a variety of modelling methods and decision support concepts that can assist with various aspects of river rehabilitation planning and management. Here, we bring all of these tools together into an Integrative River Rehabilitation Model (IRRM) that links management actions, through morphological and hydraulic changes, to the final ecological and economic consequences. The IRRM is formulated as a probability network and represents the relevant cause-effect relations among important biotic and abiotic factors, leading to attributes (model endpoints) of concern to river system stakeholders. Together with a model of the stakeholders' preference structure for different levels of these attributes, the IRRM is intended to provide a comprehensive basis for supporting river rehabilitation decisions. While many opportunities for further model improvement and uncertainty reduction exist, we believe that the present version of the model provides a flexible framework that can be adapted and refined according to local project-specific needs and data availability. We exemplify model application to three large planned or recently completed rehabilitation projects in Switzerland.

Keywords: Bayesian Network; Uncertainty; Decision Analysis; Stakeholders; Integrated Assessment; Restoration; Morphology and Hydraulics; Benthos; Fish; Economics

1. INTRODUCTION

In recent years, rehabilitation of channelized river systems has become increasingly common, with some countries spending billions of dollars to improve flood protection for adjacent land uses while enhancing ecological condition. Often, rehabilitation involves the creation of localized 'river widenings' in which levees are moved back to allow a more natural channel movement within a limited area [Rohde *et al.*, 2005]. Within the widened reach, the river might shift and adjust, possibly re-establishing the range of riparian habitats that were found prior to channelization.

As rehabilitation becomes more common, integrative modelling tools are essential to help stakeholders understand the morphological, economic, and ecological consequences of the rehabilitation activities. Such predictions can provide the basis for planning and management efforts that attempt to balance diverse interests [Reichert *et al.*, 2007]. In previous publications, we have described a variety of submodels and decision support concepts applicable to river rehabilitation planning and management. Here, we bring all of these tools together in the form of a probability network [Pearl, 1988]. The resulting Integrative River Rehabilitation Model (IRRM) links management actions, through morphological and hydraulic changes, to the final ecological and economic consequences. Together with a preliminary model of the stakeholders' preference structure for different levels of these attributes [Hostmann et al, 2005], the IRRM is intended to provide a comprehensive basis for supporting river rehabilitation decisions.

2. PROBABILITY NETWORKS

Probability (or belief) networks have been used in a variety of settings to compile knowledge from multiple sources to generate probabilistic predictions. A key element in their use is a graphical representation of the causal relationships described by the model. The interesting feature that is made explicit by the graph is the conditional independence implied by the absence of connecting arrows between some nodes. These independences allow the complex network of interactions from primary cause to final effect to be broken down into sets of relations which can each be characterized independently [Pearl, 1988]. This aspect of belief networks significantly facilitates their use for representing multi-disciplinary models such as the IRRM.

Characterization of the relationships in a probability network consists of constructing conditional distributions that reflect the aggregate response of each variable to changes in its immediate "up-arrow" predecessor, together with the uncertainty in that response. It is often convenient to write these conditional relationships in a functional form that includes uncertainty in the model's parameters and an error term capturing unexplained variability. This method of expressing conditional probabilities is consistent with the perspective of most process-based modeling and facilitates computer simulation. Once all relationships in a network are characterized, probabilistic predictions of model endpoints can be generated conditional on values (or distributions) of any "up-arrow" causal variables. These predicted endpoint probabilities, and the relative change in probabilities between decision alternatives, convey the magnitude of expected system response to management while accounting for predictive uncertainties.

3. MODEL DESCRIPTION

3.1 Model Endpoints

A model designed to support rehabilitation management decisions should have endpoints that address the key concerns of system stakeholders. Therefore, our model development started with the identification of river stakeholders and their rehabilitation objectives. Key stakeholder groups include recreational organizations, forest managers, industry representatives, environmental organizations, farmers, local communities, and federal or regional administrations [Hostmann et al, 2005]. A stakeholder elicitation exercise in Switzerland found that the objectives held by these groups could be organized into broad classes related to physical river integrity, chemical water quality, biological integrity, and economic value, including minimization of project cost and maximization of ecosystem services [Hostmann et al, 2005; Reichert *et al.*, 2007]. Some objectives, such as those related to water quality, are usually not strongly impacted by local rehabilitation actions and were therefore not considered further in our project. The remaining objectives were assigned attributes, which are measurable variables that can be used to assess attainment of objectives (Table 1). These attributes, which were understood to represent long-term steady state conditions over a reach scale, were used as endpoints of the predictive model.

Table 1: Key stakeholder objectives and corresponding attributes used as model endpoints.

Category	Objective	Attributes
Physical River Objectives	Natural river morphology	Morphological type (braided, alternating, or straight)
	Natural river hydraulics	Joint distribution of velocity and depth
		Percent area riffles, runs, and pools
		Gravel movement and siltation
Biological Objectives	Abundant benthic organisms	Summer density of periphyton
		Summer density of invertebrates
	Abundant shoreline fauna	Summer density of beetles
		Summer density of spiders
	Abundant fish	Density of salmonids
		Density of cyprinids
Economic Objectives	High flood protection	Estimated flood frequency
	Low project costs	Implementation costs
		Maintenance costs
	Positive impact on local employment	Net change in short-term service and construction jobs due to project implementation
		Net change in long-term agricultural and service jobs due to changes in land and recreational use

3.2 Physical River Objectives

The physical characteristics of a river reach are important stakeholder concerns on their own and are also fundamental factors influencing most biological and economic attributes. To predict how these characteristics would change as a function of local river widening, we developed a synthesis model based to a large degree on the results of work published by other research groups [see Schweitzer *et al.* 2007a for details].

3.2.2. River Morphology

To predict whether a river will tend towards a braided or single-threaded morphology after the release of lateral constraints, we used the logistic regression model of Bledsoe and Watson [2001], in which the probability, p_m, of a multi-thread pattern can be estimated as,

$$p_m = \frac{\exp\left[3.00 + 5.71 \cdot \log_{10}\left(J_V \cdot \sqrt{Q_a}\right) - 2.45 \cdot \log_{10}\left(d_{50}\right)\right]}{1 + \exp\left[3.00 + 5.71 \cdot \log_{10}\left(J_V \cdot \sqrt{Q_a}\right) - 2.45 \cdot \log_{10}\left(d_{50}\right)\right]} \quad (1)$$

where J_V is valley slope (-), Q_a is mean annual flood discharge (m^3s^{-1}), and d_{50} is median gravel diameter (m). This probabilistic expression could be used directly as a conditional distribution in the probability network model.

To determine the effects of any remaining width constraints on final morphology, we used the pattern diagram of da Silva [1991] which predicts whether a river section will be braided, meandering, alternating or straight, conditional on gravel size, width constraints, and mean depth at bankfull discharge. Finally, gravel transport calculations based on Meyer-Peter and Müller [1948] (for a single-threaded morphology) and Zarn [1997] (for a braided river morphology) were implemented to determine whether there is sufficient deposition in the widened reach to form the gravel structures required for a braided or alternating gravel bar morphology.

Gravel movement and substrate siltation is a crucial ecological attribute because fish and benthic species depend on the interstitial gravel zones for shelter and egg development. We modeled siltation as a process of fine sediment accumulation that occurs over time at a rate which depends on hydraulic and bed characteristics [Schälchli 1995]. This process is disrupted by high floods accompanied by high bottom shear stress. This disturbs the gravel bed matrix and clears it of fines. The threshold shear stress for bed movement can be calculated according to Günther [1971] and converted to a critical discharge using Strickler's formula for single-thread rivers and Zarn's (1997) formula for braided rivers. The frequency of river bed clearance can then be determined from the hydrograph. This frequency together with the rate of fine sediment buildup determines the temporal extent and severity of clogging.

3.2.2. River Hydraulics

To predict the joint distribution of flow velocity and depth in a rehabilitated reach after widening, we developed a statistical model based on point data from 92 stream reaches [see Schweitzer *et al.* 2007b for details]. We found that, for reaches with a braided or gravel bar morphology, the bivariate distribution of relative velocity and relative depth could be described by a mixture of two end-member distributions, one normal and the other lognormal, each with fixed parameters. The contribution of each shape for a particular reach at a particular discharge could then be related to the reach mean Froude number, the reach mean relative roughness, and the ratio of the survey discharge to the mean discharge. For straight morphologies, we found that the joint distribution of relative velocity and relative depth could be described by fixed beta-distributed marginals correlated with a rank correlation coefficient of 0.94.

The proportions of a reach consisting of pools, runs, and riffles can be calculated directly from the predicted bivariate distributions, using quantitative definitions of these hydraulic units in terms of point depth and velocity. Following Jowett [1993], we defined pools as having values of the Froude number less than 0.18 and a velocity/depth ratio less than 1.24 s^{-1}, riffles as having Froude numbers greater than 0.41 and a velocity/depth ratio greater than 3.20 s^{-1}, and runs as having intermediate values.

3.3 Biological Objectives

3.3.1 Benthic Organism Abundance

Periphyton and invertebrates dominate the first levels of the trophic pyramid in many small and intermediate size rivers and therefore can influence the complete ecosystem of running waters. They also influence water colour, clarity and odour by utilizing nutrients and organic material. Finally, anglers also rely on macroinvertebrates as the main source of food for sport fish.

To predict periphyton and invertebrate density in rehabilitated rivers, we used simple models that were mechanistically motivated but have lower data requirements than detailed simulations [see Schweizer *et al.* in review for details]. They describe the density of periphyton and various invertebrate functional feeding groups based on days since the last bed-moving flood, mean water depth, substrate size, mean flow velocity, and day of the year. Model parameters were estimated using a combination of literature results and statistical fit to survey data from a set of Swiss and French rivers (Figure 1). Considering their simplicity, the models show a remarkably good fit to time series measurements. For periphyton, total invertebrates, collector-gatherers, and predators, R^2 values ranged from 0.52 to 0.71. Scrappers were modelled less well (R^2=0.26), and shredders and filterers were too scarce in our data sets to be modelled.

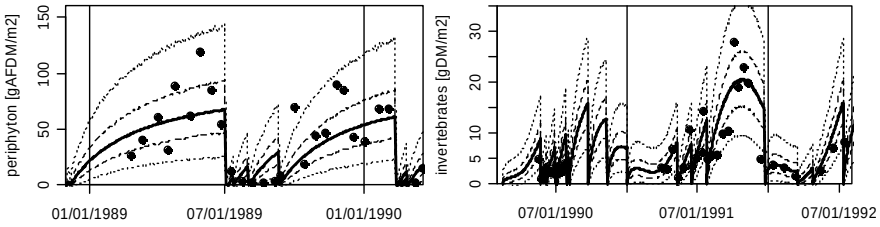


Figure 1. Example fits of benthic model to data on periphyton (left) and total invertebrates (right). Solid lines represent best estimates, dashed lines bound the 50% predictive intervals, and dotted lines bound the 90% predictive intervals. Solid circles represent measured data. Data from the Necker Aachsäge (left) [Uehlinger 1991] and Sihl (right) Rivers [Elber et al. 1996].

3.3.2 Shoreline Fauna Abundance

Riparian arthropod density is an important indicator of shoreline fauna abundance. Arthropods contribute significantly to overall riverine biodiversity and represent a functionally important component of river ecosystems. Our model focuses on predicting the abundance of three major arthropod groups (spiders, ground beetles, and rove beetles) as well as total arthropod abundance

We used multiple regression analyses to relate the variation in each species' abundance to the river morphology and shoreline embeddedness (Figure 2) using data from twelve, differently-impacted, river sections of seven, mid-size to large, rivers in Switzerland and Northern Italy [Paetzold and Tockner, in review]. We used a backward stepwise regression procedure to assess which variables and interactions explain most of the variation. All regressions were performed using the square root transformation of abundance data to improve the normality of model residuals.

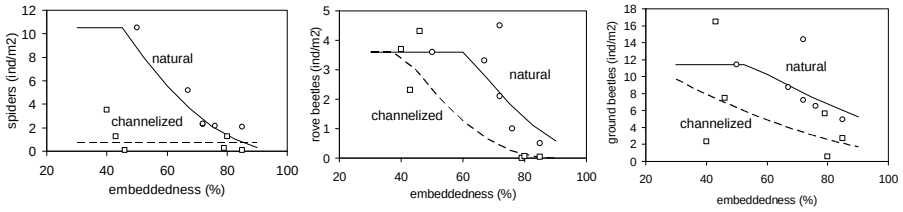


Figure 2. The dependence of arthropod density on shoreline embeddedness and river morphology. Circles and solid lines represent the data and model fit, respectively, for natural (braided or gravel bar) rivers. Squares and dashed lines represent the data and model fit for channelized rivers. Data from [Paetzold and Tockner, in review].

We found that for all species there were significant differences between natural and channelized river sections. Additionally, embeddedness reduced the abundance of all species similarly in both types of morphologies, except for spiders at channelized sites which were already so low that embeddedness had no further effect. Rove beetles were the most precisely predicted, with an R_{adj}^2 value of 0.80, and ground beetles were the least precise with an R_{adj}^2 of 0.29.

3.3.3 Fish Abundance

Salmonids and cyprinids are two key families of fish in many large rivers. They are fished and farmed for food across Eurasia and are the major species of fish eaten in many land-locked countries. Salmonids are also an important recreational species for anglers.

To model an important salmonid, brown trout, we started with a dynamic, age-structured population model [see Borsuk *et al.* 2006 for details]. This model is characterized by population parameters, such as growth, survival, and reproductive rates, which were linked to external indicators of habitat quality and anthropogenic influence using experimental and field data, literature reports, and the elicited judgment of scientists. Important influences relevant to river rehabilitation included physical habitat conditions (e.g. % riffles, depth and velocity variability, and substrate size), flood frequency, stocking practices, and angler catch. Effect strength and associated uncertainty were described by conditional distributions directly encoded in the probability network model. The model was tested using data from populations at twelve locations in four Swiss river basins. First applications of the model involved predicting the effect of candidate rehabilitation measures at these twelve sites.

A model for cyprinids is still being developed. Because this family is less well studied than salmonids, it is likely that this model will be more empirical than mechanistic in its structure. We anticipate using habitat suitability data as the basis for model relations.

3.4 Economic Objectives

3.4.1 Flood Protection and Project Costs

In most river rehabilitation projects, flood protection level is specified as a constraint on the minimal expected return period of a flood for which adequate protection must be provided. Project costs then follow from this flood protection level as well as the project design. Costs include both the initial construction cost, as well as ongoing costs for maintenance.

3.4.2 Local Employment Impacts

To estimate the impact of river rehabilitation on short-term employment in the construction sector and long-term employment in the service and agricultural sectors, we used an input-output model parameterized for the local economy [see Spörri *et al.* 2007 for details]. This type of model uses an input-output table of the goods and service flows between different sectors of the economy to calculate the change in output and jobs per sector resulting from a specified demand change (in the construction or service industries, for example) [Miller and Blair 1985]. Reductions in agricultural employment caused by changes in land use are accounted for by assuming that the agricultural sector is constrained by the land available and that the residual local demand for agricultural goods is compensated by imports.

3.5 Model Implementation

The submodels described in the above sections were implemented using the software package Analytica (Figure 3), a commercially available program for evaluating probability network models [Lumina, 1997]. The inputs to the model can be determined for a river system of interest from historical data, and the decision variables can be set to values corresponding to various rehabilitation alternatives. A large sample of realizations is then drawn for each marginal and conditional probability distribution using random Latin hypercube sampling. These samples are propagated to model endpoints to generate distributions of results which represent uncertainty and natural variability. When combined with a model of stakeholder preferences, these endpoint distributions provide a rational basis for stakeholders to decide among rehabilitation alternatives or to improve a certain alternative. [Reichert *et al.* 2007].

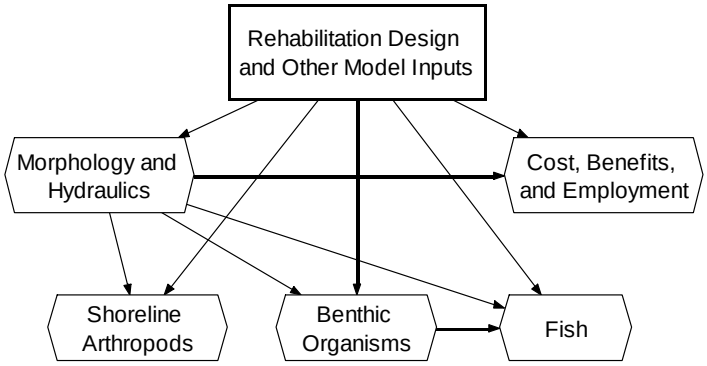


Figure 3. Schematic of the Integrative River Rehabilitation Model. The rectangular box represents rehabilitation design variables (e.g., river width constraints, flood plain and levee height, distance between levee) and other model inputs (e.g., slope, gravel size). Hexagons represent submodels predicting key endpoints. Arrows represent causal influences.

4. CASE STUDIES

4.1 Site Descriptions

We present three case studies to demonstrate application of the IRRM to different locations. The first is a rehabilitated section of the Moesa River in the Swiss canton of Graubünden. This section was originally channelized in the years 1896-1912 to protect the Rhätischen train line and to provide agricultural area. After the region was listed as an area of national importance, a rehabilitation project was financed in 1999. Along a section where it would not present an immediate risk to adjacent populated areas, the river was relieved of its side constraints for 600m along the right bank and 280m along the left bank. The river is now free to expand and run its natural course along this section. We will use the model to generate predictions of the current rehabilitated status and compare these predictions against actual conditions.

The second and third case studies concern two rehabilitation projects (one accomplished and one planned) along the Thur River in the Swiss canton Thurgau. Historically, annual floods of the Thur prevented settlement along its banks. In 1890, a first correction of the river involved straightening meanders and building levees on either side. However, occasional large floods continued, and riverbed erosion worsened on the majority of the river course. The monotonous channel also impaired breeding grounds for birds, fish and other aquatic organisms. To overcome these problems, the Thur has been rehabilitated in some places over the past 10 years. In 2004, a widening was conducted near Niederneunforn at the border with the canton of Zürich. In this 1.5 km section, where mean discharge is 49 m^3s^{-1}, the river was widened from 50 m to 120 m. For this location we will also compare model predictions to actual conditions.

Finally, we will generate predictions for a planned widening of the Thur between the towns of Weinfelden and Bürglen. This is a 4 km long, 30 m wide section, with an average discharge of 41 m^3s^{-1}. It is being proposed to widen this section to up to 200 m. We will evaluate the potential of such a widening to meet stakeholder objectives.

4.2 Model Predictions

Model results show very different predicted outcomes of widening at the three locations (Table 2). The Moesa is most likely to take on a braided or alternating gravel bar form, with a mix of riffles, runs, and pools and an associated variety in velocity and depth. This is predicted to support abundant periphyton, invertebrates, and arthropods, as well as an abundant brown trout population. For comparison, after rehabilitation this section of the Moesa has indeed taken on a blend of braided and alternating gravel bar morphologies, with about 33% of the area classified as riffles, 33% as runs and 33% as pools. Unfortunately, there have not been measurements of periphyton, invertebrate, or arthropod densities, however brown trout surveys have revealed densities between 123 and 192 ind/ha.

The Thur at Niederneunforn is predicted to be alternating or straight, with a predominance of runs and a less diverse depth structure. The resulting high frequency of bed-moving floods leads to a low predicted periphyton density, although invertebrate and arthropod densities are predicted to be fairly high. The river at this location is too large and warm to support brown trout. Observations show that this section actually has an alternating gravel bar morphology and has about 25% riffles, 60% runs and 15% pools. There are no post-rehabilitation measurements of periphyton, invertebrate, or arthropod densities against which to compare predictions. Fish population surveys have found maximum brown trout densities of only 19 ind/ha.

After widening, the Thur at Weinfelden is predicted to remain straight, primarily because there seems to be insufficient gravel input to develop braided or alternating gravel bar structures. Therefore, velocity and depth are expected to stay fairly monotonous dominated by runs. Construction costs of 31 million CHF are expected to lead to short-term employment of about 49 full time equivalents (fte), while changes in land and recreation use will only add about 1 or 2 long-term fte. The Thur at Weinfelden in not expected to support brown trout after rehabilitation.

Table 2. Summary of model predicted outcomes for three implemented or planned river rehabilitation projects.

Attribute	Moesa	Thur - Niederneunforn	Thur - Weinfelden
Morphological type (probability of braided, alternating gravel bar, or straight)	0.46 braided, 0.34 alternating, 0.20 straight	0.0 braided, 0.56 alternating, 0.44 straight	0.29 braided, 0.08 alternating, 0.63 straight
Coefficient of variation of velocity and depth[a]	0.7 velocity, 1.0 depth	0.7 velocity, 0.7 depth	0.38 velocity, 0.55 depth
Percent riffles, runs, and pools	43% riffles, 45% runs, 12% pools	12% riffles, 63% runs, 25% pools	4% riffles, 96% runs, 0% pools
Summer density of periphyton (g AFDM m^{-2})	26.0 ± 18.5	7.5 ± 3.8	6.5 ± 9.1
Summer density of total invertebrates (g dry wt m^{-2})	20.9 ± 7.8	18.6 ± 7.2	7.4 ± 4.3
Summer density of arthropods (beetles+spiders, ind m^{-2})	26.5 ± 5.7	26.5 ± 7.4	14.1 ± 7.0
Density of adult brown trout	180 ± 132	0	0
Implementation costs (million CHF)	0.8	9.9	31[b]
Net change in short term employment (fte)	NA[c]	16.1 ± 0.8[d]	49 ± 2.6
Net change in long-term employment (fte)	NA[c]	-3 ± 0.5[d]	1 ± 1.3

[a] this result and those for all lower rows are reported for the most likely morphology only
[b] rough cost estimation for demonstration purposes only
[c] relevant economic data not readily available as model input for region surrounding Moesa
[d] employment predictions made using economic data from the region surrounding Weinfelden

5. CONCLUSIONS

Additions and improvements are still being made to the IRRM, however the present version provides a coherent and flexible framework for predicting the ability of river rehabilitation projects to meet many important stakeholder objectives. Because of its modular structure, the model can be easily adapted as necessary for project-specific needs. Unfortunately, very few data are available to test the model's predictive accuracy. Collection of such data is recognized to be an important need for assessing rehabilitation project success [Woolsey *et al.* 2007].

To form a more complete and quantitative basis for rational decision making, probabilistic model predictions can be combined with a formal description of stakeholder preferences in the form of multiattribute utility functions [Keeney and Raiffa, 1993]. Preliminary such functions are reported by Hostmann et al. [2005], and we are currently working to elicit more detailed preference structures from stakeholders and scientists.

ACKNOWLEDGEMENTS

This study was supported by the Rhone-Thur project, which was initiated and funded by the Swiss Federal Office for Water and Geology (BWG), the Swiss Federal Institute for Environmental Science and Technology (EAWAG) and the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL). We thank the many Rhone-Thur project partners and members of the Laboratory of Hydraulic, Hydrology, and Glaciology (VAW) for stimulating discussions, comments, and datasets.

REFERENCES

Bledsoe B.P. and C.C. Watson, Logistic analysis of channel pattern thresholds: meandering, braiding, and incising, *Geomorphology, 38*, 281-300, 2001.

Borsuk, M. E., P. Reichert, A. Peter, E. Schager, and P. Burkhardt-Holm, Assessing the decline of brown trout (*Salmo trutta*) in Swiss rivers using a Bayesian probability network, *Ecological Modelling, 192*, 224-244, 2006.

da Silva, A.M.A.F, *Alternate Bars and Related Alluvial Processes*, Master of Science Thesis, Queen's University, Kingston, Ontario, Canada, 1991.

Elber F, Hürlimann J and Niederberger K. Algenmonitoring als Grundlage für das Abflussmanagement in der Sihl. *Wasser, Energie, Luft 88*, 55-62, 1996.

Hostmann M., M. Borsuk, P. Reichert, B. Truffer, Stakeholder values in decision support for river rehabilitation, *Large Rivers, 15, Arch. Hydrobiol. Suppl. 155*, 491-505, 2005.

Jowett I.G., A method for objectively identifying pool, run, and riffle habitats from physical measurements, *New Zealand J. of Marine and Freshwater Research, 27*, 241-248, 1993.

Keeney, R.L, and H. Raiffa, *Decisions with Multiple Objectives*, Cambridge University Press, New York, 1993.

Lumina. *Analytica, Version 4.0*, Lumina Decision Systems, Los Gatos, CA, 2007

Meyer-Peter, E. and R. Müller, Formulas for bedload transport, Proc. IAHR 3rd Congress, Stockholm, 1948.

Paetzold, A., C. Yoshimura, and K. Tockner, Riparian arthropod responses to flow regulation and river channelisation, in review.

Pearl, J., *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Mateo, CA, 1988.

Reichert P., M.E. Borsuk, M. Hostmann, S. Schweizer, C. Spörri, K. Tockner, and B. Truffer, Concepts of decision support for river rehabilitation, *Environmental Modelling & Software, 22*, 188-201, 2007.

Rohde S., M. Schutz, F. Kienast, P. Englmaier, River widening: An approach to restoring riparian habitats and plant species, *River Research and Applications, 21*(10). 1075-1094, 2005.

Schälchli, U., Basic equations for siltation of riverbeds, *Journal of Hydraulic Engineering, 121*(3), 274-287, 1995.

Schweizer, S., M.E. Borsuk, and P. Reichert, Predicting the morphological and hydraulic consequences of river rehabilitation, *River Research and Applications, 23*, 303–322, 2007a.

Schweizer, S., M.E. Borsuk, I. Jowett, and P. Reichert, Predicting joint frequency distributions of depth and velocity for instream habitat assessment, *River Research and Applications, 23*, 287–302, 2007b.

Schweizer, S., M.E. Borsuk, U. Uehlinger, S. Bouletrêau and P. Reichert, Predicting the biomass of periphyton and macroinvertebrate functional feeding groups in stream reaches, in review.

Spörri, C., M. Borsuk, I. Peters, P. Reichert, The economic impacts of river rehabilitation: a regional input-output analysis, *Ecological Economics 62*, 341-351, 2007.

Uehlinger U. Spatial and temporal variability of the periphyton biomass in a prealpine river (Necker, Switzerland). *Arch. Hydrobiol. 123*, 219-237, 1991.

Woolsey, S., F. Capelli, T. Gonser, E. Hoehn, M. Hostmann, B. Junker, A. Paetzonld, C. Roulier, S. Schweizer, S.D. Tiegs, K. Tockner, C. Weber, A. Peter, A strategy to assess river restoration success, *Freshwater Biology 52*, 752–769, 2007.

Zarn B., Influence of river width on interactions between discharge, morphology and transport capacity. *Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie*, 154, ETH, Zürich, Switzerland, 209pp. 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

Participatory Approaches in Developing a Model to Assist Water Resource Management in a Catchment in the Solomon Islands

T. Chan[a], B. Powell[b], S. Hoverman[c] and H. Ross[c]
[a] Water Studies Centre, Monash University, PO Box 23, Victoria 3800, Australia (terence.chan@sci.monash.edu.au)
[b] International WaterCentre, Brisbane, Australia (b.powell@watercentre.org)
[c] The University of Queensland, Brisbane, Australia (helen.ross@uq.edu.au, suzanne.hoverman@uq.edu.au)

Abstract: This paper describes the practical application of a participatory approach used in developing a model for assisting water resource management in the Kongulai catchment in the Solomon Islands. In collaboration with local water resource managers, the Kongulai was selected as the study site as it provides up to 60% of the water for Honiara, the capital. Management of this resource is complex, with potentially competing uses for the water and the catchment, including drinking, domestic, agricultural and industrial uses, as well as multiple threats from contamination, changing land-use and variable hydrology. Additional system considerations come from the multifaceted socio-economic and institutional arrangements. Stakeholder consultation was a key element in the model development process. The three main stakeholder groups, the customary landowners, the government, and non-governmental organisations, were consulted separately in May 2007, to ensure openness in identifying stakeholder concerns and to elicit each groups' understanding of the catchment and how it worked with respect to water resources. During a further visit in October 2007 all stakeholder representatives were brought together and preliminary results combining outputs from the May consultations were presented for discussion and feedback, and prioritisation of concerns and issues to be included in a quantitative model. Because of its intuitive graphical basis, a Bayesian belief network was considered an appropriate tool, and is being developed based on the stakeholders' conceptual diagrams. Involvement of representative stakeholders and accounting for their concerns as well as using their local knowledge of the system was intended to build trust in the model development process and in any outcomes, as well as facilitate relationships between the different groups affecting, and affected by, the catchment. Inclusion of local knowledge is also essential to model development in cases such as this, where little quantitative data is available.

Keywords: Participatory Approaches; Conceptual Model; Catchment; Water Resources.

1. INTRODUCTION

This paper describes the practical application of a participatory approach currently being used in the development of a model for assisting water resource management in the Kongulai catchment in the Solomon Islands. Local water resource managers identified the Kongulai as the priority catchment, as it provides up to 60% of the water for Honiara, the capital. Management is complex due to multiple uses for, and threats to, the water, and as managers need to evaluate, prioritise and find solutions to the multiple, potentially serious hazards, a risk assessment approach was deemed appropriate [e.g. Hart ***et al***. 2006]. As part of this approach, a quantitative model is needed to integrate these issues and allow prioritisation of on-ground management actions. Quantification will also provide a more

solid comparison against competing interests (for example, logging) for funding improved water management.

Transfer of modelling technologies to environmental managers and ensuring their application in improving management is a challenge [Matthews *et al*. 2006], and much of the process described in this paper was intended to address this. Stakeholder consultation is a key element in model development processes. Increasing the involvement of on-ground managers and local communities in developing models of a particular system is an important factor in encouraging uptake of models and their outputs [Castelletti and Soncini-Sessa 2007]. Because of its intuitive graphical basis, a Bayesian network (BN) is intended as an appropriate tool to develop, based on the stakeholders' conceptual diagrams. Bayesian techniques are also appropriate in this case because of the scarcity of data, incomplete understanding of the relatively inaccessible system, and the high uncertainties involved [Cain 2001]. The quantitative model development is currently underway.

2. METHODOLOGY

2. 1 Study Site

A situational analysis of the Solomon Islands and discussions with local water managers indicated that Kongulai catchment, just west of Honiara, the capital of the Solomon Islands, on the island of Guadalcanal, was an appropriate study site [Wairiu and Powell 2006]. The catchment is under traditional, or customary or ownership. It is approximately 50 km^2, and is mainly comprised of mountainous forest, although there has been some logging, and there is subsistence agriculture near the settlements that occur toward the coast. In addition to the values, threats and catchment complexity being of interest, the accessibility of the site and considerations of data availability were also important in site selection.

2.2 Ecological Risk Assessment

A catchment-based Ecological Risk Assessment (ERA) approach similar to that used by Hart *et al.* [2006] is being used to guide development of a model to quantitatively assess threats and hazards to water in the catchment and to fit into the management process. The ERA approach is illustrated in Figure 1. It can be considered to be made up of two primary phases, the Problem Formulation phase and the Risk Analysis and Characterisation phase.

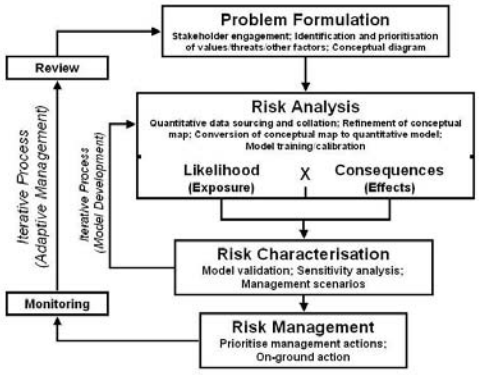


Figure 1. The Ecological Risk Assessment Process

The Problem Formulation phase involves consulting the stakeholders to determine the scope and focus of the risk investigation and ensure the relevancy of the project. Stakeholder consultation gives stakeholders an understanding of the risk assessment process, encourages their input, and provides a basis for ownership of the project and its outcomes. Group discussion and selection of the key value helps the stakeholders gain consensus on the management objectives and narrows the project scope, and discussion of the threats in the region identifies the primary risks that should be considered. Diagrams of stakeholder understanding ("conceptual diagrams") of the system can be constructed around the key values, incorporating the primary threats/hazards, and revealing what the stakeholders consider important and also how they understand the system to work.

The Problem Formulation is then used as the basis for the Risk Analysis and Characterisation phase of the ERA, analysing the consequences and likelihood of each of the risks identified in the Problem Formulation. These are then combined (in this case, by conversion to a Bayesian network, which is being done based on work done by Cain [2001]) to compare, rank and prioritise the risks on the basis of seriousness relative to management objectives. This phase is iterative, as understanding of the catchment improves during analysis and characterisation of risks, the model can be improved.

The Risk Analysis and Characterisation then feeds into Risk Management, involving identification and implementation of the best on-ground management actions. Monitoring and review of any management actions implemented makes the entire ERA process iterative and cyclic. Once the risk in the current cycle is successfully managed for, a next iteration of Problem Formulation can identify the next focus for management to address, i.e. the process results in adaptive management.

This paper focuses on the role of the participatory process in the risk assessment project, primarily occurring in the Problem Formulation phase, and how it assists in model development. It should also be noted that because of the developing and foreign context, the environmental/ecological aspects of the region were secondary to human health issues and needs.

3. STAKEHOLDER PARTICIPATION

Participatory processes are the key element of the Problem Formulation phase, which determines the scope of the risk investigation and the type of management information it needs to provide. This also provides the basis for the quantitative modelling of the following phase. Part of the interdisciplinary research team also focused on the qualitative aspects of the stakeholder input to gain an integrated qualitative system understanding of the catchment. The initial situational analysis reviewed the relevant literature, particularly local studies and reports, to prevent repetition of effort, and to provide an idea of the issues and a basis for discussion with stakeholders. This review also provided a first step in identification and mapping of the stakeholder groups and possible representatives.

3.1 Stakeholder engagement

An initial site visit in September 2006 built relationships with local partners, the Solomon Islands Water Authority (SIWA) and the Division for Water Resources, and helped identify further interest groups within three main areas: government and resource management, non-governmental organisations (NGOs), and the catchment community. A list of potential stakeholder groups and representatives was gathered from contacts in the Australian Agency for International Development (AusAID) and the local partners. Discussions with these organisations and individuals identified further potential stakeholder representatives, which subsequently revealed yet more parties of interest. The preliminary discussions with these networks of people were very important in identifying the most appropriate stakeholder groups and representatives. The discussions also helped delineate appropriate methods for culturally sensitive initiation of contact. Finally, this networking put us in contact with a respected government employee who was also related to the Kongulai landowner clan, and who was willing and able to act as a cultural interpreter, and facilitate our contact and interactions with them as well as act as translator where necessary.

There is a history of tension between the three main stakeholder groups, and as a result, three separate sets of stakeholder consultations were held in May 2007, firstly with the customary landowners, secondly with representatives from the relevant government departments and water management groups, and finally with non-governmental organisations (NGOs). Separate consultations also ensured openness in identifying stakeholder concerns and in elicitation of each groups' understanding of the catchment and

how it worked with respect to water resources. As these sessions progressed, perspectives and information from the other stakeholder groups were also introduced and discussed, and the common aspects between groups emphasized to set the stage for combined consultations in October 2007.

3.2 Community consultations

Two large information sessions were held with the two sub-tribes of the Kongulai customary landowner group (30-35 people each). Although almost all participants understood and spoke English, some felt more comfortable in their local dialect, particularly when speaking, so the community facilitator also translated each way where necessary. The procedure for these consultations began with our community facilitator introducing us to the community group. The research team then introduced the project, with pauses for translation, setting out our ethics procedures, describing our aims in identifying how to improve management of the catchment, how they could help us, how the project might help them, and how it might be generally useful. We were explicit that we were independent researchers and not working on behalf of the local authorities with whom they might normally interact on land and resource issues. We also explained that we had no say over further funding beyond this particular project. Questions and discussion followed. Subjects discussed included: whether it was in their interest to be involved, the problems and tension they have with the local authorities, and possible compensation for their time. These large information sessions were vital in clarifying what the project was about and allowing the stakeholders to air their concerns. Much of the discussion was within their group and did not require input from the researchers. Finally, over a communal lunch, each of the sub-tribes selected male and female representatives with a range of ages for four subsequent, small-group consultation sessions (6-9 people each).

The small group discussions were held with male and female representatives separately, at times and venues convenient to them over the following week, and were based around a large map of the catchment with only a few primary features displayed. The participants were asked what more they knew about the catchment, particularly with regard to water. Additional features were gradually added. Note that although mapping by the Department of Lands covers this area, much of the official detail has not been ground-truthed. The landowners are the only people who regularly visit the area and a number of discrepancies from the official map were revealed by this exercise. Questions were used to encourage further discussion, these included prompts about catchment features and landuse; how water moved through the catchment; the location and behaviour of rivers, springs, sinkholes and runoff; personal water collection and use; other local uses for water; threats to water; what they valued in the catchment and any associated threats; and what relationships they saw between different factors.

A primary finding was the significant differences in identification of the location and extent of logging operations by those familiar with the area, which when combined by a lack of government monitoring is of concern. Additionally, some rivers thought to be perennial were discovered to be intermittent, and there was significant uncertainty in the location of sinkholes and springs affecting the water supply.

3.3 Government and NGO consultations

The government and NGO consultations were more structured because of the educational and work backgrounds of the participants, and there was more time available due to less remote venues and because there was no need for translation. In addition to discussions similar to those described for the community consultations, elicitation of explicit conceptual maps or influence diagrams for how these groups understood the Kongulai Catchment to work, also occurred.

Government (19 representatives) and NGOs (16 representatives) each had a one-day workshop. After project and team introductions, ethics clearances and catchment map discussions to elicit details as described in §3.2, discussion of values for water in the Kongulai was followed by a formal (blind) vote to identify the priority value.

Discussion of the threats to water was followed by discussion of known or suspected sources of data for any aspects of the Kongulai. Smaller groups of 5-8 stakeholders were formed (3 groups for the government workshop, 2 for the NGOs) to construct conceptual diagrams of how the stakeholders understood the water in the catchment to work, based around the priority value selected earlier, and including those threats they thought were most significant for the Kongulai, as well as any additional factors they considered important. The conceptual diagrams were then presented and discussed in the larger group.

3.4 Stakeholder conceptual diagrams

The small-group conceptual diagrams were merged for government and NGO groups, and community stakeholder variables and linkages added where these were not already included, producing the diagram shown in Figures 2 and 3. This overall stakeholder diagram is too complex for direct use as a Bayesian network structure, and also requires refinement where certain aspects may have been overlooked because they were too obvious, e.g. lack of a link between rainfall and water quantity, or because of the limited timeframe of the stakeholder consultations. However there are several elements of interest in these diagrams.

Figure 2 highlights the differences in the views and focuses between the stakeholder groups. There is a strong government focus on infrastructure such as leakage, pipes and pumping; technical aspects such as geology, and also on the problems they see with end-users, such as overuse (the blue variables). In contrast, the NGOs (in red) focused more on socio-economic aspects such as the costs to end-users and community values, and political factors such as leadership. The landowner focus (in yellow) was more at small-scale use and impacts of water. Additionally, because of the shorter consultation time with the landowners (see §3.3), fewer variables were obtained from them overall, although as discussed below, many of the central features of the diagram were shared with the other stakeholders. Differences in perspective were acknowledged and discussed with the

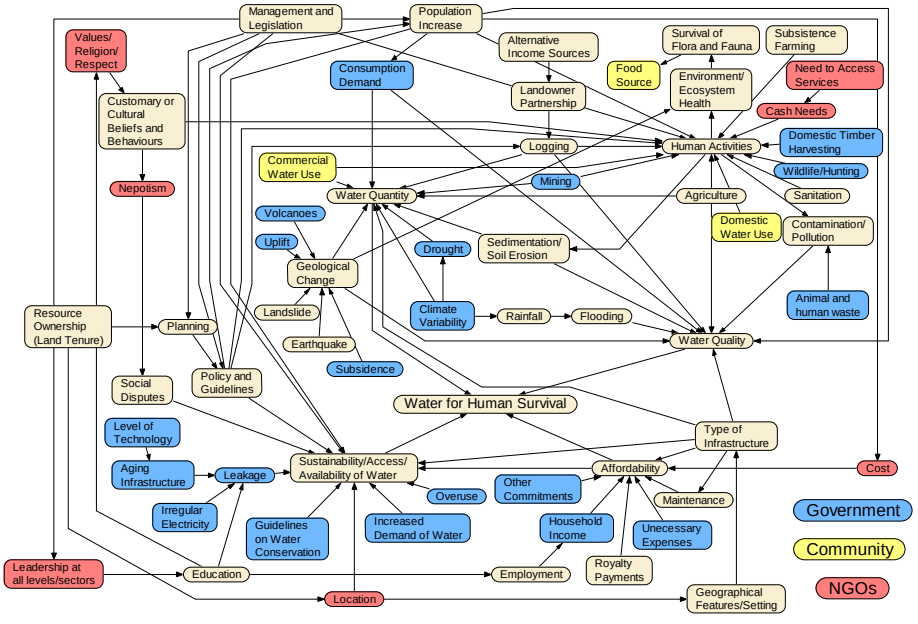


Figure 2. Stakeholder conceptual diagram with differences highlighted. Government-only variables are in blue, NGO-only variables in red, community-only variables in yellow.

stakeholders to get them to think about whether outlier variables were vital factors, and understand why certain factors may not be included in further model iterations.

The overlap in the viewpoints of the different stakeholder groups is highlighted in Figure 3. A number of elements are common to all groups (in grey), including the most central aspects (linking directly to the priority value) water quality, water quantity, sustainability and availability, and affordability. Additionally, all groups regarded the key threats as population growth, the impact of geological change such as earthquakes and landslides, and the landuse impact of logging and agriculture. These similarities were emphasized in the separate consultations to build common ground, and to make it more difficult for some

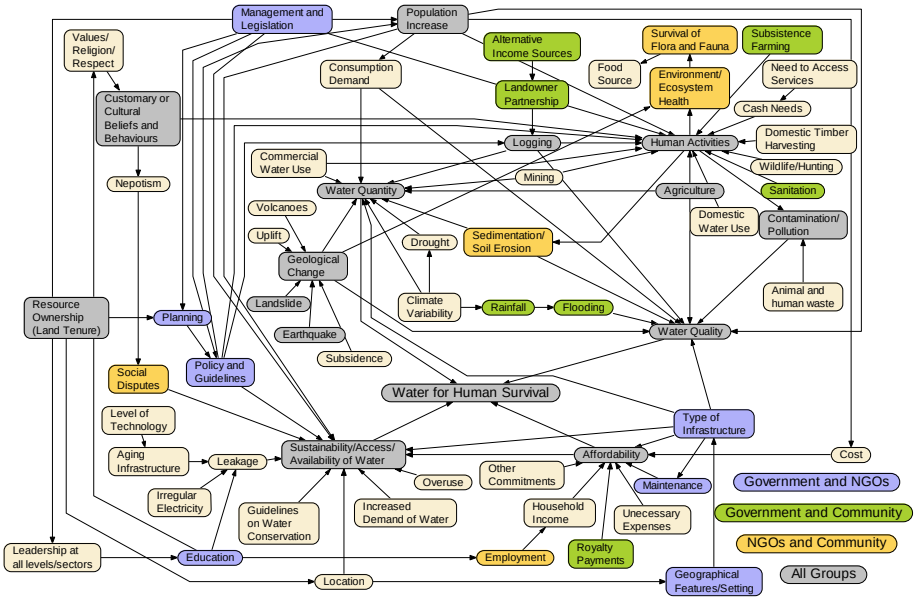


Figure 3. Stakeholder conceptual diagram with similarities highlighted. Variables common to all three groups are in grey, variables common to government and NGOs in purple, to government and community in green, and to NGOs and community in orange.

groups to dismiss the viewpoints of others as coming from ignorance rather than from different priorities. Additionally, these common variables represented consensus on the important factors in the catchment, to be used in the next model iteration.

3.5 Stakeholder follow-up and variable prioritisation

During the subsequent visit in October 2007, a selection of the original stakeholder representatives (28 people) were brought together and preliminary results from the May consultations were presented (note, the fully merged diagrams were not yet complete), emphasizing commonalities while acknowledging the different priorities amongst stakeholders, and also providing reassurance that the problems they've encountered are common to water resource management around the world. A series of small-group tasks, combined with a gradual remixing of the initially self-selected groups, was used to encourage interaction between the different stakeholder groups. The exercises also provided an opportunity for the exchange of views (for example on the relative importance of different factors to water), and also to discuss valid alternative conceptual models of the same system, while promoting consensus and collaboration.

One of the tasks was to simplify the overall conceptual diagram. As there were too many stakeholders to directly go through the network and pare/restructure, clusters of variables were organized into categories such as: water quality, water quantity and human activities. Six mixed sub-groups of three to five stakeholders ranked up to three clusters of variables

each in order of importance to water resources in the Kongulai. The rankings were then averaged across groups, and some of these results are presented in Table 1.

Table 1. Variable prioritisation.

Rank	*Water Quality*	*Water Quantity*	*Human Activities*
1	Sanitation	Rainfall	Population increase
2	Logging	Sinkholes/springs	Water management
3	Pollution/rubbish/waste	Geology/natural disasters	Land management
4	Animal waste	Season	Pollution/rubbish/waste
5	Agriculture	Pipes/infrastructure	Sanitation
6	Flooding	Climate	Logging
7	Natural disasters	Logging	Agriculture
8		Agriculture	Climate change

This prioritisation is intended to allow paring of variables in different sections of the diagram and allow conversion to a Bayesian network. The further steps required for this process are proposed in §4.

3.7 Ancillary Benefits

There were a number of ancillary benefits observed throughout the participatory process. A primary positive outcome was the re-establishment of contact and communication between government and community landowners, via the simple act of officially bringing together the different groups and providing an independent forum and facilitator to discuss water resource issues. This improved relationship was made concrete with funding provided by government for meetings between landowners to discuss water and catchment issues. During the May consultations there was recognition amongst the separate stakeholders that a partnership between government/resource managers and the landowners was needed, and during the October consultation the authors facilitated an initial brainstorming with the stakeholders on how to implement a shared management plan. In July 2007, logging in the Kongulai was halted, and one landowner working for the main logging company in the catchment claims it was a direct response to the May consultation, and that the landowners are waiting for the outcomes of this research before allowing logging to resume. Saw-milling of previously logged wood continues. Finally and unexpectedly, the community facilitator who works in the public health field recognized potential uses for the conceptual diagram in showing colleagues and patients the connections between what occurs in the catchment to water quality and human health.

4 FURTHER WORK

Streamlining of the conceptual diagram for conversion to a Bayesian network is currently underway, based on the variable prioritisation. It should also be noted that the development context of the Pacific region necessitated care in the selection of tools appropriate to the resources of the local community. The science or engineering educational background of many of the local managers was an important factor in deciding Bayesian networks were feasible. Additionally, the research team made arrangements with Norsys, the creators of the Bayesian network software Netica, to provide software access to the on-ground managers.

Further steps required to complete the Risk Analysis and Characterisation phase will involve a further site visit in May 2008, and consultations with a small group of water managers and local experts to confirm model structure plausibility and proposed measures and variable states. Expert elicitation is also required for initial conditional probability tables. A similar procedure to that suggested in Cain [2001] will be used. The available quantitative data for the Kongulai has already been sourced and will then be used for network learning. The trained network will be used to provide an assessment of risk from the potential threats. An analysis of network sensitivity, consequences to key values, and

comparison of different management scenarios will be run past the resource managers for plausibility and usefulness before presentation to the wider stakeholder group.

5 CONCLUSIONS

Involvement of representative stakeholders and accounting for their concerns as well as using their local knowledge of the system built trust in the model development process and facilitated relationships between the managers and others affecting and affected by the catchment. Inclusion of local knowledge is essential to model development in cases such as this where little data is available. A further field trip is planned in May of 2008 for final model refinements and testing with the on-ground managers, and to update the wider stakeholder group.

Lessons from the participatory process include:

- Stakeholder input will require interpretation (model refinement). Ideally, this interpretation should be checked with the participants, firstly to ensure the interpretation still represents their thinking, and secondly to bring them along the process of model development so all stakeholder views develop concurrently.
- Relationship building is important in addressing problems of management and model use/uptake – possibly more than specific model-building processes.
- Expectation management of the participants is important – models are simplifications of reality, not everything can or should be represented.
- A wide range of stakeholders is important – e.g. the few stakeholders who knew of logging sites could easily have been missed in the consultation process.
- However, the larger the stakeholder group, the more unwieldy the results and the more difficult to reconcile different worldviews and aims. A balance is needed.

ACKNOWLEDGEMENTS

The authors wish to thank the Division of Water Resources, its Director, Charlie Bepapa, and Senior Hydrologist, Isaac Lekelalu, the Solomon Islands Water Authority, its General Manager, John Waki, and Jacob Houtarau, local liaisons Gabriel Vagi and Donald Marahare, as well as all stakeholders who participated in the research. This project was funded by the Australian Agency for International Development (AusAID).

REFERENCES

Cain, J., *Planning improvements in natural resource management: Guidelines for using Bayesian networks to support the planning and management of development programmes in the water sector and beyond*. Centre for Ecology and Hydrology, 124 pp., Wallingford, Oxon, 2001.

Castelletti, A. and R. Soncini-Sessa, A procedural approach to strengthening integration and participation in water resource planning, *Environmental Modelling & Software, 12(10),* 1455-1470, 2006.

Hart, B.T., M. Burgman, M. Grace, C. Pollino, C. Thomas and A.J. Webb, Risk-based approaches to managing contaminants in catchments. *Human and Ecological Risk Assessment, 12(1),* 66-73, 2006.

Matthews, K.B., M.G. Hutchins and G. Hill, Bridging the design-use gap for DSS in environmental policy and practice. In: Voinov, A., Jakeman, A.J., Rizzoli, A.E. (eds). *Proceedings of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software*". International Environmental Modelling and Software Society, Burlington, USA, July 2006. CD ROM. ISBN 1-4243-0852-6 978-1-4243-0852-1.

Wairiu, M. and B. Powell, *Situational Analysis Report for the Solomon Islands*. International WaterCentre and the Australian Water Research Facility, 53 pp, Brisbane, Australia, 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

Improving communication in urban planning using TOPIC-COHESIE

H. van Delden[a], P. Valkering[b], P. Luja[a], M. van de Lindt[c] and V. Neevel[d]

[a] Research Institute for Knowledge Systems, P.O. Box 463, 6200 AL Maastricht, The Netherlands (hvdelden@riks.nl, pluja@riks.nl)
[b] International Centre for Integrated assessment & Sustainable development, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands (p.valkering@icis.unimaas.nl)
[c] TNO Environment and Geosciences / Innovation and Environment, P.O. Box 49, 2600 AA Delft, The Netherlands (martin.vandelindt@tno.nl)
[d] Netherlands Ministry of Housing, Spatial Planning and the Environment, P.O. Box 20951, 2500 XP Den Haag, The Netherlands (valentin.neevel@minvrom.nl)

Abstract: Decision making in spatial planning often takes place in a complex and ill-defined context. It includes a large number of actors, factors and –often uncertain– relations. Each of the actors or stakeholders has his or her individual perceptions, values and priorities. This paper presents the TOPIC system, which aims to support this process by interactively collecting and managing the information in a participatory stakeholder process. We focus on a specific version of TOPIC – named COHESIE – which includes functionality for qualitative, participatory modelling. By representing actor perspectives in an explicit and transparent way, the modelling process facilitates the discussion and improves the communication between the different parties. This paper will illustrate the use of TOPIC-COHESIE with a practical example related to urban planning in the Netherlands. The participants of the case study have evaluated TOPIC_COHESIE a useful tool for early stakeholder involvement, and obtaining a more integrated plan.

Keywords: Participatory modelling; Urban planning; Qualitative modelling; Fuzzy Cognitive Maps; Planning Support System.

1. INTRODUCTION

Good communication is crucial in participatory planning processes. Far too often problems arise because stakeholders:

- have their own interpretations of the terminology used;
- have implicit knowledge, which they assume everyone has;
- reason from fundamentally different perspectives;
- lack a shared understanding of the problem and possible solutions;
- have different goals, values and core assumptions; and
- have hidden agendas.

TOPIC (Thematic Orientation of Problem definition in an Interactive Context) has the explicit aim to provide support for interactive planning and decision-making. It improves communication through structuring information, by making implicit knowledge explicit and individual thoughts transparent. Its main focus is on the support of the first phases of the planning process: the development of a well-defined problem definition, a set of objectives, an inventory of possible alternatives and a first assessment of the impact of those alternatives.

This paper focuses on a specific version of TOPIC named COHESIE. TOPIC-COHESIE extends the TOPIC toolbox by including a module for qualitative modelling. It was developed and applied as part of the COHESIE project (ICIS, 2006; ICIS, 2003) supported by the Netherlands Ministry of Housing, Spatial Planning and the Environment (VROM).

We illustrate its application for a case study of city planning in the city of Kerkrade, The Netherlands, in which the system and the accompanying methodology have been used and evaluated. The paper serves as an example of applying participatory qualitative modelling in practice, and aims to illustrate the potential for TOPIC-COHESIE for supporting participatory planning processes. It is to a large extent a summary of the final report of the project (ICIS, 2006).

2. THEORY

2. 1 Participatory modelling

In the application described in this paper use is made of 'participatory modelling' (Van Asselt, 2002), also known as 'group model building' (Vennix, 1996). In this approach model users are actively participating in the model development process.

The main objective of participatory modelling is to provide insight in:

- The system: Because of the inherent complexity of the real-world system, most participants have difficulties to oversee the whole system in which they are operating; relevant (weak and strong) relations, possible feedback loops and sight-effects of actions are often unclear. The aim of the modelling session is to create a mutual understanding about the real-world system. The wide input of stakeholders can very well be used to include different elements of this system.
- Each other's opinions: Participatory modelling facilitates a structured dialogue between stakeholders. The process contributes to an improved understanding of each others objectives, problems, position and perception and stimulates trust amongst the participants. In doing so it can help to reach consensus within the group and can play a significant role in the process of social learning (Ridder, 2005).

Modelling in general, and participatory modelling in particular, encompasses objective and subjective elements. Through participatory modelling, subjectiveness is made explicit since the different perspectives of the participants form the basis of the model. Moreover, uncertainties and differences in opinions become apparent.

In a participatory modelling process different types of models can be used, ranging from conceptual to quantitative. The approach used in TOPIC-COHESIE lies in between those two extremes and can be expressed as qualitative modelling. This technique combines the advantages of transparency and 'soft' elements in conceptual modelling with the possibility to carry out what-if analyses as is normally done in quantitative modelling.

2. 2 Qualitative modelling

The methodology for qualitative modelling in TOPIC-COHESIE is based on the concept of Cognitive Mapping (Axelrod, 1976). A cognitive map is a graphical representation of (the interpretation of) a system and an example of a conceptual model. System variables are represented as boxes; relations between variables are represented by arrows. The term 'cognitive' is used to indicate that the 'map' represents the cognitive interpretation of the system. This can be an interpretation of a scientist, a stakeholder or a policy-maker. Despite of a certain degree of compatibility, the interpretations of complex systems will generally differ from each other at essential points. This should not be seen as a weakness of the approach, it rather is its strength to make those differences explicit.

Fuzzy Cognitive Mapping (Kosko, 1986) is an extension to Cognitive Mapping that allows to qualitatively calculate the impacts of changes in the system. The term 'Fuzzy' is used because the value of the variables is neither numeric nor exact, rather it is interpreted in a linguistic way. The state of a variable is therefore not expressed in absolute values, but is provided in an ordinal scale. In TOPIC-COHESIE, the safety of a neighbourhood is, for example, represented on a scale of seven classes, ranging from 'very safe' (+3), via 'neutral' or 'average' (0), to 'very unsafe' (-3). The strength of the causal relations between

the variables is expressed on the same ordinal scale from 'very positive' (+3), via 'slightly positive' (+1) and 'slightly negative (-1) to 'very negative' (-3).

The traditional Fuzzy Cognitive Mapping methodology only uses causal relations. In practice, however, other relations can be important as well. The housing corporation could, for example, indicate that they will support the plan if there is a minimum number of new housing units being developed. This statement then takes the form of a threshold. To model these relations TOPIC-COHESIE is equipped with an 'if-then' feature (ICIS, 2006). This feature is in line with the concept of Rule Based FCM (Carvalho, 2000).

3. TOPIC-COHESIE

3. 1 The system

TOPIC-COHESIE is a software instrument developed to support participatory, qualitative modelling. It can be used in every interactive policy-making process, decision-making process, planning process or management case with multiple stakeholders. The domain and context of a new case can be set-up easily and can be adapted at any time. The information gathered in discussions with stakeholders (during interviews and workshops) can be stored, ordered, accessed and displayed easily. Moreover, it can be linked by making use of relations. Besides textual information, also other data like documents, drawings, photos and media recordings can be added as information. Structuring the information in this way leads to more insight in the implicit knowledge of stakeholders, deepens the level of information and facilitates a common understanding of the definitions and terminology used.

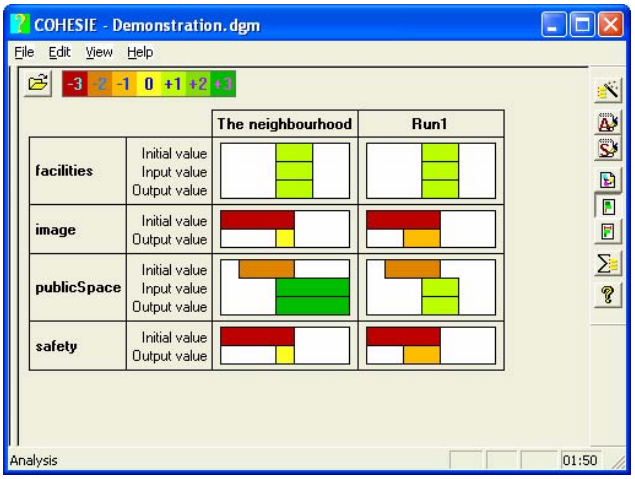


Figure 1. Comparison functionality in TOPIC-COHESIE: the screenshot shows the results of two alternatives -the neighbourhood and run 1- on the indicators facilities, image, public space and safety.

Besides storing and accessing the information in a textual way, it can also be represented in a graphical, conceptual model. Based on this conceptual model, simple qualitative calculations can be made. Stepping through the results of the calculations the stakeholders can see the first, secondary and higher order effects of proposed measures and compare the impacts of different alternatives. From this graphical model, the stored textual information can be accessed and adapted. Furthermore, the perception of different stakeholders towards various alternatives can be visualised. As an extension to the qualitative calculations, an optimization algorithm, a visual comparison module and a sensitivity analysis module are available. The system is complemented with a reporting functionality for quick generation of custom designed reports from the entered information, ready for further editing in word processors.

Although the system is developed for policy-makers and decision-makers at local, regional, and national administrations, also (project) managers, researchers and modellers can benefit from the functionalities of the system. It can be used to provide support in different phases

of the decision-making process, but –because of its systemic approach and focus on causal relations– also as a starting point for the development of quantitative integrated models. More information about the TOPIC-COHESIE system can be found in RIKS (2006). RIKS is also the contact point for any inquiries about the system.

3.2 Modelling city planning with TOPIC-COHESIE

For TOPIC-COHESIE we have developed a standard model for neighbourhoods (ICIS, 2003; ICIS, 2005). This model was developed using expert interviews and experience from previous case studies. It is used as a template when developing specific applications.

The standard model consists of two parts: the physical neighbourhood model and the actor model. The model for the physical neighbourhood includes the main elements that have an impact on the image of the neighbourhood such as: the participation of inhabitants, the fraction rental houses, the safety on the streets, the amount of public space, the availability of facilities, the prices of rental and private properties and the number of vacant houses. In the actor model we have included the various elements that a stakeholder takes into account before giving his or her approval to a plan: the extent to which the plan meets the objectives, its costs and the trust of the stakeholder towards the process and the other stakeholders.

The model development process follows the following procedure: First, the standard neighbourhood model is used as a starting point. Next, we make an inventory of the actors involved and their objectives. Then, we investigate how the neighbourhood system functions, based on the perceptions of the different stakeholders. We discuss with the stakeholders what measures and alternatives can be proposed as part of the plans for restructuring the neighbourhood as well as the costs for each of the stakeholders associated with these measures. Finally, we investigate what is important for the different actors to agree to a measure. Based on all the information above a case specific model can be developed. This model can be fine-tuned by running model simulations and by further interaction with stakeholders.

4. PRACTICAL APPLICATION

Dohmenplein is a neighbourhood in the municipality of Kerkrade in the Netherlands. Although the neighbourhood was characterised as a pleasant living environment a few decades ago, it now experiences problems: a poor living quality, bad image, low quality housing and limited social coherence between residents. Recently, the housing corporation has taken the initiative to revitalise the neighbourhood. It thereby aims to establish broad support for the revitalisation plan amongst the stakeholders involved. In this context, the specific aim of the case study described here was to reach a common view on the problem and its solutions among the key stakeholder group: the housing corporation, the city of Kerkrade and the association of inhabitants.

4. 1 Process

The process followed for participatory modelling with TOPIC-COHESIE is presented in Figure 2. The design of this process is derived from an existing methodology developed by Vennix (1996). The elements in bold (problem definition, interviews, analysis and model development and workshop) are part of the process. The other elements (individual models, joint model, conclusions and evaluation) are results of the process.

The methodology starts with individual interviews and participatory modelling (mapping diversity, diverging), followed by a common stakeholder workshop to bring views together in an integrated assessment, and to reach consensus (consensus building, converging). This set up has the following advantages (see also Vennix, 1996):

- Mapping diversity: In personal interviews a stakeholder can express his or her views without being influenced by the other parties;

- Getting acquainted with the methodology: During the interview session stakeholders became familiar with the system and the process of model development;
- Preparing input for the workshop: The models developed during the individual sessions are valuable input for the workshop.

We carried out two individual interviews per stakeholder. In the first interview (~ 2 hours) we presented a short questionnaire to collect the main objectives, problems and social network relations. Furthermore, the interview contained an interactive modelling session which was used as a first step to develop the individual model. The first interview provided a wealth of information that was processed as part of the analysis and model development phase and resulted in a first version of the individual models. In a second interview (~ 1 hour) this model was brought back to the stakeholder with the aim to ensure that the stakeholder maintains ownership of his model and to reflect on the model, make final changes, ask for definitions of variables included and discuss if-then relations. After the second session the models formed a representation of the perception of the individual stakeholders. They were also perceived by the stakeholders in this way.

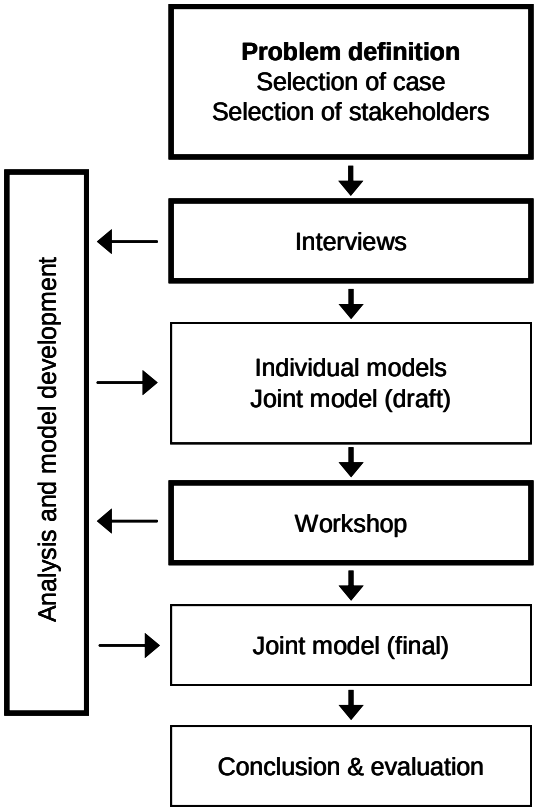


Figure 2. Methodology for participatory modelling with TOPIC-COHESIE.

At the final workshop all representatives of the stakeholders were present. During this workshop the three separate models, developed by the individual stakeholders, were presented, resulting to an improved understanding of each others perceptions. Subsequently an integrated model was presented based on the communalities. Moreover, bottlenecks were defined and discussed. Through the discussion some conflicts were resolved and for others the impact on the overall system was made explicit. At the end of the workshop a model was developed that included all elements relevant to the problem at hand as well as their causal relations. With this final model different alternatives for restructuring the neighbourhood can be assessed on a common platform agreed upon by all stakeholders.

4. 2 Results

In this section we describe the most important results of the case study: the *models* including the three individual models and the joint model as presented in figure 3, an integrated *problem analysis* including agreements and fundamental disagreements (‘bottlenecks) amongst the stakeholders involved, and the *social relational results* of the modelling process.

Models

The individual models are a clear reflection of the perspective of the respective stakeholders. In the model of the municipality, for example, the revitalisation process is

seen in the larger picture of the development of the city. Besides general objectives of quality, affordability and image, the municipality also aims for a lower housing density and puts the focus on age appropriate housing.
The housing corporation has to ensure that new plans fit within her mission as well as her business perspective and financial constraints. Regarding the first, the quality and affordability of the housing stock together with the quality of life are seen as important objectives. To obtain the necessary financial benefits it is important to have a low vacancy rate and a good rental price. In the ideal plan of the corporation there will be a balance between social housing and houses in the private sector.
The perspective of the inhabitant is that of quality of life and affordability. In this model two aspects become apparent. First of all it points out the problem of public nuisance: trash on the streets, noise pollution and gatherings of teen-age groups. Secondly the inhabitant is the only actor that has included trust in its final approval of a plan. At present, trust is low, because of the long duration of the process and the poor communication of this duration. The corporation and the municipality were not aware about the inhabitants' feelings and this was a learning point during the case study.
The objectives of the different stakeholders are summarized in table 1.

Table 1: Objectives of the stakeholders.

	Corporation	**Municipality**	**Inhabitant**
Quality of the housing stock	+	+	+
Affordability	+	+	+
Image	+	+	+
Quality of the public space		+	
Age appropriate housing		+	
Housing density		-	
Vacancy rate	-		
Rental price	+		

Based on the agreements between the models of the different stakeholders a first joint model was developed and presented during the workshop. Using this as a starting point, differences in terminology and points of view were discussed and in an interactive process the model was gradually expanded to its final version (see Figure 3). This version, however, does not yet present a complete consensus between the stakeholders. Although there is agreement on the variables and relations, the relative weight and strength of the relations is not always the same. An example of this is the image of the neighbourhood, which is dependent on the visual appearance, the quality of life and the affordability. The stakeholders agree on these components but give a different weight to the relative importance of these elements.

Problem analysis
The modelling process conducted so far served to clarify, both agreements and fundamental disagreements amongst the stakeholders. It became clear, for example, that all stakeholders shared the opinion that the main focus areas should be housing quality, affordability and image. There is also consensus about the statement that 1) the quality of the housing stock has to improve, to improve the image of the neighbourhood, and 2) a balance has to be found between the prices of the rental properties and the quality, to ensure the affordability of the houses.

Bottlenecks that became clear during the workshop discussions were:

- Mission versus business perspective: The housing corporation in general experiences a tension between her mission (building according to needs of social housing and cheap private housing) and her business perspective. Housing development solely in the social sector does not provide sufficient financial benefits.
- Reducing the housing stock: The first bottleneck is exacerbated by the desired reduction of the housing stock caused by the expected population decline. Reducing the housing stock has negative impacts on the financial benefits for the housing corporation.

- Insufficient insight in the housing needs: Although both the corporation and the municipality value a differentiated population, a clear definition of the target group is missing.
- Lack of trust: the low level of trust of the inhabitant forms a bottleneck in the redevelopment process. According to the inhabitant this problem has occurred because of the long duration of the process. The other stakeholders realise this and will strive to improve the communication about the duration of these types of processes.
- Improvement of public space: While the municipality saw a solution in placing lampposts and benches, the inhabitant agued that this would attract nuisance by teen-age youth and thus aggravate the feeling of unsafety.

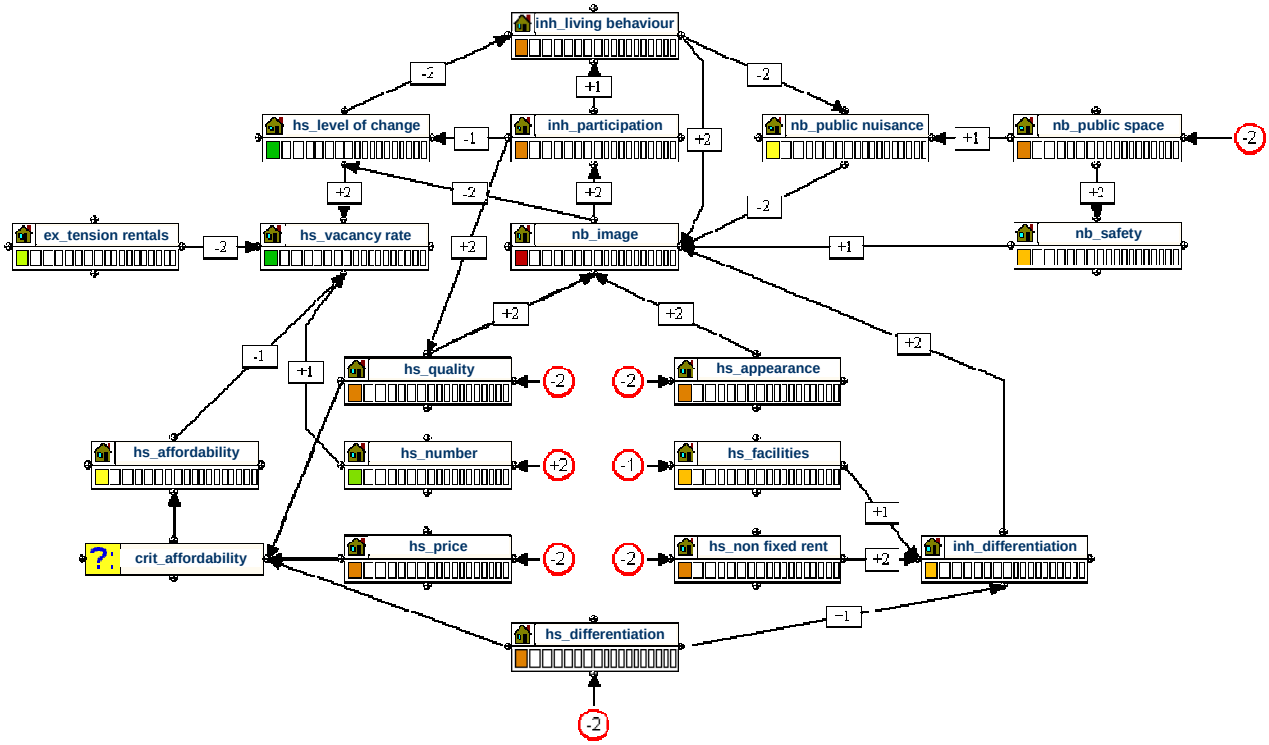


Figure 3. The joint model of the neighbourhood and its processes, agreed upon by all stakeholders. Variables are grouped into those linked to the housing stock (hs), the inhabitants (inh), the neighbourhood (nb) and exogenous (ext).

Social relational results

The participatory process in the case Dohmenplein has been too short to expect major social relational results. However, we would like to describe some initial findings:

- Improved understanding of the core problems: With the support of TOPIC-COHESIE it was possible to present individual perspectives in a transparent manner. Based on this, the main bottlenecks could be identified.
- Changes in perspective: In our opinion, a better understanding of the problems has not –yet– lead to a change in perspectives. In a more elaborate process this could however be an important result.
- Relations and trust: the open and transparent discussion seems to have contributed to improved relations and trust between the stakeholders.

5. CONCLUSION

At the end of the workshop, the participants engaged in an open discussion to evaluate the TOPIC-COHESIE approach. Overall, this evaluation was positive. They see the tool as a useful instrument to include stakeholders in an early stage of the planning process, start the

communication between the stakeholders in an efficient manner and –as a result– obtain an integrated plan. In particular, the graphical modelling functionality –defining variables, providing definitions, creating relations– has proven to be an added value. The bilateral development of the model, as a joint effort of the stakeholder and the modeller is a good step to obtain insight in the position and objectives of the individual stakeholders and brings stakeholders closer together. The workshop session in which the individual models were discussed led to an improved understanding of each others position and objectives.

Critical notes have also been made. According to the stakeholders the tool is too abstract, which might lead to misunderstanding instead of transparency, especially when using it with inhabitants. Secondly, the qualitative calculations have to be interpreted with care. Especially for complex case studies involving a high number of variables, it takes time and knowledge to develop a consistent model in TOPIC-COHESIE. Interactive model building and model explorations during a participatory modelling session can therefore be problematic. The development of illustrative sub-models might be a good way forward in improving the consistency of the qualitative computational models. Third, the TOPIC-COHESIE approach does depend on stakeholders that are willing to cooperate in a transparent and open dialogue. Stakeholders with serious 'hidden agendas' are most probably not eager to participate.

Overall, the use of TOPIC-COHESIE in a participatory setting has been successful. Even though it will never be able to solve all communication problems mentioned in section 1, it has shown its added value as a tool that improves the communication. TOPIC-COHESIE increases the *efficiency* of the communication process (by coming directly 'to the point'), the *transparency* (because the models reveal the actual perceptions and objectives), stimulates *reflection* (because the participants reflect on their own perceptions), supports reaching *consensus* (by creating a common understanding of the system, making each others perceptions and aims explicit, creating trust and stimulating an open discussion) and is useful for the *documentation* of the individual perceptions and objectives as well as the shared understanding of the system. To further develop the tool and the methodology, it is recommended to carry out more practical exercises to experience what users find difficult or inconvenient and to discuss if and how those problems can be solved.

REFERENCES

Axelrod, R., Structure of Decision, Princeton, Princeton University Press, 1976.

Carvalho, J.P., Tome, J.A.B., Rule based fuzzy cognitive maps-qualitative systems dynamics. This paper appears in: Fuzzy Information Processing Society, 2000. NAFIPS. 19th International Conference of the North American.

ICIS, COHESIE: **CO**mmunicatie-instrument **HE**rstructurering **S**teden in een *I*nteractieve **C**ontext. ICIS, Maastricht University, Netherlands, 2003 (www.icis.unimaas.nl).

ICIS, Tussenrapportage COHESIE II. Handelingsmotieven: input voor een gedragsmodel. ICIS, Maastricht University, Netherlands, 2005 (www.icis.unimaas.nl).

ICIS, Eindrapportage COHESIE II: **CO**mmunicatie-instrument **HE**rstructurering **S**teden in een *I*nteractieve **C**ontext. ICIS, Maastricht University, Netherlands, 2006 (www.icis.unimaas.nl).

Kosko, B., Fuzzy Cognitive Maps, In: International Journal of man-Machine Studies 24 (1), 1986.

Ridder, D., E. Mostert, and H.A. Wolters (Eds.), Learning together to manage together, Osnabrück, Germany, University of Osnabrück, 2005.

RIKS, User manual for COHESIE 2, RIKS bv, Maastricht, 2006. (www.riks.nl/projects/cohesie)

Van Asselt, M. B. A. and N. Rijkens-Klomp, A Look in the Mirror: Reflection on participation in Integrated Assessment from a methodological perspective, Global Environmental Change vol. 12: 167-184, 2002.

Vennix, J., Group Model Building: Facilitating team learning using system dynamics, Chichester, England, John Wiley & Sons Ltd, 1996.

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

ARDI: a co-construction method for participatory modelling in natural resources management

M. Etienne[a], D. Du Toit[b] and S. Pollard[b]
[a] *Ecodevelopment Unit, INRA, Site Agroparc, 84914 Avignon Cedex 9 (etienne @avignon.inra.fr)*
[b] *Association for Water And Rural Development, Acornhoek 1360, Limpopo, South Africa (derick@award.org.za, sharon@award.org.za)*

Abstract: The outcomes of a series of tests of the ARDI (Actors, Resources, Dynamics and Interactions) method in complex cases or conflict-ridden situations is presented. ARDI is part of a companion modelling approach that makes it possible to engage a broad spectrum of stakeholders in the design and development of land and water management plans. It is essentially based on participatory workshops that set out to collaboratively imagine a future open, dynamic management system, capable of adaptation and anticipation, by gathering the various affected stakeholders in a partnership dedicated to preserving the natural resources and promoting a sustainable development. Its originality lies in the co-construction of a "conceptual model" of the functioning of the territory, according to a main negotiated development question.
The approach is based on the collective articulation of the key elements of a territory and context by affected stakeholders such as managers, representatives, socio-professional technicians, NGOs, experts and scientists, and local policy makers. This sharing of representations is done by means of a series of collective workshops during which Actors, Resources, Dynamics and Interactions (ARDI), making up the stakes of the territory are identified and clarified. This work of co-construction is conducted within a precise methodological framework that we present in a step-by-step format. The method is also illustrated with concrete examples gleened from the tests carried out by the authors during the last 5 years. Finally, the need for skills development and pitfalls to avoid when applying the method are discussed.

Keywords: Participatory modelling, co-construction, conceptual model, natural resources management, facilitation

1. INTRODUCTION

The application of simulation models in collaborative decision-making for the management of natural resources is one of the characteristics of adaptive management (Holling, 1978; Walters, 1986). But the use of these models to stimulate the participation of stakeholders in the development of management scenarios is much rarer (Costanza and Ruth, 1998; Bousquet et al., 2002). The progressive shift from management plans based on an authoritative or rationalist model towards tools for mediation based on a democratic approach (Van den Belt, 2004) calls for the emergence of new tools of co-construction and sharing of information and understanding.

Following a series of tests of a method, implemented in complex cases (natural areas with multiple use, Biosphere Reserves, Regional or National Parks) or in conflict situations (Heritage Sites, urban-forest interfaces), a companion modelling approach making it possible to involve stakeholders in the design of land and water management plans was developed (Etienne, 2006). It is based on participatory workshops set up to imagine a more open, dynamic management, capable of adaptation and anticipation, by gathering the various stakeholders together to preserve natural resources and promote sustainable

development. Its originality lies in the co-construction of a "conceptual model" of the functioning of a territory, according to the negotiated main development question.

The approach is based on the collective articulation of the key elements of a territory and context by affected stakeholders such as managers, representatives, socio-professional technicians, NGOs, experts and scientists, and local policy makers. This sharing of representations is done by means of a series of collective workshops during which Actors, Resources, Dynamics and Interactions (ARDI), are identified and clarified. This work of co-construction is conducted within a precise methodological framework that we present and illustrate by means of concrete examples resulting from the tests carried out by the authors during the last 5 years.

2. KEY QUESTION AND KEY PARTNERS

The success of the participatory modelling process depends on three key points being directly addressed when initiating the process. These points have to be discussed during one or more preparatory meetings among the mandatory partners and the facilitators of the approach. The first point involves identifying the different types of stakeholders and clearly defining the territory under question. Secondly, one or several facilitator(s) must be identified and their aptitude and legitimacy to carry out the debates during the process of design-validation-use of ARDI tools will have to be appointed. Thirdly, it is necessary to pay special attention to the convocation of the working group: choice of the partners, place of the meetings, periodicity of the workshops, modality of invitation. This is mainly because the representativeness of the participants and thus the richness and relevance of the conceptual model depend on that point.

The ARDI method was tested under a varied set of conditions, questions and territories. It was mainly applied by French researchers working in the field of companion modeling (Collectif Commod, 2006) but several agents of regional natural reserves were trained to apply it in France, and mediators are currently being trained in Western Africa Biosphere Reserves. The success of this approach to natural resources management lies in the relative independence of an external scientific agent, and the familiarity and skill of such a person in the handling of the methodological aspects. However, there is a distinct advantage to engaging a researcher as facilitator who is skilled in both the ecological sciences and social sciences with basic experience in facilitating debates between researchers and managers. But a communication expert can also easily play this role.

Finally, several criteria have to be considered when choosing participants for the exercise. Even if this choice is flexible (it is possible to invite a new participant in the course of the exercise), the process gains from having access to an initial "core group" that will be present throughout the process of co-construction. Three types of situations were confronted during the testing process:
1 - priority given to a global understanding of the system: the participants chosen from extension services of the territory whose local experience legitimizes their position to speak on behalf of the stakeholders that they frequently come into contact with. It is important not to forget any relevant activity according to the defined question, and to avoid over-representing an activity (for example inviting three foresters because there are three forest companies working in the territory).
2 - priority given to the involvement of local stakeholders but by maintaining a global view of the system: the participants are sorted from local stakeholders representatives chosen for their legitimacy (elected democratically, leader of a professional organization) and for the relevance of their activity in relation to the initial question.
3 - priority given to the involvement of local stakeholders whilst seeking to appreciate the diversity of the system: the participants are local stakeholders selected for originality of their practices compared to classical or formal stakeholder groups.

The position and status of researchers in the process is variable and is still being debated amongst the companion modeling community. The general rule is that researchers carrying knowledge of the context and major processes (social, technological, economic, ecological,

and political) be engaged. Some bring expertise to the initial stage whilst others will be integrated at a specific workshop, (frequently the discussion on system dynamics or the design of the interactions diagram), if the participants feel there is a need for an expertise on a particularly topic. As much this differentiation is relatively easy in the field of the ecological sciences, it is problematic in the field of social sciences where the researcher may play the role of the expert who holds a global vision of the social relationships or economic flows. The choice of the venue, the duration and the periodicity of the meetings depend on many factors external to the exercise itself (availability, schedules of obligation, levels of responsibility). But some principles should be negotiated and respected if the method is to be successfully applied. For example, the method is facilitated if the place is easily accessible to participants, and on neutral ground. If not, it must be clearly identified as the legitimate place of the partner who convenes to the exercise or raises the question. Each meeting must at least last 2 hours and the participants must remain centered on the collaborative exercise. The ideal is to conduct all the workshops over a period not exceeding 1 month and the meetings may take the form of: a) a 2-day and a half workshop, b) one half-day per week, c) three separate days.

3. THE ARDI METHOD

3. 1 Co-constructing a common representation

The first step of the companion modeling approach follows the ARDI method (or any similar one), in collectively identifying the principal stakeholders concerned with the key question, their management entities, the resources used and the main processes driving changes affecting these resources. With this intention, the group that takes part in the co-construction of the model must answer the three following questions (the formulation of which is adapted here to the establishment of a sustainable development project):

1. What are the principal resources of the territory and what is the key information to guarantee a sustainable use of these resources?
2. Who are the main stakeholders involved in the use or duty to decide the management practices of this territory?
3. What are the main processes that drive strong changes in resource dynamics?

Dependent on the extant and complexity of the territory concerned, the collective response to each of these three questions can take between 1 and 3 hours. Depending on the level of detail required, this can be between one half-day to one day and a half workshop. It is important that the order of questions be respected and the facilitator must take care that each one participant has the opportunity to deliver an opinion. In the sessions we facilitated, the following simple procedure was adopted: a) a drawing, on an interactive white board, easy-to-see by all the participants, b) for each element of ARDI, each participant has, in turn, the opportunity to respond, c) only one concept to be proposed at a time.

To facilitate sharing mental models and representations, the answers to the questions are formulated as lists of words, with a minimum of coding making it possible to easily classify the information. The workshop is generally led by two people: a facilitator and a secretary. The role of the facilitator is essentially the “hand” of the group and intervening only when the response is formulated either in a too generic form (i.e. to refuse systematically the term manager to define a stakeholder), or with a polysemous word or a term that can lend to confusion (i.e. wood can be the place where trees stand but also the material resulting from the exploitation of these trees). The role of the secretary is to keep track of the exchange between members of the group, or between one participant and the facilitator. Among the key aspects to monitor, three are particularly important: attitudes of the participants to each other as a way to reveal social links, arguments developed to support a proposal or to contradict it as a way to measure the strength of the assessment, and reasons advanced for changing a previously accepted proposal or terminology as a way to follow up the group dynamics. The first will permit to identify social networks, the second to better understand individual mental models, and the third to keep the track of the path followed to reach an agreement.

3. 2 Identifying key stakeholders (“A for actors” in ARDI)

The first stage of the ARDI process culminates in the “actors” diagram (“A” from ARDI) which is composed of the list of stakeholders and the corresponding management entities and the links between them (Figure 1). The exercise proceeds in 3 stages. Initially the participants simply list the stakeholders whom they consider associated with the question. As long as new suggestions for stakeholders are proposed, the facilitator goes on with the next participant or begins a new round from the table. Each “actor” proposed must be a direct stakeholder (people who use or whose practices have a direct impact on key resources of the territory), or an indirect stakeholder (people whose actions will encourage the direct stakeholders to change their practices). Each input is added to the interactive board by the facilitator as a new label, using colors to distinguish the category to which they belong (black case for the direct ones, blue for the indirect ones). The facilitator may suggest to precise certain types of actors (i.e. farmers be subdivided into stockbreeders and wine growers) or challenge the assignment to a category if there is not consensus in the room. A typical example of this type of intervention is the status given to the entity "herd". Certain participants will position it as a resource, others will regard it as an actor. When the grazing impact on grassland dynamics is a significant process, the facilitator may ask whether participants think that the herd is autonomous (it decides where, when and how much it will graze), or if it depends mainly on the decisions of the shepherd. In the first case, one will retain the herd as a stakeholder, in the second case, it will be listed as a resource managed by the shepherd.

Next, the organizer will ask the participants to specify the links which exist between the identified stakeholders and to clarify in a simple way this relationship. Progressively, the facilitator adds arrows according to suggestions made by the participants. He also progressively shapes the diagram by bringing closer the stakeholders who have many relations and moving those away that do not have any. When the participants consider that the main interactions between actors are represented, the facilitator can put the finger on incongruities and gaps (i.e. no link between the stockbreeder and the shepherd) or point out stakeholders without any relation with any other. The facilitator then launches a discussion on the relevance to retain this “actor” in the diagram, while the secretary keeps record of the decisions taken by the group and the justification for the decision (the landowner is the typical example of a stakeholder who does not have a link with anybody but that is often retained in the diagram because he can easily block the development of the activities of another stakeholder).

Lastly, always according to the principle of the negotiatin, the participants must identify and clarify the management entities used by each direct stakeholder. Those can be spatial entities (forest plot, grazing unit, water catchment), or not (herd, cash).

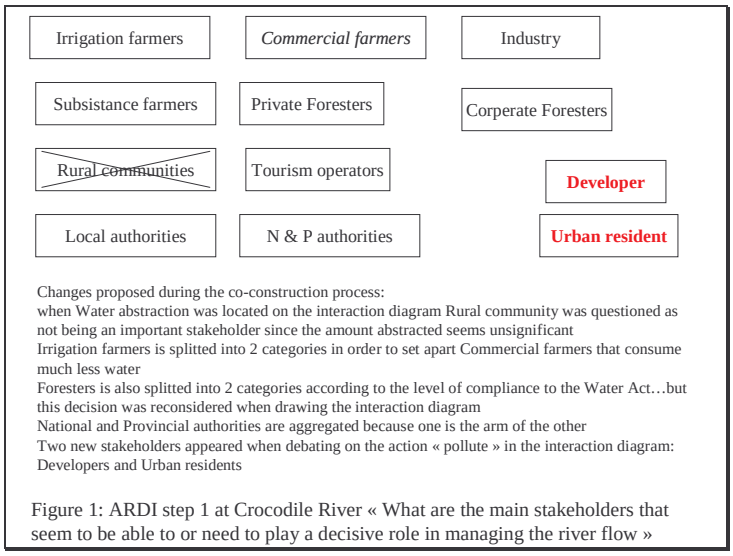


Figure 1: ARDI step 1 at Crocodile River « What are the main stakeholders that seem to be able to or need to play a decisive role in managing the river flow »

3. 3 Identifying key resources ("R" in ARDI)

The second stage consists of listing the relevant resources of the territory according to the key stakeholders previously identified, the word resource applying exclusively to goods or products used by any of the stakeholders (Figure 2). During the collaborative construction of the list, the principal types of resources are often gathered within five main categories (infrastructure, water, minerals, plants and animals). For each resource mentioned, the speaker is brought to justify his/her choice and is encouraged to specify which indicator seems to be the most relevant to make management decisions regarding that resource. Participants are encouraged to explain which characteristics of the resource they evaluate before taking a decision on that resource. This indicator can be quantitative or qualitative and if there is debate or non-agreement, several indicators may be applied to a particular resource. As certain resources are temporary, one may have to specify the period of existence (season, favorable year) and/or long-standing (lifespan of a building, time for filling of a dam). The resources functioning as exogenous variables but whose characteristics are critical in operating the system can also be mentioned (i.e. the rainfall in arid or dry zones). This set of indicators will be used afterwards, during the model implementation and the development scenarios steps, to visualize and compare the stakeholders' points of view (Etienne et al., 2003).

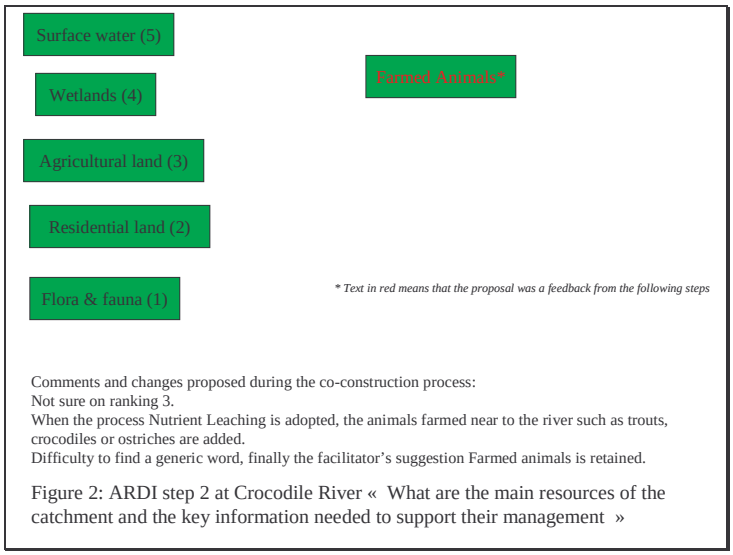


Figure 2: ARDI step 2 at Crocodile River « What are the main resources of the catchment and the key information needed to support their management »

3. 4 Identifying key processes ("D for dynamics" in ARDI)

The third stage consists of listing the main processes that drive change in the territory in relation to the question (Figure 3). These processes can deal with ecological dynamics (i.e. vegetation transitions or water flow), economic dynamics (i.e. market price-changes, subsidies amount) or social dynamics (i.e. social cohesion, knowledge transfer). If the list is large, the facilitator asks the participants to rank the 10 main processes giving 10 to the most important one and 1 to the least. Then he sums up the scores given by each participant and selects the 5 processes that get the highest score. For these processes, diagrams are drawn to explain what forces are driving changes, with respect to which resources.

When dealing with ecological dynamics, participants may agree to the successive states taken by the vegetation and specify the factors which cause the transition from one state to another including the time required to move from one state to the next. The diagram can either be designed "in situ", or be a response to a proposal designed by an expert. In the two options, it must clearly distinguish the dynamics related to the human actions (effect of the techniques currently implemented), from natural dynamics (consequence of the abandonment of the uses). A similar diagram can be applied to the dynamics of water.

At the end of this phase, it is advised to review and revise the diagrams and to identify possible gaps. Three types of gaps may be identified. 1) An activity or a resource was identified but no participant carried enough knowledge about it. The group then agrees to

call upon an expert and nominates the person charged to identify and mobilize the expert. 2) An important actor was forgotten at the time of the preparatory phase and the group is concerned by this absence. The group then agrees to invite the person to the next phase. 3) An actor, a resource or a dynamic process are the subject of a total disagreement between two or several participants. The group then agrees on the choice of an expert and the type of information required from him in order to solve this dead-lock.

DF Drought frequency
CP Crop production
NL Nutrient leaching (N)
WH Water heating
CM Chemical modification
UPI Urban population increase
WA water abstraction

SFRA stream flow reduction activity
FR flow regulation
WP water purification

LS life support

** Text in red italics means that the term comes from the facilitator*

Comments and changes proposed during the process:
When eliciting the impact of afforestations on river, the acronym SFRA is chosen because it corresponds to the terms of the law.
When arguing on the importance of Wetlands in the interaction diagram, 2 new processes arised: FR and WP
No way to elicit the process that makes sense with Fauna and Flora, after a long discussion, the facilitator's statement Life support is accepted

Figure 3: ARDI step 3 at Crocodile River « What are the main processes that drive changes in the Crocodile Catchment that affect the river flow »

3. 5 Eliciting interactions

The last phase of the ARDI method consists of synthesizing answers to the three preceding questions by stressing the interaction between users and resources. It is a pivotal of the exercise since it leads to the conceptual model representing all interactions related to the tackled question. It is advised to devote more time to this phase since it generally takes one half-day for a simple diagram (3-4 direct actors, 3-4 resources), and one day for a more complex diagram (5-8 direct actors, 5-10 resources). The group must then answer the following central question:

How does each stakeholder use the resources and modify the processes?

The facilitator will begin this stage by distributing and summarizing the diagrams carried out during the previous stages, by making a particular effort of clarification if new people were integrated to the group. When the diagrams are relatively simple, he directly invites the participants to collectively construct an interaction diagram. For that, the facilitator puts the main resource in the middle of the diagram and proposes to position the direct stakeholders related to this resource. Each participant chooses, in turn, to add an interaction between a stakeholder and a resource or between a stakeholder and another stakeholder. He can either add a link on the collective diagram, or ask to add one of the stakeholders of the list not yet included on the collective diagram. Each new interaction suggested must include a verb which specifies the type of action that generates the link. The proposer must justify his choice and indicate, when he knows them, the type of information used by the actors to make the corresponding decision (i.e. I authorize a new allotment because the request for residences exceeded 50; I withdraw my flock from this paddock because it remains less than 300 kg of fodder; I will look for an agreement with the Regional Park because more than 30% of the inhabitants complain about the area covered by fallow lands). Finally, when all the arrows are drawn, the participants locate on the diagram the key processes identified at the previous stage, by writing down their acronym besides the arrow representing an interaction that is supposed to strongly affect them.

When the diagrams become too complex, it is preferable to proceed in a segmented fashion by cutting up the exercise into several phases. Two options are possible. If several stakes were clearly identified during the co-construction process, the facilitator proposes to carry out a diagram of interactions for each of these stakes and leads the procedure described in the preceding paragraph as many times as is necessary to complete the diagram. In this

case, he must take care that the resources and the stakeholders mentioned by the participants continue to relate well to the chosen stake, and in case of doubt, to clarify the considered link. If stakes are not clearly identified, the facilitator proposes to gather the resources into categories, and then constitutes working groups on the 3 or 4 categories which appear most important to the participants. In this case, it is necessary to add a phase of pooling and comparison between the 3 or 4 built diagrams.

The role of the facilitator during the "interaction step" is particularly important and delicate since he constructs an easily accessible and recognizable diagram at the same time as facilitating the interactions and inputs (taking care to avoid confusing representations or crossed arrows, etc). He needs also to ensure clarity of inputs from participants (whilst avoiding putting them in delicate or uncomfortable positions) and regularly revisit those inputs that are not integrated into the diagram (i.e. boxes without arrows), without forcing the participants too much). The facilitator simultaneously assumes three objectives: a) to gradually prepare a common diagram comprehensible to all, b) to identify clear and indisputable interactions, and c) to leave the possibility of repairing lapses of memory. Additionally, the facilitators role are to oblige each participant to reformulate their input so as to avoid uninformative verbs (i.e. the herd grazes, the farmer farms his field, the manager manages his budget) or to retain only the interactions which make sense according to the question (i.e. in an exercise on fire prevention and urbanization, the interaction between the cereal farmer and his crop field was restricted to ploughing the stubble after harvest, because it is the only one that impacts land sensitivity to fire).

This phase is generally the richest and most interesting of the co-modeling process, but to benefit maximally from this richness, it is essential to keep a record of the process of the construction of the four diagrams. There is specific value to knowing why and how a particular actor, or particular resource, or particular interaction, was mentioned, retained, eliminated or transformed. It is possible to use many means to reach this goal: audio recording (very comprehensive but very time consuming to analyze), a secretary dedicated to this task (very effective because they can quickly give an account of the sequence followed and how decisions were justified but it demands an additional person), the use of an interactive table or a digital camera allowing progressively to take a series of instantaneous diagrams with their construction (very demonstrative but requires either particular equipment, or a person partially dedicated to the exercise).

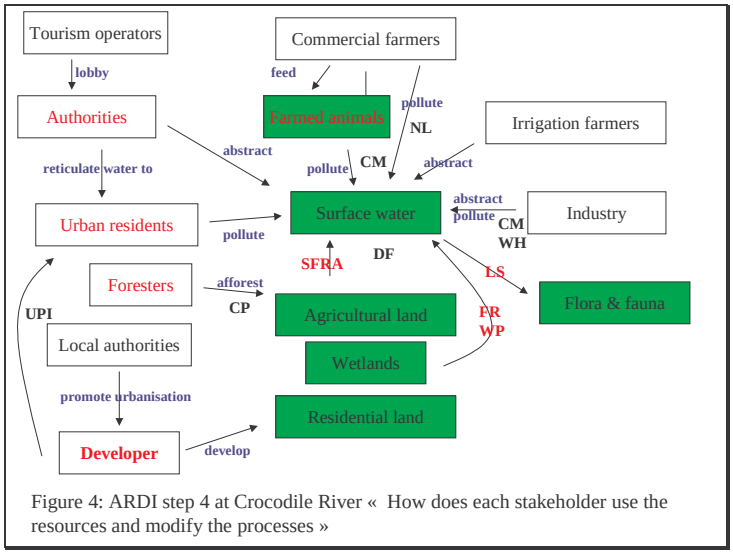


Figure 4: ARDI step 4 at Crocodile River « How does each stakeholder use the resources and modify the processes »

4. TAKING THE PERSPECTIVES FURTHER

The completion of these four stages leads to the establishment of a conceptual model. This model is a critical output of the ARDI process as it is a graphical representation of how the stakeholders perceive the system to function. This has fundamental implications for the next stages: designing and implementing a management plan for the territory based on the

collaborative established understanding captured in the diagrams. Two options arise for the working group: a) to work out a proposal for a management plan based on the conceptual diagram (concerted research plan, charter of sustainable development), or b) to develop a computer simulation model that will assist in decision making and dialog. In the first case, the thinking will be focused on the territory and its priorities of development, education and research. In the second case, the thinking will focus on the implementation of a computer model or a role-playing game to help stakeholders to transport themselves to the future and imagine and vision collectively adaptive co-management scenarios. In both cases, the ARDI method is valuable and useful as it works with a collectively established conceptualization of the territory and provides a concrete tool for applying the concepts of adaptive management.

ARDI method has many similarities with Problem Structuring Methods such as the use of a model as a transitional object, the emphasis put on the group process and the importance of facilitation skills (Eden & Ackermann, 2006). But it is concentrated on the preliminary issue conceptualization stage of modeling and on the visualization of a shared mental model as other methods developed for systems thinking (Hodgson, 1992; Richardson & Andersen, 1995).

ACKNOWLEDGEMENTS

The authors wish to thank the French Institute for Biodiversity (IFB) and the UNESCO-MAB programme for the financial support provided to French and Western African case studies, and the South African National Parks and Water Research Commission for the south African case study. They also want to acknowledge the participants to the many workshops organized in France, and in Western and South Africa to develop and validate this methodology.

REFERENCES

Bousquet, F., Barreteau, O., d'Aquino, P., Etienne, M., Boissau, S., Aubert, S., Le Page, C., Babin, D. and Castella, J.C., Multi-agent systems and role games : an approach for ecosystem co-management. In Janssen M (Ed), Complexity and ecosystem management: the theory and practice of multi-agent approaches, Elgar Publishers, Northampton, 248-285., 2002.

Collectif ComMod, Modélisation d'accompagnement. In Modélisation et simulation multi-agents : applications aux sciences de l'homme et de la société, Amblard, F. et Phan, D. (eds), Hermès Sciences, Londres, 217-228, 2006.

Costanza, R., Ruth, M., Using dynamic modelling to scope environmental problems and build consensus. *Environmental Management, 22, 183-195*, 1998.

Eden, C., Ackermann, F. , Where next for problem structuring methods. *Journal of the Operational Research Society* 57(7), 766-768, 2006

Etienne, M ., Le Page, C., Cohen, M.,. A step-by-step approach to building land management scenarios based on multiple viewpoints on multi-agent system simulations. *Journal of Artificial Societies & Social Simulations* 6(2), 2003 <http://jasss.soc.surrey.ac.uk/6/2/2.html>

Etienne, M., Companion modelling: a tool for dialogue and concertation in Biosphere Reserves. UNESCO-MAB, Paris, *Technical Notes 1, 44-52*, 2006.

Hodgson, A., Hexagons for systems thinking, *European Journal of Operational Research,* 59(1), 220-230, 1992

Holling, C., Adaptive environmental assessment and management, John Wiley, London, 1978.

Richardson, G., Andersen, D., Teamwork in group model building, Systems Dynamics Review, 11(2), 113-137, 1995

Van den Belt, M., Mediated modelling: a system dynamics approach to environmental consensus building. Island Press, 2004

Walters, C., Adaptive management of renewable resources, New York, McGraw Hill, 1986.

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 Evolutionary Bayesian Belief Network-based Methodology for Adaptive Water Management

R. Farmani[a] and D. Savic[b]

Centre for Water Systems, University of Exeter, Exeter, United Kingdom

[a](r.farmnai@exeter.ac.uk)

[b](d.savic@exeter.ac.uk)

Abstract: This paper presents an approach for constructing and testing a decision analysis process for adaptive water management under uncertainty. Water resources management as a complex dynamic system contains nonlinearities, feedback loops, and delays. Qualitative system dynamics modelling (e.g. causal loop diagram) is employed within a participatory integrated framework (integrating social, environmental and economic elements) to identify major drivers and their trends, potential evolutionary paths and their interdependencies, and also possible actions that can be taken to reduce impact of these drivers. An evolutionary Bayesian belief network-based methodology is developed to guide stepwise decision making during the transition process taking into account key uncertainties. Causal loop diagrams, as directed graphs, have no restrictions with feedback loops. Loops in causal maps are usually the result of dynamic relationships between variables across multiple time periods. However, Bayesian belief networks are hierarchical acyclic graphs, therefore have no means of handling feedback loops. The proposed methodology addresses this major shortcoming of Bayesian belief networks.

Keywords: System Dynamics, Feedback loop, Evolutionary algorithm, Bayesian belief network, Adaptive management, Uncertainty, Learning.

1. INTRODUCTION

Sustainable management of water resources in light of global and climate changes is one of the most pressing challenges of the 21st century. This requires approaches that take into account full complexity of the systems to be managed and the need to develop adaptive and integrated management approaches (Pahl-Wostl, 2006). This requires planning and managing water resources in a holistic manner. In order to succeed, it is important to take into account a wide range of (e.g. physical, environmental, economic, social and political) factors that impact on the water resources. It is equally important to identify the best way of linking these factors together and to simulate the interactions between them. Uncertainty is another important problem, which is an inherent feature of environmental systems. These systems are rarely well structured (Simonovic, 1996) as there is no definitive formulation, no true or false solution, and no test of a solution for these problems. This has earned them the title of wicked problems (Rittel and Webber, 1973). Causal Loop Diagrams (CLD) as system dynamics tool can provide a framework within which the environmental structure can be developed and the interactions and relationships among different variables can be investigated. System dynamics is important in understanding the cyclical behaviour of a system. In general, a feedback control system exists whenever the environment causes a decision that in turn affects the original environment (Forrester, 1958). System dynamics

introduces the possibility that a system may display non-equilibrium behaviour as it flips between positive and negative feedbacks. The result is much more complex patterns of movement over time (Stacey, 2002). Nowadays CLDs are mainly used for articulation of dynamic hypothesis of the system as endogenous consequences of the feedback structure (Sterman, 2000).

CLDs can be a good start for system modelling, however, in dealing with complex systems with high uncertainties other tools are required. Bayesian belief networks (BBNs) are used to simulate domains or systems containing some degree of uncertainty caused by imperfect understanding or incomplete knowledge of the state of the system (Jensen, 1996). BBNs have the advantage of dealing with uncertainties while avoiding overly complicated mathematical methodologies. BBNs are directed acyclic graphs; therefore transformation of CLDs with feedback loops to BBNs is not a straightforward process. In what follows, an evolutionary Bayesian belief network-based methodology is presented. The suitability of the developed integrated methodology is discussed in facilitating generation of robust management options as well as the way delayed feedback loops are handled.

2. ADAPTIVE WATER MANAGEMENT

Adaptive management is a systematic process for improving management policies and practices by learning from the outcomes of implemented management strategy (Pahl-Wostl, 2007). This requires incorporation of iterative learning cycles in the overall management approach. Considering and analyzing different hypotheses and scenarios about system behaviour under uncertain future development can be used as a guiding process in adaptive management. Figure 1 shows an adaptive management cycle including consideration of scenarios and hypotheses as learning process in an iterative policy cycle.

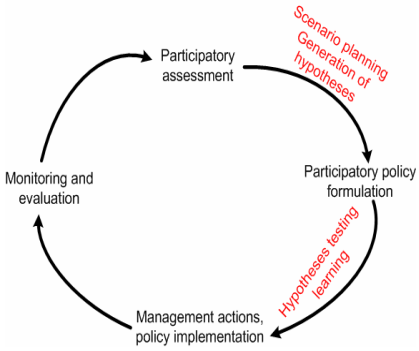


Figure 1. Adaptive management cycle

Adaptive management increases adaptive capacity by shifting linear decision making process to a cyclic learning process that iteratively integrates problem bounding and definition, policy formulation, implementation and monitoring in order to track and manage changes (Sendzimir et al., 2007). This requires a number of decisions along any path of change, the consequences of which are uncertain and evolutionary. Such consequences can be modelled in system dynamics using feedback loops that show ways in which a system can unexpectedly shift its behaviour. Feedback loop simply means that the outcome of a previous action is fed back as information that guides the next action in such a way that the discrepancy between the actual outcome and the desired one is reduced until it disappears to reach equilibrium state of behaviour for a system. A number of different forecast scenarios should be prepared to take into account unforeseeable events (Stacey, 2002). A loop can dominate the system's behaviour until accumulating influences suddenly allow another feedback loop to take over control. Even though feedback loops add to dynamic complexity of systems, all learning depends on them. As discrepancies between

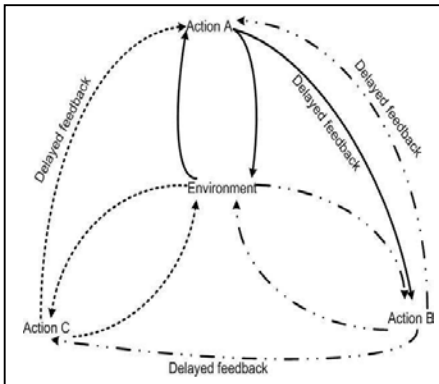


Figure 2. Delayed feedbacks as drivers of policy resistance and learning.

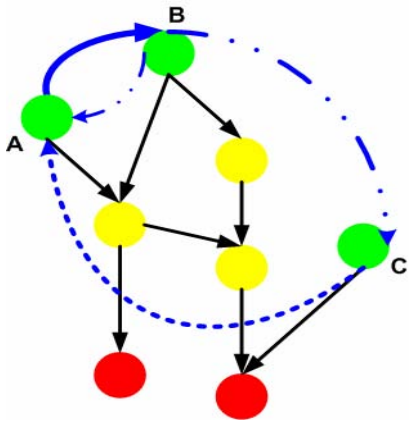
desired and actual states is perceived, actions are taken that hopefully will cause movement towards the desired state (single-loop learning). On the other hand double-loop learning results in more deep changes i.e. changes in mental models, goals and values (Sterman, 2006).

Figure 2 shows that each intervention has a consequence. Actions not only alter the environment and the future decisions, they also can have delayed effects that need to be addressed by other actions in order to restore equilibrium in the system.

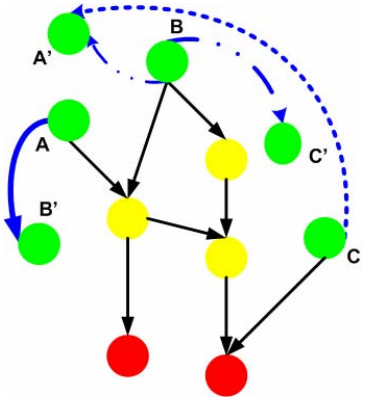
3. EVOLUTIONARY BAYESIAN BELIEF NETWORK METHODOLOGY FOR ADAPTIVE MANAGEMENT

Increasing uncertainties require a more adaptive and flexible management approach to realise a faster coping cycle that allows the rapid assessment and implementation of the consequences of new insights. Adaptive management can be defined as a systematic process for improving management policies and practices by learning from the outcome of implemented management strategies. Being adaptive thus means being able to constantly change internal structures in order to respond to external changes. This requires innovative approaches to facilitate improved learning and adaptation in addition to control (Pahl-Wostl, 2007). Robustness is a key criterion for good decisions under uncertainty (Rosenhead, 1993). The most effective form of adaptive management employs management programmes that are designed to experimentally compare selected policies or practices by evaluating alternative hypotheses (Gunderson, 1999). In general, there is no single solution for complex and uncertain problems. There are often trade-offs that require choices. Scenario planning is a strategic method that can be used to make flexible long-term plans. Scenarios represent the outcome of the feedback loops with complex interactions and long delays based on a set of assumptions about key driving forces. They assist in the assessment of impacts, adaptation and mitigation processes.

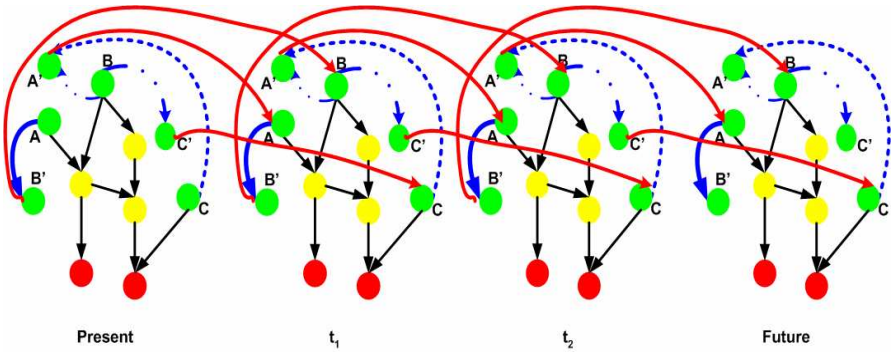
To learn effectively in a world of dynamic complexity when evidence cannot be generated through experiments, virtual worlds and simulation become the only reliable way to test hypotheses and evaluate the likely effects of policies. The virtual worlds are models or simulations in which decision makers can conduct experiment, rehearse decision-making and play. They can be physical models, role-plays, or computer simulations (Sterman, 2006). The proposed methodology, which is based on the integration of evolutionary multiobjective optimisation algorithm and Bayesian belief, facilitates design of robust and flexible management strategies through an iterative decision making process. The two software are linked via Microsoft Excel where all the data exchange takes place. In this methodology, first different management strategies are identified (Fig.3.a). This is followed by identification of future states of the system based on scenarios, which has been done by introduction of new nodes (nodes A', B' and C', Fig.3.b). Scenarios represent possible consequences and effects of each action solution on other aspects of the system through feedback loops (Fig.3.b). A Bayesian belief network is set up for each time step. In the simplest from, on one hand, this is similar to the temporal extension of BBN which means that the network structure or parameters do not change dynamically, but that a dynamic system is modeled. On the other hand, as it consists of time-slices (or time-steps), with each time-slice containing its own variables that are generated using EMO, it resembles single loop learning where only actions and strategies can be changed. However in complex systems with a large number of feedbacks, not only it models temporal nature of the problem, but also introduces changes to the next time step as they are identified in each time step. Changes here refer to those that will affect structure or parameters of the existing Bayesian belief network. From decision making point of view, the former deals with sequential decision making task while the latter, so called dynamic decision making task, is more concerned with controlling dynamic systems over time.



a. BBN with hypothetical feedback loops



b. hypothetical feedback loops replaced with BBN nodes



c. output from each time step is fed as input of next time step.

Figure 3. Handling feedback loops in BBN

The developed Bayesian belief networks are considered simultaneously in identification of robust decision paths (Fig.4). The outcomes of each time step are the inputs of the following time step (Fig.3.c). The trade-offs between different objectives are evaluated. The stopping criterion for the algorithm is defined as identification of a management strategy that is reinforced by other strategies enabling its growth and stabilization. The evolutionary based model facilitates this and identifies, based on the concept of survival of the fittest, the robust pathways in a co-evolving environment. Figure 5 demonstrates the main steps of the proposed methodology. The algorithm starts by initialising action or strategy nodes using randomly generated values from EMO software. This change will then have a knock on effect throughout all those nodes linked to it. In this way the impact on the whole system can be evaluated. The criteria for stopping this part of algorithm are that either several consecutive decisions support similar actions or a predefined large finite time

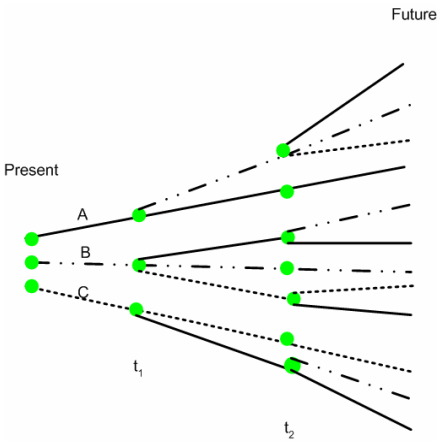


Figure 4. Decision tree.

horizon has reached. If the former criterion is not satisfied and depending on the information provided by additional nodes representing the impacts of the feedbacks, two possibilities exist. If additional nodes indicate no need for change in structure or parameters of the system, the next step action plans generated by EMO will be implemented otherwise the changes will be fed back to latest Bayesian network and the process will be continued. This process will be repeated for all the solutions generated by EMO. The evaluated results will be ranked based on their objective function values. The procedure will be repeated until no improvement is made on Pareto optimal front or maximum number of generations is reached.

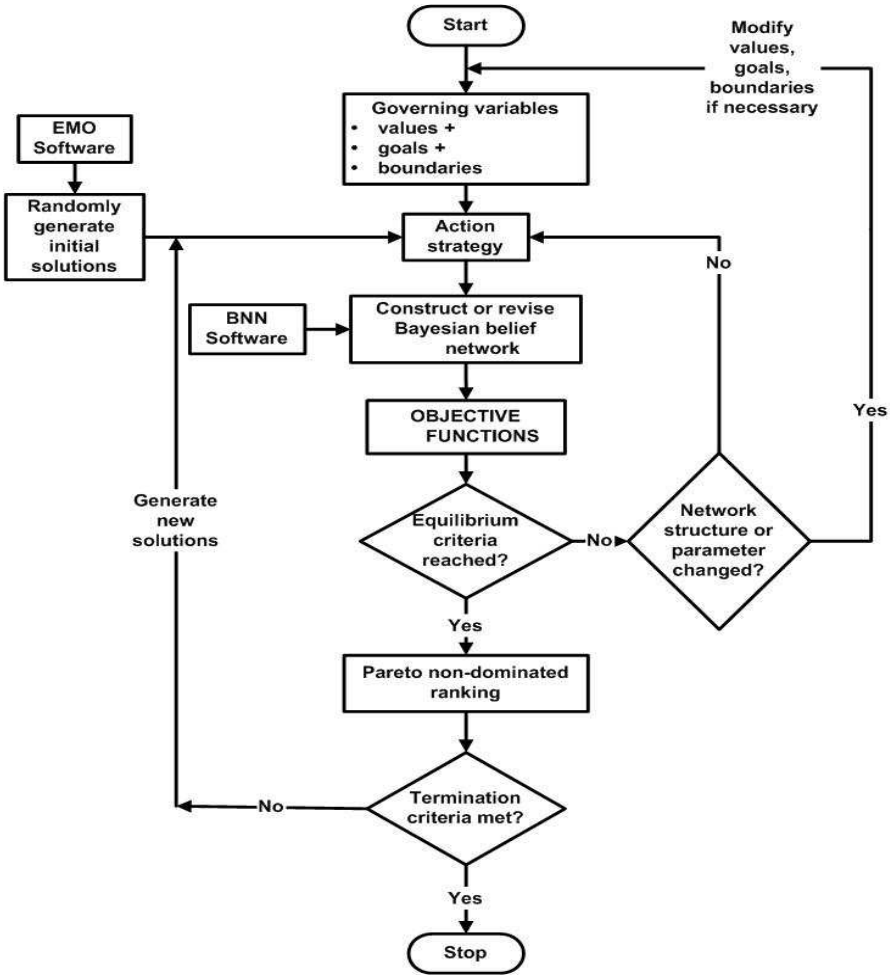


Figure 5. The proposed methodology

The methodology proposed in this work is not only anticipatory but also exploratory. Anticipatory in a sense that it starts with prescribed vision of the future and then works backwards in time to visualize how this future could emerge (focusing on long term). On the other hand it is exploratory as it starts in the present and explores possible trends into the future. This methodology is similar in a way to transition management (Rotmans et al., 2001) which involves long-term planning process in small and incremental steps. These planning and management methodologies take uncertainty and complexity as starting point rather than as closing entry; they take learning as guide rather than fixed goals and are co-evolutionary. Evolutionary planning and decision making process is aimed at different interventions at different levels in time and space (Rotmans, 2006).

Despite our efforts to present the methodology by application to a flood plain management problem, we were not able to quantify our developed conceptual models due to lack of data.

We are hoping presentation of it at the conference would results in some interest from participants and possibly a suitable case study to better illustrate the methodology.

4. CONCLUSIONS

High complexities due to nonlinearities, feedbacks and delays in environmental decision making problems require more advanced techniques. In this paper an evolutionary Bayesian belief network methodology is proposed to guide stepwise decision-making during the transition process taking into account key uncertainties. In the proposed methodology, complexities are considered as uncertain information and treated as elements of Bayesian belief networks. The effects of delayed feedbacks are modelled using scenarios and hypotheses. The outcome of each time step is fed back as input for next time step. Simultaneous consideration of the all time steps under different feasible interventions using evolutionary algorithm will result in a set of adaptive decision options that trade-off between different objectives.

The proposed decision analysis approach allows decision makers to use computer models to plan a wide range of feasible paths into the long term. Decisions made in this way are robust because they are adaptive as they are explicitly designed to evolve in response to new information.

ACKNOWLEDGEMENTS

This work was supported by the European Commission under the contract No 511179 for the Integrated Project NeWater.

REFERENCES

Forrester, J.W., Industrial dynamics: a major break-through for decision-making, *Harvard Business Review*, 36(4), 37-66, 1958.

Gunderson, L., Resilience, flexibility and adaptive management: antidotes for spurious certitude?, *Conservation Ecology*, 3(1), 1999.

Jensen, F.V., An introduction to Bayesian networks, 1 Edition, UCL Press, London, 178 pp, 1996.

Pahl-Wostl, C., NeWater newsletter, 2006.

Pahl-Wostl, C., Requirements for adaptive water management, *CAIWA book*, 2007.

Rittel, H.W.J., and M.M. Webber, Dilemmas in a general theory of planning, *Policy Science*, 4, 155-169, 1973.

Rosenhead, J., Rational analysis for problematic world: problem structuring methods for complexity, uncertainty and conflict, ISBN 0 471 92285 4, 1993.

Rotmans J., Tools for integrated sustainability assessment: A two-track approach. *The Integrated Assessment Journal, Building science and policy*, 6(4), 35-57, 2006.

Rotmans J., R. Kemp, and M.B.A. Van Asselt, More evolution than revolution: Transition management in public policy. *Foresight*, 3(1), 15-31, 2001.

Sendzimir, J., P. Magnuszewski, P. Balogh, and A. Vari, Anticipatory modeling of biocomplexity in the Tisza river basin: first steps to establish a participatory adaptive framework, Environmental Modeling and Software, 22, 599-609, 2007.

Simonovic, S.P., Decision support systems for sustainable management of water resources: 1. General principles, *Water International,* 21, 223-232, 1996.

Stacey, R.D., Strategic management and organisational dynamics: the challenge of complexity, ISBN 0-273-65898-0, 2002.

Sterman, J.D., Learning from evidence in a complex world, Public Health Matters, 96(3), 505-514, 2006.

Sterman, J.D., All models are wrong: reflections on becoming a systems scientist, *System Dynamics*, 18(4), 501-531, 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

The integration of Expert and Stakeholder Cognitive Models to support Environmental Monitoring

R. Giordano[a], **S. Liersch**[b], **M. Vurro**[a] **and V.F. Uricchio**[a]

[a] *National Research Council – Water Research Institute, via De Blasio, 5 – 70123, Bari, Italy (raffaele.giordano/michele.vurro/vito.uricchio@ba.irsa.cnr.it)*

[b] *Helmholtz-Centre for Environmental Research UFZ, Leipzig-Halle, (stefan.liersch@ufz.de)*

Abstract: The transition towards an adaptive approach to the water resources management (AM) is leading to a growing demand of information in environmental decision-making. In an adaptive approach, monitoring becomes the primary tool for learning about the system and assessing the management strategies. Nevertheless, professional monitoring faces several important challenges, especially in countries where financial and human resources are limited. In our work, the usability of local knowledge to support environmental monitoring has been investigated. Several shortcomings are hampering the use of local knowledge. The data credibility is among the most important ones. To address this issue, we propose a methodology based on the integration between stakeholders knowledge and expert knowledge. Cognitive modelling has been used to disclose individual perceptions and understanding of the system. The integrated cognitive model is used as basis for the development of an expert system able to assess the reliability of local knowledge. The methodology is under experimental implementation to support soil salinity monitoring in the Amudarya River Basin using local knowledge for a qualitative assessment. The preliminary results of the experimentation are described in this contribution.

Keywords*:* Adaptive management; Monitoring system; Local knowledge; Cognitive modelling; Expert system.

1. INTRODUCTION

The increasing awareness of complexity and uncertainty of the real world is posing new challenges to environmental resources management, and, consequently to the production of information to support decision making.
One approach, based on learning process, developed to face complexity and uncertainty is Adaptive Management (AM) (Holling 1978). AM approaches explicitly recognise the existence of uncertainty, and create an experimental management framework based on flexibility to test and refine a range of management approaches over time, based on a careful comparison of results. Using intervention as experiments, managers can learn-by-doing in a structured process. Learning in AM leads to a focus on the role of feedback from the implemented actions to stimulate a better understanding about the environmental system and its responses. Such feedback-base learning models stress the need for monitoring the discrepancies between intentions and actual outcomes (Fazey et al., 2005). Monitoring becomes the primary tool for learning about the system and assessing management strategies.
Thus, a monitoring system for AM has to be able to support the identification of changes in system behaviour. To this aim, the collection of information on trends plays a fundamental role. In fact, the availability of time series data of different variables allows the definition of the behaviour of the system variables and the trajectory of the system. The detection of trends facilitate the identification of system thresholds, i.e. breakpoints between two states of a system.

An important issue need to be dealt with, i.e. the cost of environmental monitoring. On one hand, the detection of trends in system dynamics requires the collection of long time series of data. On the other hand, in order to take into account the complexity and the strong interaction between different spatial and temporal scale, AM often results in a demand to monitor a broad set of variables, with prohibitive costs if the monitoring is done using only traditional methods of measurement, impeding the economic sustainability over time of the monitoring system.

The issue of monitoring cost is crucial in countries where, due to limited financial and human resources, monitoring may cover only small part of the territory, or the number of monitored factors is incomplete, or data gathering process may be too costly to allow collection of long data series for trend detection.

Therefore, developing an affordable monitoring system for AM requires substantial innovations in methods and approaches (Walters, 1997). In this contribution, the usability of local knowledge to support environmental monitoring is investigated.

Local environmental knowledge refers to the body of knowledge held by a specific group of people about their local environmental resources (Scholz et al., 2004; Robertson and McGee, 2003). Local knowledge should not be seen as the simple counterpart of the scientific knowledge; they can be combined as partialities of a whole knowledge, leading to a hybrid and broad view of local resources management issues (Robbins, 2003).

Many efforts have been made to utilize local knowledge in environmental monitoring and management. Robertson and McGee (2003) propose to utilize the memory of a local community to support the wetland rehabilitation in Australia. They demonstrate the role of oral history, integrated with established scientific knowledge, to lead to environment management that is in tune with the ecosystem dynamics. Scholz et al. (2004) integrate fishermen's knowledge in geospatial analysis to support marine protected area planning process in California. Their work aims to collect local knowledge to be integrated with scientific knowledge, in order to fill important data gaps. Interesting experiences demonstrate the capability of local communities to monitor soil degradation using their knowledge about plants species (Hambly, 1996). In several cases, local knowledge has been used to monitor the biodiversity at local level (Danielsen et al., 2000).

According to the scientific literature, several benefits related to the involvement of local communities in environmental monitoring for both the communities and the environmental management agencies. From the communities side, the benefits obtainable through the public involvement are main related to the promotion of the public awareness of environmental issues, the enhancement of collaboration and cooperation and the promotion of a "two-way" information exchange. On the other side, the environmental management agencies could increase the available information reducing the cost of information collection; they could base their strategies on a more integrated knowledge; the implementation phase could be facilitated since the conflicts could be reduced.

Nevertheless the use of local knowledge in environmental resources monitoring is still limited because of several shortcomings. Among them, the credibility of data collected by local communities is of outmost importance. Local knowledge is not subjected to a peer – review process of validation, nourishing the scepticism of both scientists and policy makers. Moreover, the local knowledge is qualitatively and unstructured, based on experiences and stories, and therefore not easily comprehensible for the decision makers and functional for the decision process.

To overcome these shortcomings and enhance the usability of local knowledge for environmental monitoring two important phases have to be considered, i.e. the validation and structuring of the local knowledge.

In this contribution the integration between experts and local communities cognitive models is proposed as a methodology to deal with these two issues.

The paper is organized as following. Section two describes the potentialities of cognitive modelling for local knowledge structuring and validation. In section three the preliminary results of the implementation of the methodology for soil salinity assessment are described.

2. COGNITIVE MODELLING FOR KNOWLEDGE STRUCTURING AND KNOWLEDGE VALIDATION

As reported in previous section, the structuring of the local knowledge is a fundamental step in order to enhance its usability to support environmental monitoring. To this aim, the potentialities of cognitive modelling to disclose individuals' understandings of the environmental system and its main properties are particularly interesting.
A cognitive model can be defined as a representation of thought process for how something works in the real world. Most of the techniques for cognitive modelling may be viewed as composed by three main phases: identify concepts, refine concepts and identify links. A common characteristic of these approaches is a focus on obtaining the views of people in the problem environment.
Starting from these premises, a methodology has been define to develop a cognitive model to be used as a basis for the design of a monitoring system able to integrate local knowledge as a source of information. The methodology is composed by three main steps, i.e. the structuring of local knowledge, the structuring of expert knowledge, and the integration among conceptual models.
The first step aims to disclose the perceptions of local communities about the system to be investigated and the environmental problem to be monitored. This step is based on the premises that, due to their knowledge and experiences, local communities are able to evaluate the state of environmental resources and to detect possible changes. Moreover, in our approach using local knowledge for environmental monitoring and management is not simply a "data-collection" operation, carried out by researchers and in which local people have a passive role. The basic idea is to involve local communities in a process of co-producing environmental information. Furthermore, to facilitate a long term engagement of local communities in monitoring, it's important to integrate the monitoring activities within the daily activities of the community. Thus, it becomes fundamental to start the design of the monitoring system defining the mental model used by the members of the communities to make a judgement of the state of the environmental resources. This mental model is structured in the local community conceptual model.
The definition of this conceptual model is based on two phases. In the first phase, semi-structured interviews are conducted involving members of the local communities in order to obtain preliminary ideas about the factors used by local people to assess the state of the environmental resources and causal links among them. The number of interviews to be made is determined considering the number of new concepts included in the model after each interview (Ozesmi and Ozesmi, 2004). The cognitive model is concluded when no new variables emerged after a number of interviews. As a result, a preliminary local community conceptual model is defined, which provides a preliminary list of factors to be monitored and their degree of importance. To this aim, the analysis of CM can provided information about the relative importance of the different variables, by analysing the complexity of the causal chain. Those nodes whose immediate domain is most complex are taken to be those most central and, thus, the most important according to the perception of the local communities. The members of the local communities are then involved in a group model session, in which the results of the interviews are used as basis for the debate. During the debate, the cognitive model and the importance degree for each factor are discussed until a consensus is achieved among the participants.
At the end of this phase, the local communities conceptual model is defined. This model allows to structure the local knowledge, making it easily understandable for the decision makers. The conceptual model is used as basis for the definition of the protocol for data collection. This protocol takes into account the factors highlighted by local communities and the terms used by them to describe their possible states.
The usability of local knowledge to support the decision making requires a validation phase.
In our work, the validation is based on the integration among local communities knowledge and experts knowledge. To this aim, local experts can be involved in a cognitive modelling exercise which aims to collect and structure their understanding about the considered environmental problem. Similarly to the previous step, the expert cognitive model is developed by semi-structured interviews and group discussion. In this case, the interviews are not focused to the factors used by expert to assess the state of environmental resources, but to the factors that, according to their opinion, can influence the state of the

environmental resources. Therefore, the conceptual model developed at the end of this phase aims to structure experts' understanding about the risk that the environmental problem happens.
The next step is to integrate the two cognitive models trying to define the possible impacts of "experts" factors on "farmers" factors. In our approach, an impact means that there is a relationship between the experts' concept and the farmers' concepts. Therefore a relationship could be found between the values assumed by experts and farmers concepts. The links have been defined according to the results of experts' interviews. These relationships can be used in the validation phase.
Therefore, the validation is based on the comparison between the assessment made by local communities and the assessment made considering the factors defined by the experts. To this aim, data from local community are collected according to the protocol. Using these data and the local community conceptual model, the state of environmental resources can be assessed. To be validated, this assessment has to be compared with the risk that a certain state of the environmental resources occurs. This risk is assessed considering the experts conceptual model.
Then, the experts' assessment and the farmers' assessment can be compared in order to define a degree of congruence. The congruence degree is assessed using the fuzzy semantic distances. Therefore, the formula (1) is applied.

$$S_d(1,2,x_i) = \left|\mu_1(x_i) - \mu_2(x_i)\right| \quad (1)$$

Where S_d (1, 2, xi) represents the semantic distance of the values of the element x_i in experts' assessment and local community's assessment; $\mu_1(x_i)$ represents the value of experts' assessment and $\mu_2(x_i)$ represents the value of local community's assessment. The congruence degree is, then calculated according to the formula (2).

$$CD(x_i) = 1 - S_d(1,2,x_i) \quad (2)$$

The congruence degree (CD) provides information about the reliability of the assessment made by local community, supporting the validation of local knowledge, as requested to enhance its usability in environmental monitoring. Therefore, if the value assigned by local community is highly congruent with the value defined using experts conceptual model, then the value can be considered as reliable. Otherwise, a suggestion to deepen the investigation is provided to the monitoring system managers.
In the next section, the preliminary results of an experimental implementation of the methodology to support soil salinity monitoring are described.

3. LOCAL AND EXPERT KNOWLEDGE FOR SOIL SALINITY MONITORING IN THE AMUDARYA RIVER BASIN

During the Soviet time it was planned to make Uzbekistan the largest centre of cotton production. Due to the arid climatic conditions this aim could be achieved only by the construction of large irrigation systems (UNDP, 2007). In the early 1990s, Uzbekistan accounted for about 20 percent of world trade and thus was the third largest cotton producer in the world (ERS, 2006). The inefficiency of the irrigation network, inadequate drainage systems, and intensive agricultural production were leading to severe soil degradation (salinisation). 55% of the land in the Khorezm oblast (the study area) is medium to severe salinised (UNDP, 2007). In order to reduce the degree of salinity and to increase the agricultural productivity the soils are leached before the vegetation period requiring large amounts of (not always available) water. Based on a forecast of water availability for the oncoming vegetation period, carried out by a national authority, certain amounts of water allowed to be used for leaching and irrigation are defined at the regional scale. The regional branches of the Ministry of Agriculture and Water are responsible to allocate the available water among the Water User Associations (WUA) leading to a competitive situation between WUA. Each WUA claims for water required for agricultural management according to the degree of salinisation. Hence, an adequate soil salinity monitoring system is required for a reasonable allocation of the available water resources.

Currently, the monitoring network is based on soil sampling stations where one station is representing an area of about 50 ha. The monitoring network is managed by the Hydromeliorative Expedition (HE) that is a branch of the Amelioration Expedition, a governmental agency. Soil samples are collected each year before the harvesting time and are analyzed in the HE laboratories in order to assess soil salinity. The data are used to develop a regional map of soil salinity, which plays a fundamental role in the definition of water allocation strategies in the Khorezm oblast.

Given the scarcity of financial resources of the HE, the sampling is not carried out using a well-structured monitoring network. Moreover, in order to reduce the number of required soil samples, a preliminary phase is carried out by HE aiming to define homogenous areas. The samples are taken and analyzed from each of the homogeneous areas and the obtained degree of salinity is then extended to these areas. The definition of homogeneous areas is accomplished on the basis of one parameter only – the visual assessment of plant growth characteristics during the growing season. The weak point in this approach is that plant growth is influenced by a variety of factors such as seed quality, agricultural management practices, climatic conditions etc., to name a few only. During our fieldwork, several people working in the management of the monitoring system, water managers, and chiefs of the WUAs were interviewed about the current monitoring system. According to their opinions, the soil salinity map is reliable only at the regional scale, while at the local scale the information provided are often not correct. This means that the water allocation among farmers could be wrong. According to the interviewees' opinions, these errors are due to the wrong definition of homogeneous areas.

Therefore, the aim of our work is to support soil salinity monitoring at the local scale using local knowledge for a qualitative assessment of soil salinity. This qualitative assessment can be used as basis for the definition of the homogenous areas. In fact, as several water managers said during the interviews, farmers are able to assess the soil salinity in a qualitative way using their tacit knowledge. The methodology described in previous section was applied to collect and structure this knowledge making it usable to support environment monitoring. The experimentation is focused on the Khorezm oblast, in the delta region of the Amudarya.

Therefore, we tried to capture the cognitive models of experienced farmers during several interviews. A preliminary cognitive model about factors that are considered in the qualitative soil salinity assessment was developed by augmenting and superimposing the individual cognitive models. The farmers' cognitive model is shown in figure 1. In this model, the concepts forming the tacit knowledge of experienced farmers are included. This tacit knowledge allows farmers to assess qualitatively the degree of soil salinity.

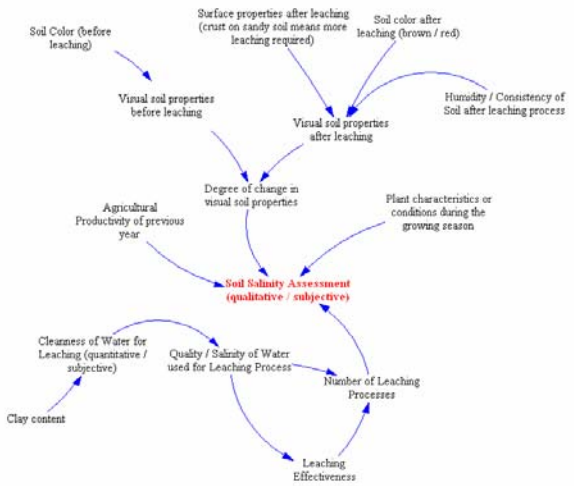


Figure 1. Cognitive model of qualitative soil salinity assessments.

Furthermore, a debate session was organized with all interviewed farmers in order to validate the cognitive model and the importance degree of the factors. At the end of this section, an agreement was reached about the definitive list of factors to be considered. Therefore, the soil salinity degree based on farmers' knowledge is assessed using the formula (3).

$$FA(x_i) = \sum_j f_j \times w_j \quad (3)$$

Where, $FA(x_i)$ is the soil salinity value for the element x_i, f_j represents the value of the j-th factor, w_j represents the importance degree of j-th factor according to the farmers conceptual model.

As reported in previous section, the conceptual model is fundamental to define the protocol for data collection. Therefore, a questionnaire was developed considering the factors highlight by experienced farmers and their possible states. The questionnaire was developed together with farmers in order to make it easily understandable. The collection of data from farmers is going to be started in next months, after the harvesting time. At the current state of the experimental implementation the process for data collection was discussed during the debate session with farmers. WUAs technicians are involved in collecting data from farmers during their normal activities. The data collection uses the field as spatial basis. Therefore, using (3), a map of farmers soil salinity assessment can be developed. A fuzzy linguistic variables "soil salinity" is defined. This variable can assume four different values, i.e. highly salinised, salinised, relatively salinised and not salinised. The values to be assigned to the variable are similar to the linguistic terms used by the farmers.

To validate the local knowledge, several local experts working on water and soil management were interviewed. The interviews aimed to collect and structure the understanding of the experts concerning the definition of soil salinity risk. A cognitive model was developed using the results of the interviews (figure 2).

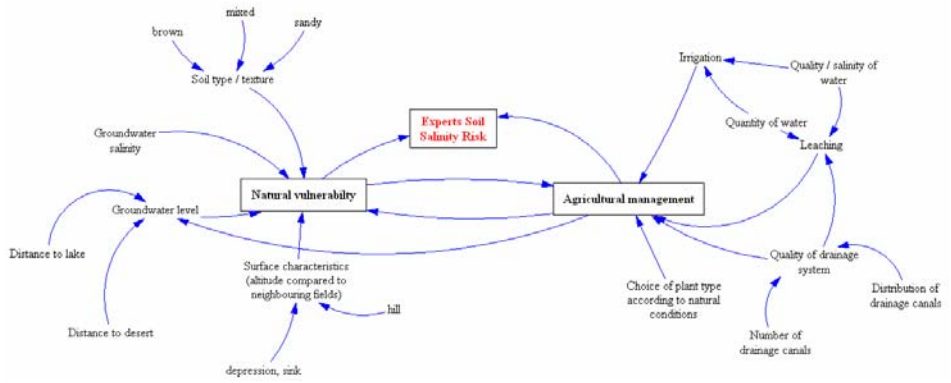


Figure 2. Cognitive model for soil salinity risk assessment

In this model, all the factors taken into account by experts to assess the soil salinity risk have been described. The importance degree of the factors was defined and validate by participants. Therefore the soil risk can be assessed using experts' knowledge according to the formula (4).

$$EA(x_i) = \sum_j e_j \times w_j \quad (3)$$

Where, $EA(x_i)$ is the soil risk for the element x_i, e_j represents the value of the j-th factor contained in the experts cognitive model, w_j represents the importance degree of j-th factor. All the factors contained in the expert cognitive model can be assessed considering "traditional" environmental data, such as groundwater level, groundwater salinity, soil type, drainage system, etc. Therefore, a map for soil salinity risk can be develop using this information. This map shows for each field the value of soil salinity risk. A fuzzy linguistic variable "soil salinity risk" is defined. At the current stage of experimental implementation, the collection of data for the factors contained in the experts cognitive model is started. The soil risk map will be ready before the initial phase of collection of data by farmers. This map will be stored in HE office and it will be used by them to validate the farmers' information, following the procedure described in previous section.

The integration between farmers' and experts' cognitive model is represented in fig. 3.

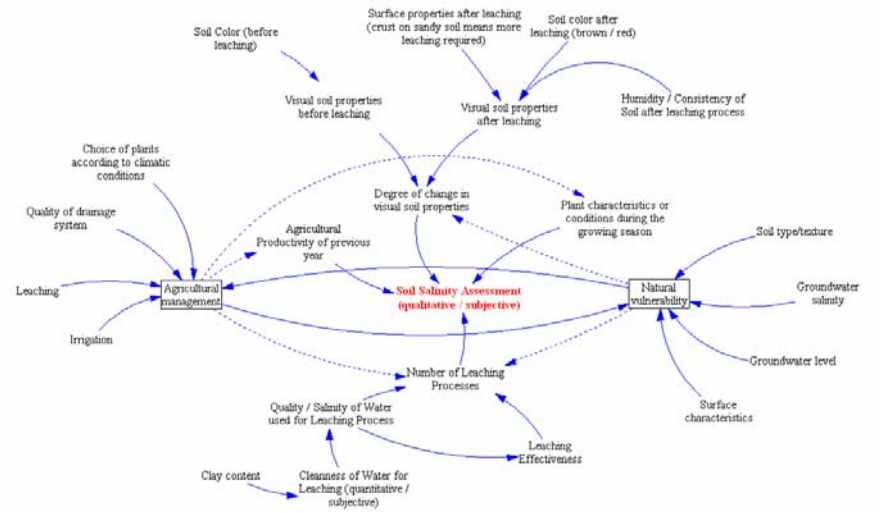


Figure 3. Integrated cognitive model

The validation of farmers' knowledge is made comparing, for each field, the value in farmers soil salinity map and the value in soil risk map. The comparison is done by HE using a prototype of a software developed by us able to produce the two maps and to evaluate the congruence degree for each field. At the current stage of the experimentation, the prototype has been developed and presented to HE to collect feedback. In next months the prototype will be installed in HE office in Urgench (Khorezm).
As reported in previous section, the congruence degree (CD) provides information about the reliability of farmers' salinity assessment, supporting the validation of local knowledge, as requested by the water managers and people working in the monitoring system management. Therefore, if the value assigned to the field x_i in the farmers' map is highly congruent with the value in the soil risk map, then the value can be considered as valid. Otherwise, a suggestion to deepen the investigation is provided to the monitoring system managers.
At the end of this process, the validated farmers soil salinity map is used to define the homogenous areas, and, following the current monitoring practices, samples will be taken in each area in order to create the soil salinity map.

4. CONCLUSIONS

The development of an affordable monitoring program to support Adaptive Management involves substantial, scientific innovation in both method and approach. Particularly interesting are methods and tools able to facilitate the integration of different sources of information. In this contribution, local knowledge is proposed as an interesting alternative source of information to support environmental monitoring. Nevertheless, local knowledge cannot be used by decision makers as it is proposed. A structuring and validation phase is needed.
The potentialities of cognitive modelling as a tool to support the structuring of local knowledge, which is tacit and not easily understandable for decision makers, are investigated in this work. The methodology is applied for the monitoring of soil salinity in the Amudarya river basin. A farmers cognitive model have been defined which highlight the factors used by members of the local community to make a qualitative assessment of soil salinity, i.e. the basis of their tacit knowledge. The design of the protocol for data collection has been designed considering these factors.
To facilitate the validation of the local knowledge, an expert system for the assessment of soil salinity risk has been defined using the knowledge of local experts. The integration among farmers and experts cognitive models and the comparison between the two assessments facilitate the local knowledge validation.

ACKNOWLEDGEMENTS

This work has been done within NeWater project, supported by the European Commission - EC contract No. 511179 GOCE.

REFERENCES

Campbell, B., J.A. Sayer, P. Frost, S. Vermeulen, M. Ruiz Pérez, A. Cunningham, and R. Prabhu, Assessing the Performance of Natural Resource Systems, *Conservation Ecology* 5(2): 22. [online] URL: http://www.consecol.org/vol5/iss2/art22/, 2001.

Danielsen, F., D.S. Balete, M.K. Poulsen, M. Enghoff, C.M. Nozawa, and A.E. Jensen, A simple system for monitoring biodiversity in protected areas of a developing country, *Biodiversity and Conservation*, 9: 1671-1705, 2000.

ERS, United States Department of Agriculture. Economic Research Service. http://www.ers.usda.gov/Briefing/Cotton/trade.htm, 2006.

Fazey, I., J.A. Fazey, and D.M.A. Fazey, Learning More Effectively from Experience, *Ecology and Society*, 10(2), 4. [online] URL: http://www.ecologyandsociety.org/vol10/iss2/art4/, 2005.

Fraser, E. D. G., A. Dougilla, W.E. Mabeeb, M. Reeda, and P. McAlpinec, Bottom up and top down: Analysis of participatory processes for sustainability indicator identification as a pathway to community empowerment and sustainable environmental management, *Journal of Environmental Management*, 78: 114-127, 2006.

Hambly, H, Grassroots indicators: measuring and monitoring environmental change at the local level. ILEIA Newsletter 12(3): 14 – 20, 1996.

Holling, C.S. (ed.), *Adaptive Environmental Assessment and Management*, John Wiley and Sons, New York, 1978.

Kuchar, J.K., D.S. Walton, and D.M Matsumoto, Integrated objective and subjective hazard risk in decision aiding system design, *Reliability Engineering & System Safety*, 75: 207-214, 2002.

McDaniels, T L, and R. Gregory, Learning as an objective within a structured risk management decision process, *Environmental Science & Technology* 38 (7), 1921-1926, 2004.

Oudwater, N., and A. Martin, Methods and issues in exploring local knowledge of soils, *Geoderma*, 111: 387-401, 2003.

Ozesmi, U., and S. Ozesmi, A Participatory Approach to Ecosystem Conservation: Fuzzy Cognitive Maps And Stakeholders Groups Analysis in Uluabat Lake, Turkey, *Environmental Management*, 31 (4), 518-531, 2004.

Robbins, P., Beyond Ground Truth: GIS and the Environmental Knowledge of Herders, Professional Foresters, and Other Traditional Communities, *Human Ecology*, 31(2): 233-253, 2003.

Robertson, H. A., and T.H. McGee, Applying local knowledge: the contribution of oral history to wetland rehabilitation at Kanyapella Basin, Australia, *Journal of Environmental Management*, 69: 275-287, 2003.

Rosenhead, J., and J. Mingers, A New Paradigm for Analysis. In Rosenhead, J., Mingers J. (eds), *Rational Analysis for a Problematic World*, John Wiley and Sons, Chichester, UK, 2001.

Scholz, A., K. Bonzonb, R. Fujitab, N. Benjaminc, N. Woodlingd, P. Blackb, C. Steinback, Participatory socioeconomic analysis: drawing on fishermen's knowledge for marine protected area planning in California, *Marine Policy* 28: 335-349, 2004.
Publishing, London, UK. ISBN: 1 84339 038 8, 2004.

UNDP, Water, critical resource for Uzbekistan's future. Publication in support of the Millenium Development Goals. Goals 7: Ensure environmental sustainability. United Nations Development Program Country Office for Uzbekistan. Tashkent, 2007.

Walters, C, *Challenges in adaptive management of riparian and coastal ecosystems*, Conservation Ecology [online]1(2):1. Available from the Internet. URL: http://www.consecol.org/vol1/iss2/art1/, 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

Participatory process management

J. Krywkow [a], **M. Hare**[b]

[a]*University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands*
(j.krywkow@ctw.utwente.nl)

[b]*Seecon Deutschland GmbH, Westerbreite 7, 49084 Osnabrck, Germany*
(matt.hare@seecon.org)

Abstract: Although the process of public and stakeholder participation continues to be intensively investigated and discussed in academic circles, the implementation of participatory methods in practice remains problematic. This can be attributed to the lack of knowledge transfer on the one hand, and the general underestimation of participatory approaches in planning processes on the other.
A possible solution - participatory process management - is introduced in this article. Participatory process management means that all participatory activities are embedded in the overall planning activities of a project. The most significant criteria for a participatory process are identified as 'objectives', 'constraints' and 'process' which together form a framework for combining generally applicable methods with local constraints and the objectives of a project. The main elements of the participatory management framework introduced here are levels and classes of participation and a generic process scheme including monitoring and evaluation of participatory processes. This work is based upon long-term experiences of consultants and scientists. However, the insights from the InterReg project TRUST are particularly valuable and confirm the hypotheses that different water management projects are comparable in terms of their participatory process performance. The participatory management framework is a step forward in closing the gap between scientific knowledge about participatory methods and their applicability in practice.

Keywords: participatory management, process planning, guideline for the use of methods, objectives, constraints, choice of participatory methods

1 INTRODUCTION

The incorporation of participatory processes in water management and planning is often poorly conceived, and the impact of stakeholder participation on an entire planning process is frequently mis-interpreted or underestimated. With the release of the European Water Framework Directive [EU, 2000] public and stakeholder participation is prescribed, and the synthesis of expert and lay knowledge is recommended.
Best practice guides on the application of participatory methods have been produced in plenty over the last two decades. However, in many cases a systematic and complete guideline is missing or a specific focus, for example on social learning, communication, democratisation and others, may superpose other significant issues in participatory processes. Although a comprehensive literature review has been done, this article only allows for a few examples:

Ridder et al. [2005] focus on social learning process, give practical advise on how to approach stakeholders and the public, and how to communicate with them; provide water managers with detailed descriptions of chosen participatory methods; and provide a number of conflict-solving approaches. Furthermore, a general structure of participatory processes, a procedural flowchart and an approach to monitor and evaluate the participatory process are identified. However, consis-

tency between methods, structure, monitoring and procedure is rather weak. Procedural schemes are mechanistic and not flexible. The descriptions of the general framework and specific methods and how-tos could be more balanced. One section lists and describes particular methods with reference to a particular phase of a participatory process. However, there is no reference to the 'levels of participation' in which these methods reside. For non-experts this might be confusing, since GIS-methods stands beside role playing games, maps, group model building and others. This book " Learning together to manage together" gives some valuable advise for those individuals and groups who want to know more in how to implement social learning processes in a planned participatory process.

Elliott et al. [2005] provide the reader with general guidelines for the design and realisation of a participatory process, and an in-depth description of a chosen number of participatory methods. The general guidelines give valuable advise for participatory managers, however, a systematic approach in terms of levels of participation as well as a general procedural guideline is missing. Costs and effort are discussed in detail, but oscillate between general remarks and pedantic recitals of how much paper, pencils or word processors are required. The description of participatory methods is detailed, however, the choice is elusive. GIS, role-playing games, mental mapping, group model building and other methods are missing. This book can be valuable, if particular methods are required, and as a source for particular tips and guidelines.

Wates [2000] is a valuable source of practical knowledge about numerous participatory methods that are especially useful for community planning. Some general hints of how to design and implement a participatory process are given, and an interesting 'participation matrix' was developed. However, the general principles have no structure and are rather superficially described. The methods part however, reveals sound knowledge and experience of the author, although it is not complete. Many, sophisticated methods such as group model building, mental mapping or similar are missing. Moreover, there is no system that relates the methods to a general frame work.

Bousset et al. [2005] is another recent guideline of how to apply participatory methods. The choice of methods is limited, and there is no overall framework in relationship to the chosen methods. However, this document holds a comprehensive description in how to implement monitoring and evaluation of a participatory process, including stakeholder feedback.

This short review is far from being complete, but the examples above demonstrate that much work has been done to produce valuable guidelines supporting 'participatory managers' in their search for designing and implementing a participatory process in their planning activities. On the other hand it reveals the complexity of integrating planning procedures such as water resource management, community planning or agricultural projects with participatory processes. This way many guidebooks are not complete or focus on particular topics such as social learning or on particular sectors such as community planning. Moreover, the consistency between a general framework, levels and classes of participation, a process scheme, the applicability of methods as well as the application of evaluation and monitoring is missing. Generally, most authors avoid generalisations that go beyond their sectoral approach. In other words case-study specific or sectoral constraints determine the guidelines to a great extent. This however, makes it more difficult for planners and managers to transfer the guide-book knowledge to their specific projects.

Moreover, planners and managers are often not trained in the field of participatory processes, and thus overwhelmed with the multitude of methods from which to choose or simply do not know of their existence. They may also lack knowledge of the implications of employing particular methods within a planning process. As a result, methods are selected based upon the experience of the responsible authority, which is often limited, and does not permit space for experimentation and expansion of available knowledge. In other words, contrary to well-educated engineers, planners or ecologists there are usually no properly-educated 'participatory managers'. The role of the participatory manager is often taken over by communication officers, engineers, planners or ecologists, who may not be explicitly trained to manage a participatory process.
The following sections present an approach that may support water managers in choosing from a

variety of available methods, find the balance between local specifics and common procedures, as well as between domain and expert knowledge such as engineering methods, ecological impact assessment and participatory approaches.

2 ASSUMPTIONS, INDICATORS AND FRAMEWORK

The design of an appropriate participatory management strategy can be characterised by three distinct indicators: (1) process (2) constraints and (3) objectives. *Process* implies the available methodology its costs and capability as well as the consequences of applying particular methods [Krywkow, 2007]. *Constraints* refer to the boundary conditions of a particular case study. These conditions include:

- the physical environment such as land use, size of a river basin, climate and weather, geology, slope and others;
- the stakeholders and lay people who are involved in a particular land-use activity or have particular interest in the management of the region under investigation;
- the available resources: budget, time and staff;
- legal constraints such as planning permissions, the right of the public to comment/object to planning proposals;
- cultural and behavioural differences which distinguish countries or regions.

Objectives are significant indicators of a participatory process. Objectives of the planning process and its participatory process have to be distinguished clearly. The latter may include knowledge elicitation, problem identification (of the overall planning process), conflict resolution, seeking consensus, finding support for maintenance. In other words, the objectives of the participatory process determine to a large extent the combination and application of participatory methods. Project goals may however, include sustainability, improving the infrastructure, increasing safety, improving biodiversity, improving drinking water quality, changing resource management and many more. Within a project it is crucial to distinguish between the overall project goals and the goals of the related participatory process. Commonly, the goals of a participatory process integrate into the overall project goals.

3 LEVELS OF PARTICIPATION AND CLASSES OF METHODS

From a methodological point of view a participatory process may reach four different levels of participation: (1) *information provision*; (2) *consultation*; (3) *active involvement* and (4) *social learning*. Levels of participation do not indicate a sequence for the application of methods, but the quality of interaction between managers, experts and lay people. For instance, information provision is a one way communication from managers to the public or stakeholders. Whereas consultation already includes a dialogue, and social learning an intensive exchange of opinions, knowledge and views that may result in a change of perspectives among individuals and groups at the end of a process [Arnstein, 1969; Mostert, 2003].
In addition to the levels of participation Hare and Krywkow [2005] summarised the main participatory methods in classes and displayed those in relationship to the levels of participation (figure 1). It became necessary to subsume methods into classes, since many different names exist for similar methods. The classes then enable users to evaluate methods that are not listed in Hare and Krywkow [2005]. Classes and levels of participation are a guideline for participatory managers. However, additionally Hare and Krywkow [2005, pp.21-48] provide potential users with an overview of indicators of methods such as: cost/effort share, expertise, moderator skills, user mode and computer application. Moreover, the effort involved in applying various methods in terms of staff resource, time, tools and miscellaneous costs is estimated in terms of in preparation, execution and analysis phase. This 'catalogue of participatory methods' together with the

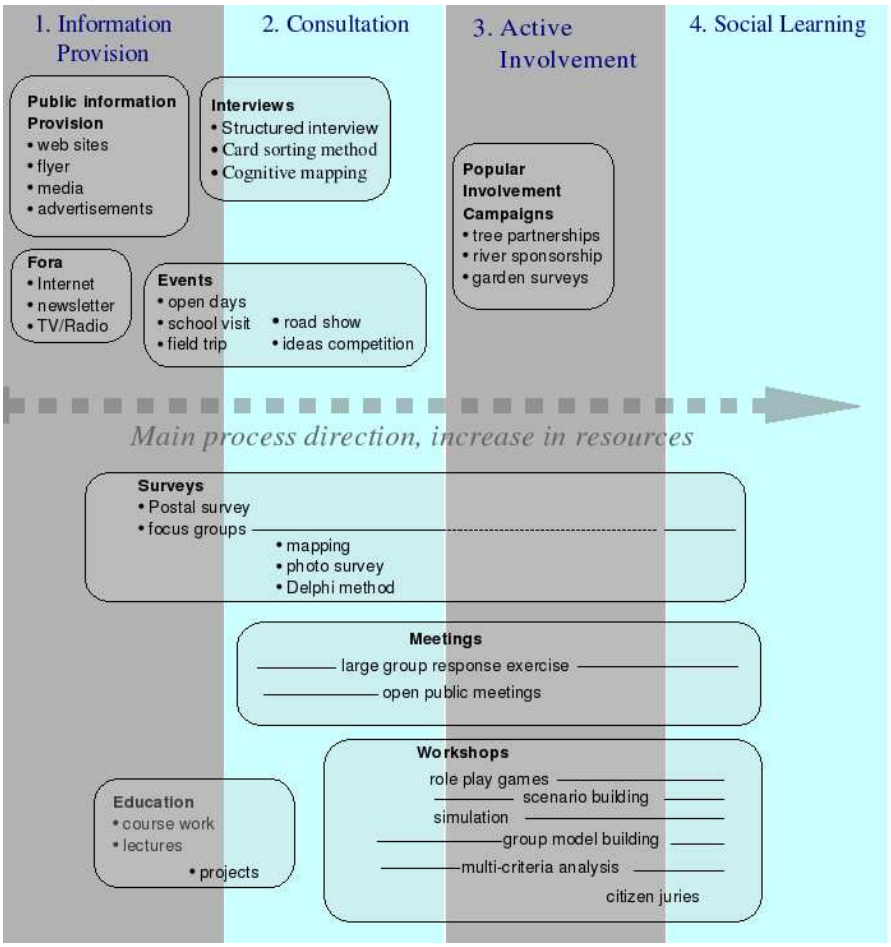


Figure 1: Levels and classes of participatory methods [Hare and Krywkow, 2005, p.19]

levels/classes overview enables the user to determine the applicability of methods within a given project.

4 PHASES OF A PARTICIPATORY PROCESS

Throughout the collaboration phase among water managers, scientists and consultants in the Inter-Reg III project TRUST[1] [Krywkow et al., 2007] it emerged that levels, classes and the catalogue of participatory methods as described above were, in practise, not easy for water managers to comprehend. The link between the process indicators and practical constraints was not strong enough. For this reason Krywkow [2007] developed a first draft of a universal framework based upon *phases* of participatory processes in water management projects. The TRUST project revealed that although the five case studies involved have different planning themes and objectives, the course of each participatory process has a similar pattern. In this way a comparison of the various participatory processes is possible. This enabled scientists and consultants to develop a set of universal, process-oriented guidelines for designing and conducting a participatory process in water management projects.
Independent of its constraints, a participatory process can be partitioned into four phases:

1. *preparation phase*: problem analysis, stakeholder analysis, resources analysis, goals analysis, drafting a participatory plan;

[1] http://www.trustpartners.org/

2. *publication phase*: introducing the public and stakeholders to the planning objectives and problems, envisaged solutions and measures including possible impacts on the social and physical environment;
3. *dialogue phase*: additional and deeper information provision, knowledge elicitation, education, detecting planning design errors and yet unknown side effects;
4. *response phase*: education, social learning, recruiting volunteers, scenario building, model validation, finding consensus or compromise, adjusting planning goals.

The list above indicates a number of tasks for the water managers as well as participatory methods that are applicable in one or more specific phases of a water management project.

4.1 Preparation phase

Before a participative process can be initiated, a sound preparation is required including the analysis of stakeholders, resources and budget as well as setting of the agenda (participatory plan). The course and, at the end, the success of a participatory process both depend on these factors. When preparing the process, managers must be aware of maintaining the balance between a well organised schedule and sufficient flexibility for applying participatory methods.

Co-planning or expert planning. Expert planning is a planning procedure where a planning draft is designed by experts (planners, engineers, ecologists, etc.) and presented to lay people. Co-planning is a process where a project idea may emerge among lay people, and is (co) designed by experts and lay people together. If an experienced planning authority can propose a planning draft that is open for discussions with the public and stakeholders, it can be as transparent and adaptive as a planning draft that is developed by a group consisting of both experts and lay people. In the case of a co-planning approach planners, however, must maintain control over the participatory process. Co-planning requires stakeholder identification before a planning draft can be designed. Otherwise, there is a risk that some groups and individuals may be excluded.
If the project uses a co-decision approach, the problem analysis must be conducted together with the participating stakeholders. If the project is an expert approach, a problem analysis comes first, and subsequently the relevant stakeholders must be found. In both cases, however, stakeholder as well as problem analyses must be flexible and repetitive, and should reach into the next phases of the participatory process.

Goals and limitations. Similar to a communication plan goals are a significant guideline for designing and implementing a participatory process. First, there are several types of goals:

1. goals of the overall project such as improving safety, environmental protection, sustainability, better access to and functionality of recreational areas, better infrastructure and many more. Within these main goals there may exist subgoals (e.g. safety may consist of risk reduction and increased protection). Note that goals and sub-goals of stakeholders and the public may significantly differ from the planning objectives of the planning organisation or the consortium involved in a particular project. Therefore, it is crucial to identify the goals of all involved stakeholders throughout a participatory process;
2. goals of the participatory process such as developing consensus or a compromise, satisfying stakeholders, increasing the reputation of the responsible organisation, identifying the appropriate measures, detecting undesired side effects, designing a transparent, democratic participatory process.

Goals of the participatory plan usually depend on the overall project goals. At the end the participatory process is supposed to support the project or planning goals. However, project goals may not be confused with the goals of a participatory process. Goals may be defined with reference

to the planning organisation's own goals, the goals of the local, regional, national or European governments and of course the goals of stakeholders and lay people. Limitations that cannot be modified throughout a participatory process may be legal, environmental or budget constraints and limited space. Since budget can be a limitation, it is advantageous to not only include participation as part of a planning process from the beginning, but also to allocate sufficient budget for the participatory process throughout the design phase of the planning process. It is helpful that all involved parties of the participatory process are aware of all the goals and limitations.

Problem identification. As soon as a first planning draft or concrete ideas for a project are available, the main issues as well as potential problems and side effects that are relevant for the planning process from the point of view of the planners should be identified and documented. This helps to avoid surprises at initial stakeholder meetings.

Stakeholder analysis. There are two purposes of stakeholder analysis: to identify the possible stakeholders in a project, and to categorise them according to a framework that can allow you to make decisions about how to involve them, or not, as the case may be. Categorisation of stakeholders also helps water managers to make explicit the assumptions about the nature of the stakeholders so that within the planning organisation the involved parties can discuss the basis of these assumptions. Stakeholder analysis is essentially a three stage, iterative process:

1. Identify stakeholders
2. Categorise stakeholders
3. Select stakeholders.

Communication. Sufficient communication is a crucial requirement for a participatory process. However, participation is more than communication. Indeed, a *communication plan* is helpful for projects that include multiple planning organisations and stakeholders. With a communication plan managers can assign responsibilities, identify tasks, provide stakeholders and the public with key contacts, keep track of what was said and written, and more. A communication plan should have a clear objective. Objectives of the communication plan must be compatible with the goals of both the overall project and the participatory process.

Tasks and resources. Before the actual participatory process starts, resources and tasks have to be allocated. Funding, expertise and man power are limited resources. The participatory process must accompany the planning process. If participation is both a vital part of the entire project, and is initiated at the outset of a planning process, then it should be feasible to allocate resources such as available staff, costs for professional advice, computer capacity, locations for meeting, printing costs etc..
Some methods that involve, for example, the use of complex models and analysis tools are not always easy to apply, and universally understandable. External experts are helpful, and despite primary costs, their recruitment can help to spare financial resources and increase the level of participatory applications. An estimation of costs of and effort required by particular methods is provided in Hare and Krywkow [2005, pp. 21-49].

Monitoring a participatory process. Monitoring and evaluation can support a successful participatory process, and may help to avoid process failure. Therefore, it should be prepared and initiated as early as possible. A planning sheet such as described in Krywkow et al. [2007] is an appropriate means for recording a participatory process. This approach serves two purposes: 1) a planning sheet can be seen as a logbook for a project and its participatory process, that helps planners to keep track of all phases of the project; and 2) regular stakeholder feedback can be planned and incorporated in the process from the outset. When setting the agenda of the participatory process (participatory plan), dates for monitoring activities should be incorporated. Stakeholder feedback on the participatory process may not be confused with feedback on planning options. An

in-depth description of monitoring and evaluation of participatory processes is provided in Rasche et al. [2006].

Designing a participatory plan. Once problem identification, stakeholder identification, goals, resources and a communication plan, are settled, a first draft of a participatory plan can be designed. Planners are in a dilemma, since stakeholders might expect both a well-organised and efficient agenda as well as sufficient flexibility to incorporate new stakeholder perspectives and eventually new criteria and side effects. In other words water managers need an adaptive management agenda both for the entire planning as well as for the participatory process. However, a number of activities can be planned well ahead:

- *Information provision*: managers have to choose the means of informing the public and stakeholders such as websites, announcements in newspapers, face-to-face contacts, flyers, posters, etc.;
- *Surveys and interviews* can help to understand the perspectives and interests of stakeholders and the public;
- *First meeting with stakeholders or the public*: the main purpose should be providing more information in greater depth and knowledge elicitation. This meeting is important, since it can determine the further course of the project;
- *Miscellaneous activities*, depending on the results of the first meeting: space should be given for additional activities such as site visits, training for volunteers, popular involvement campaigns, survey or events ;
- *Response meeting*: This meeting should be used to display the results of knowledge elicitation and consequences of chosen solutions on the physical and social environment, as well as introducing new solutions and discussing them, if necessary;
- *Further activities* such as workshops, voluntary work, educational activities, etc.;
- *A final event* such as an on-site festivity, a final conference etc..

The participatory plan can be seen as a link between between theory and practice, between objectives, local constraints and the available methodology. With the participatory process the participatory manager is able to adjust available methods and resources to the given circumstances of a specific location.

4.2 Publication phase - Information provision

The publication phase falls, of course, within the realm of communication. Sufficient information provision requires a well-functioning communication process. Information provision is one-way communication where information about a new project is published. In most cases legally prescribed means of information provision such as an official announcement in the local newspaper or a weekly neighbourhood journal is not sufficient. A comprehensive website is nowadays one of the most suitable ways of providing information, and should be used as a basis for all supplementary information such as brochures and fliers. The website should not be a mere advertisement, but also provide detailed and up-to-date planning documents, maps, calendar of (planned) events and any available information that may be published according to the communication plan. For people without Internet access there must be a well-known location to view and browse the same documents. Depending on the available budget, regular advertisements in conjunction with reference to more detailed information in local media are preferable to flyers and colourful booklets.
Individuals and groups must have a chance to react. These reactions have to be documented. The documentation helps to recognise possible design errors at a very early stage of a planning procedure, identify 'difficult stakeholders' and (potential) conflicts.
In the recent years planners have started to provide the public not with a first and only draft of a

plan, but with a number of planning alternatives. This is a step forward towards more flexibility and adaptive water management, but it entails two contradictory problems: 1) a part of the public may be confused, and may have problems in comprehending these alternatives; and 2) the public may only choose between the provided alternatives, which introduces a new inflexibility. One solution might be a very explicit explanation of the options and the public should be repeatedly be asked for their level of awareness. Another approach could be to design just one map indicating the variable elements, and have a precise description of the variability of these elements in addition to the map.
Furthermore, the 'information provision' phase can be used to find yet unknown stakeholders and interest groups. As indicated the section 'communication', there should be a central contact address and person to whom interested new individuals and groups can refer. This is especially important for large projects that involve a consortium of planning organisations.
Additionally, individual and personal contact 'in the field' is always useful. On-side work can be efficiently coupled with publicity. Every project employee should be able to communicate with local stakeholders and the public, and be able to point out at least relevant contacts on specific issues. Besides identifying new stakeholders, this will build up trust and consolidate contacts.

4.3 Dialogue phase

Surveys and interviews. Planners have to make a decision how knowledge elicitation can be accomplished in the best way. If most stakeholders as well as current public opinion on relevant issues are known, the next step can be a meeting. However, it may be preferable to interrogate the public and/or stakeholders to find out about their perspectives on the relevant issues, before bringing them together. The results may be used in preparing a meeting and other subsequent participatory methods.

The first public meeting. After all available electronic and print media have been employed to inform people, a public or stakeholder meeting is recommended as a next step for several reasons:

- awareness raising;
- relevant information can be provided to the public or stakeholders in more depth;
- first questions can be answered, misunderstandings can be cleared up;
- yet, unknown stakeholders can be identified;
- stakeholder perspectives and opinions can be identified and categorised;
- side effects, and previously neglected problems or anticipated problems may be identified;
- opinions on the effectiveness of information provision can be collected (monitoring).

However, project managers and planners have to decide in which phase of the participatory process a meeting is advisable. Significant indicators may be:

- Has all collected information been pre-processed and available?
- Is the staff who performs and supervises the meeting appropriately prepared?
- Are all relevant parties able (and willing) to participate in the meeting?
- Ensuring that all individuals and groups will not be excluded from subsequent participatory activities?
- Can eventual mutations such as the transition from public to stakeholder participation be anticipated and handled?

A public meeting must be well prepared, and methods especially for knowledge elicitation, must be thoughtfully selected. When applying knowledge elicitation it is crucial to distinguish between asking stakeholders for their knowledge, preferences, and interests about the (desired) state of their environment and preferred measures to be implemented.
Small groups can be more easily handled than larger groups, and a variety of methods for this management is available. For larger groups a 'large group response exercise' is advisable.
The performance of a (first) meeting can be trend-setting for the course of the subsequent participatory process. For this reason the support of a professional moderator is recommended, if the planning organisation or consortium itself does not have a sufficiently trained or neutral moderator. They should be prepared to collect information about the quality of the meeting and the participatory process thus far (monitoring). The meeting itself must have clear goals and provide every involved person with tangible results. One result must be a programme for the participatory process (a participatory plan).

Analysis of public meetings and related activities. A thorough and well-structured analysis of the meeting and the preceding participatory process is a determining factor for the further participatory process from now on. At this point all relevant stakeholders should be known. All problems, new criteria and new findings must be identified, analysed and categorised. Further steps depend on the results of this analysis. In other words, this analysis helps the water manager decide how to proceed. The combination of the goals of the planning procedure and the participatory process provides the manager with a number of options. Here are some examples:

- Public or stakeholder participation: At this point it should be clear if a participatory process with stakeholders, the public or a combination of both is the dominating process;
- Stakeholder categorisation: After the identification of stakeholders, before or after the initial provision of information a categorisation of stakeholders is highly useful. This provides an overview of relevant stakeholders, and distinguishes them according to their perspectives and interests. Particularly, when a large number of groups and individuals manifest their interest in the participatory process, it is important to have an overview of stakeholders and their representatives;
- Is the involvement of volunteers useful and helpful? If so, this should go into the planning of further participatory activities including course work, workshops, creative activities, etc.;
- Consensus or controversy: this is one of the most sensitive issues in a participatory process. If conflicts are detected, apply 1) social learning methods such as role playing games, scenario building exercises, citizen juries, nominal group technique, group model building or similar methods, and 2) response methods: views and perspectives of particular stakeholders may be tested with models and simulations in order to display the consequences of individual management options on the project;
- Has the meeting revealed (planning) design errors such as missed criteria, side effects, hidden costs, etc.? If so, the original plan or draft should be modified, and introduced to the public and stakeholders in a new meeting or information campaign;
- If stakeholders have a significant knowledge deficit, educational activities or awareness raising activities are recommended.

At this point a definite participatory plan including a planning sheet should be developed.

4.4 Response phase

The design of the response phase within the participatory process depends on the analysis of a first meeting and related activities. The following methods may be applied:

- *educational activities*, if there are still knowledge gaps among stakeholders, but also professionals;
- *events* can help to increase the public awareness and popularity of the project;
- *popular involvement campaigns* can recruit volunteers, and take advantage of the labour and creativity of lay people. Besides idealistic, artistic and monetary benefit, these activities may increase the sense of ownership significantly;
- for a *follow-up meeting*, collected, analysed and processed records and other information must be available, so that this activity does not repeat the effort of a previous meeting.

Stakeholder analysis and categorisation including the various views of the stakeholders should be completed. In this particular meeting stakeholders may already be confronted with scenarios or planning options that include various perspectives, and indicate the consequences of particular options on the individual stakeholder or groups as well as the variety of possible consequences on the affected community. Appropriate methods are (environmental or economic) models, maps with planning alternatives, 3D models, story lines and other. Stakeholders should able to comprehend the methods, and draw conclusions.

Working with models and scenarios. The use of models and scenarios can be part of a follow-up meeting. Whereas stakeholders and lay people were asked to provide their domain knowledge to planners, experts have to review their own plans and models. In a way this step within both the planning process as well as the participatory process can be seen as a validation and verification of the plans and the available models respectively. If stakeholder perspectives have proven that the plans are not viable, the plans should be modified and employed in a new validation round with stakeholders. All perspectives of stakeholders, however, should be processed with the available methods such as models, simulations, thought experiments or scenarios, and displayed as well as discussed with the stakeholders. Participants must be chosen in a way that those individuals are able to comprehend the methods and all groups are sufficiently represented.
In particular, interactive model approaches require a great deal of expertise and moderation skills. In many cases it is advisable to employ an external expert.

Final event. At the very end of the implementation phase of a project it is advisable to run an event that brings together all stakeholders and managers. This may be a final conference, a street party, field trip, etc. The purpose is to create a sense of ownership and community. The latter is important for acceptance, maintenance and sustainability of the overall project.

5 LESSONS LEARNED FROM THE TRUST PROJECT

As mentioned in section four, even after providing the project partners with a document containing guidelines, tips, a systematic overview of participatory methods and the catalogue of methods [Hare and Krywkow, 2005], the water managers still had problems to apply these methods to their local projects. It was simply too abstract for planners, engineers, communication officers and ecologists. At the same time the listing of the various professions indicates that in water management projects, independently from the planning context, most professions are well cast except the position of a 'participatory manager'. At the start of the project the impact of a participatory process was underestimated or misunderstood. One project manager viewed the participatory process as risk management in the beginning of TRUST, but completely changed at the end. The co-development of participatory plans with consultants was a step forward from the point of view of water managers. However, the consultants' support was still required. Afterwards the managers still asked for a 'cookbook' of participation, which resulted in the procedural approach indicating what methods and tasks are most useful at which phase of the project. A cookbook is not possible and would be too inflexible and mechanistic. However, the integration of the guidelines [Hare and Krywkow, 2005] with the procedural framework [Krywkow, 2007] was better comprehensible for most of the partners. Probably, the most significant insight of the TRUST project was the fact that

the same participatory management scheme was applied to five water management projects coping with five completely different project goals (canal restauration; enlargement of a recreation area; building a pumping station in combination with canal broadening; building a freshwater reservoir and building water purification ponds in a park) spread over for different European countries. In other words the impact of local and regional constraints is not the prevailing factor for the design of a participatory process. This fact was underpinned by the effective and successful process of capacity building, the comparability and multi-national exchange of experience in the field of participatory processes. A comprehensive documentation of the TRUST experience is provided in Krywkow et al. [2007].

6 SUMMARY AND CONCLUSION

This paper is an attempt to provide practitioners in the field of water management with tangible advise how to organise a participatory process within a water management project. There is much emphasis on the implementation of methodological knowledge (process) in the specific physical and social circumstances of a local or regional project (constraints). Much effort in research and consultancy work was required to generate this framework for 'participatory process management'. However, it has been the collaboration between water managers, consultants and scientists in the TRUST project that has contributed to this effort, and revealed a number of gaps between available participatory process guide books and the day-to-day work of water managers.
The guidelines such as described here can be seen as sort of 'optimisation' of a participatory process, and is fully consistent with section 14 of the Water Framework Directive [EU, 2000, L327/16]. The novelty of the participatory management framework is the process oriented guidelines with the main criteria: (1) the systematic view of projects including the criteria *objectives*, *constraints* and *process*, (2) the close interrelationship between project planning and the participatory process, (3) monitoring and evaluation of the process with the possibility of stakeholder feedback, and (4) a generic process scheme that is applicable to many cases independent of their constraints.
The literature examples in the introduction of this article as well as the experience of the TRUST project indicate that providing useful guidelines for managers, planners and scientists who are not experts in participatory process management remains a challenge. The virtue of this participatory management scheme is that individuals and organisations who have no experience in participatory processes are enabled to obtain an overview of available methods within a systematic framework, and quickly learn about the required steps and tasks to undertake within a procedural framework. Furthermore, together with evaluation and monitoring efforts managers are encouraged to integrate the participatory process in the other planning and implementing activities. This holds also for scientists who apply models in a human-environment context. The participatory management scheme may help those modellers to better embed their models in a participatory process. The expected benefit is an improved communication and social learning process between modellers and stakeholders, a better understanding of the model as well as an appropriate model design in regard of stakeholder requirements. This article can only give a brief overview of guidelines for water managers. A more comprehensive document will follow this one.
In summary it can be said that an appropriate participatory management process enables managers as well as modellers to improve the results of a planning process and its maintenance, increase the acceptance of results and prevent or at least minimise unwanted side effects. This all assumes an adaptive and transparent governance style that truly accepts the participation of lay people, and allows engineers or planners to adjust their calculations by incorporating domain knowledge.

ACKNOWLEDGEMENT

This study was possible through the European InterReg III project TRUST. Very special thanks to Karina Speil. We would like to thank all participants of TRUST, theme group III, and Caroline van Bers for careful editing and thoughtful suggestions.

REFERENCES

Arnstein, S. R. A ladder of citizen participation. *Journal of the American Institute of Planners (JAIP)*, 35(4):216–224, July 1969.

Bousset, J.-P., C. Macombe, and M. Taverne. Participatory methods, guidelines and good practice guidance to be applied throughout the project to enhance problem definition, co-learning, synthesis and dissemination. Technical Report SEAMLESS Report No.10, Cemagref, 2005. SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org.

Elliott, J., S. Heesterbeek, C. J. Lukensmeyer, and N. Slocum. *Participatory Methods Toolkit. A practitioner's manual*. King Baudouin Foundation and the Flemish Institute for Science and Technology Assessment (viWTA), 2005.

EU. Directive 2000/60/ec of the european parliament and of the council. *Official Journal of the European Communities*, OJ L 327:72, December 2000. WFD.

Hare, M. P. and J. Krywkow. Participatory processes for the design of water storage areas. Technical report, Seecon Deutschland GmbH, Osnabrück, 2005. Seecon TR no. 01/2005 for the TRUST project.

Krywkow, J. Participatory management - a brief guideline for water managers. Technical Report Seecon Report #7 2007, Seecon Deutschland, Osnabrück, Germany, 2007.

Krywkow, J., K. Rasche, E. Moss, T. Mitchell, K. Vancleemput, R. van der Kroef, R. Hotting, A. Rodenbach, and C. Noordanus. Public and stakeholder participation. Theme Group III Final Report of the TRUST Project, Rotterdam, October 2007.

Mostert, E. The challenge of public participation. *Water Policy*, 5:179–197, 2003.

Rasche, K., J. Krywkow, J. Newig, and M. Hare. Measuring the intensity of participation along six dimensions. Edinburgh, June 2006. The Macaulay Institute. Proceedings of the PATH conference, http://www.macaulay.ac.uk/PATHconference/outputs/PATH_abstract_3.3.3.pdf.

Ridder, D., A. Mostert, and H. Wolters. *Learning together to manage together. Improving participation in water management*. University of Osnabrück, Institute of Environmental Systems Research., Osnabrück, 2005.

Wates, N. *The community planning handbook*. Earthscan Publications Ltd, London, 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

The Event Bush as a Potential Complex Methodology of Conceptual Modelling in the Geosciences

Pshenichny C. A.[a], Nikolenko S. I.[b], Carniel R.[c], Sobissevitch A.L.[d], Vaganov P. A.[e], Khrabrykh Z. V.[f], Moukhachov, V. P.[g], Shterkhun V. L.[h], Rezyapkin A. A.[i], Yakovlev, A.V.[j], Fedukov, R.A.[k], and Gusev, E.A.[l]

(a) Geological Mapping Division, VNII Okeangeologia, Angliisky Prospect, 1, St. Petersburg 190121, Russia – Faculty of Geology, St. Petersburg State University, Universitetskaya Naberezhnaya, 7/9, St. Petersburg 199034, Russia (pshenich@kp1306.spb.edu, cpshenichny@yandex.ru)

(b) St. Petersburg Department of the Steklov Mathematical Institute, St. Petersburg, Russia (sergey@logic.pdmi.ras.ru, snikolenko@gmail.com)
(c) Dip. Georisorse e Territorio, Università di Udine, Udine, Italy (roberto.carniel@uniud.it; carniel65@gmail.com)
(d) Laboratory of applied geophysics and volcanology, Schmidt Institute of Physics of the Earth, Russian Academy of Sciences; Moscow, Russia (alex@ifz.ru)
(e) Ecological Geology Dept., Faculty of Geology, St. Petersburg State University, St. Petersburg, Russia (pvaganov@hotmail.com)
(f) Laboratory "Magma and Volcanoes", University of Blaise Pascal, Clermont-Ferrand, France (zkh@mail.ru)
(g) Logic Dept., Faculty of Philosophy, St. Petersburg State University, St. Petersburg, Russia (mukhachev@mail.ru)
(h) Karpinsky All-Russia Geological Research Institute (VSEGEI), St. Petersburg, Russia (Victoriya_Akimova@vsegei.ru; v-shterkhun@yandex.ru)
(i) Karpinsky All-Russia Geological Research Institute (VSEGEI), St. Petersburg, Russia (alex_rez@pisem.net)
(j) St. Petersburg State University of Information technologies, Mechanics, and Optics, St. Petersburg, Russia (a.yakovlev@creatstudio.com)
(k) St.Petersburg Research Centre for Ecologic Safety, Russian Academy of Sciences, St. Petersburg, Russia (fedrom26@mail.ru)
(l) Geological Mapping Division, VNII Okeangeologia, St. Petersburg, Russia (gus-evgeny@yandex.ru)

Abstract: The event bush is a new method of artificial intelligence proposed specially to meet the needs of geosciences, from basic research to communication with non-professionals. First results of its application as well as combination of event bush with mathematical and logical formalisms are encouraging, and the method has shown some advantages to existing approaches, as well as an ability to unite them and become a complex integrated methodology. However, since the very first steps, it turned out that the way of thinking that underlies event bush differs from virtually all existing pathways of thought perceived in the Earth sciences. This paper aims to provide the practical guidelines for building event bushes. For this, it describes the method through an example, chosen to be clear to any geoscientist and even non-scientist. The procedure of inference in the event bush is considered in positional (graphic) and semantic aspects. Interrelations of event bush with existing formalisms and ways of thinking are discussed. Recommendations are formulated to make the best use of event bushes in practical reasoning.

Keywords: Event bush; Conceptual; Artificial intelligence; Geoscience; Knowledge; Reasoning.

1. INTRODUCTION

A new method of artificial intelligence, the event bush, was proposed specially to meet the needs of geosciences, from basic research to communication with non-professionals (Pshenichny and Khrabrykh, 2002). It aims to act as an intermediary between hardly formalized, loose, and largely intuitive knowledge, from one side, and strict methods of mathematics and logic providing quantitative computation and logical inference, from the other. It also intends to bring our knowledge into an intuitively clear framework to make it better understood by students and non-professionals.

First results of its application to various geoscientific tasks (Pshenichny and Fedukov, 2007; Behncke and Pshenichny, submitted) as well as combination of event bush with mathematical and logical formalisms (Nikolenko and Pshenichny, 2007; Pshenichny et al., 2008) are encouraging. However, since the very first steps, it turned out that the way of thinking that underlies event bush differs from virtually all existing pathways of thought perceived in the Earth sciences, be it essentially inductive and non-strict traditional geological consideration (for example, reconstruction of geological history), modelling in terms of physical parameters, building of a single-root event/probability tree (Newhall and Hoblitt, 2002), or compilation of a Bayesian belief network (BBN; Jensen, 2001; Aspinall et al., 2003) based on the expert's knowledge and intuition.

The purpose of this paper is to provide some guidelines to allow one to build an event bush. For this, the principles and application of the method are viewed through an example, which is deliberately taken far from the geosciences. We hope that this will make the procedure clear even to a non-scientist and prove the wide scope of the presented method.

The event bush itself is a firmly defined formalism based on the graph theory (e.g., Tutte, 1998). It is linked to some other formalisms like BBN and allows highly formal management of terms. Still, formalization of geoscientific and alike contents is a complex and still unresolved problem. Therefore, the construction of an efficient event bush for a particular domain of knowledge until now is largely a matter of intuition and may need repetitive tries, redefinition of entities, shifting their positions and interrelations to reach the "best fit" to the modeled domain. Therefore, the example of application is, in fact, an experience of our reasoning at solving a problem with an event bush, not necessarily straightforward from the very beginning and still not sufficiently formalized. Some of the results described below should be considered working guesses awaiting their strict formulation. Nevertheless, we believe that the reasoning we follow is efficient, and would like to share our experience.

2. PRINCIPLES OF CONSTRUCTION OF THE EVENT BUSH

2.1 Getting Started

Informally speaking, the method of event bush can be compared to creating a script for a cartoon, or, yet more precisely, for a computer game. One specifies characters and background, makes assumptions stating what these characters principally may and may not do, and how the background, or environment, may affect their action. Based on this, the imagination of scriptwriter creates a family, or a cluster, of alternative scenarios plausible from the point of view of the assumptions made. These scenarios have definite beginning and end, and are put in uniform and uniformly understood, though not necessarily well-defined, terms ("gun", "monster", "labyrinth", etc.). We believe that in such a way one may adequately describe various domains of geoscientific knowledge of virtually any contents (e.g., action of pollutants in subsurface, eruptive activity of a volcano, seismic hazard, urbanization in a river valley, ore formation and extraction at the seafloor).

However, the example we are going to consider here is equally far from all the abovementioned. Suppose that you are a father of a boy who asks you to let him go for a walk in a wood near your home for a couple of hours in the afternoon. You know that a fog may occur. What would your reasoning be to make a decision?

1. First of all, let us define the area of reality. To make a decision, one should reason on "what can happen if a boy goes for a walk". This is a fair entitlement for the said area and, simultaneously, for the corresponding event bush. Note that it distinctly splits into two parts, "what can happen" and "boy goes for a walk". (Though, formal ground for this division still needs to be defined.) Let us plot the former on the top, and the latter, on the left of the future graph (Fig. 1). General flow of scenarios will be from left to right; modifications will come from top. It is solely up to us what to put where. If our concern is *what can happen* to a boy, then we should put this phrase on the left; if we want to know what can happen *to a boy*, then "boy goes for a walk" goes to the left, and the other phrase, to the top. Right column and central field remain blank so far.

2. Below each phrase, we will add the appropriate details that we know and/or want to consider. Thus, at the left, we put the boy who goes for a walk and the father who stays at home (Fig. 2). We may add other family members, neighbors, friends, accidental passers-by and whomever and whatever else we consider relevant (e.g., hungry wolf, dead tree longing to fall down onto somebody's head, and so forth). These will be the "type one entities" (Pshenichny et al., 2008) denoting the primary, not overlapping and non-unique inputs (basic objects, processes or tendencies). These inputs would predetermine any further course of events ("happenings") in the modeled area. Either we use a classification to define them, or simply take any entities we like, just making sure that we always mean different things by different words.



Fig. 1. Setting the task to compose an event bush

At the top, just to start with, we put the circumstances of the supposed walk that we know: in a wood, for a couple of hours, in the afternoon, with possible fog. These circumstances, after some revision, will become environmental entities of the event bush. They will indicate the way the environment may independently affect the primary inputs or influence their further, indirect manifestations, thus "shaping up" different "happenings" – what if the boy walks in a wood, what if he gets in a fog, and so forth.

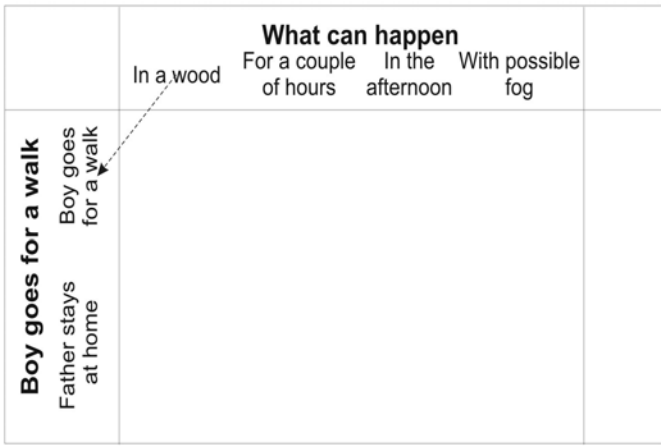


Fig. 2. Adding entities (i) and (ii) to the bush draft

3. Now we should analyze the primary and environmental entities we selected and formulate simple statements describing them. Simple statements are those that have one subject and one predicate (e.g., "Fog occurs"), and hence, no logical connectives like AND, OR, etc. With these statements, we will begin to construct the bush. Editing them still relies on intuition and is determined by the context. Environmental entities are the circumstances that may happen or not in the environment. As we are interested only in the walk in a wood, the entity "in a wood" is unlikely to be

regarded as such. At present, we add it as a property to the boy's walk (the first-type entity). If nothing peculiar comes out of it, we may reject it all.
The circumstances “for a couple of hours” and “in the afternoon” seem to make sense together: the matter is obviously not a fixed duration of walk but coming back before dark. Then, these two circumstances can be fairly well expressed by one statement, “It is getting dark” (with the truth values “yes”/”no”).
Similarly, the possibility of fog can be expressed by a simple statement “Fog occurs” (yes/no).
The first-type entities are also transformed into simple statements with assumed truth value “yes”. The fact that father is at home can be expressed directly, “Father stays/is/remains/etc. at home”. Concerning the boy, we have to pick up the circumstance expelled from the top, “in a wood” (Fig. 2). Also, we may formulate the same idea in a number of ways, “Boy goes for a walk in a wood”, “Boy is walking in a wood”, “Boy follows his way in a wood”, etc. We may make final choice later. The statements (i) of the first-type entities may change their formulation (say, become compound) but unlikely be added or removed (that means, we will go ahead with the same two), while the environmental statements (ii), vice versa, must remain simple and may not change their formulation once added to the bush, but may be added anew, as well as moved from the top or removed from the bush at all.

2.2 Developing the Bush

Now we plot the statements, which we have had, in the form of nodes (Fig. 3) and begin the procedure of inference in the bush. The inference is assumed to be represented by an “IF... THEN...” relation and is graphically expressed by an arrow (directed arc). Making the inferences from statements (i) with or without participation of (ii), we obtain various “happenings”, which we term as secondary entities (being iii, after the primary, i, and environmental, ii), and tertiary end results (iv). Their location is shown in Fig. 3.

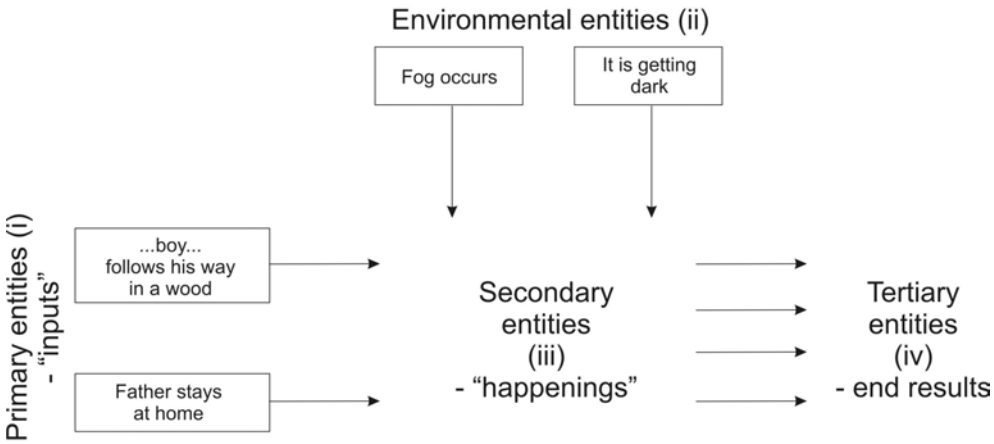


Fig. 3. Architecture of the event bush

In general,

1. statements (i) and (ii) must not lead to other (i) or (ii);
2. statement (ii) may lead to a statement only together with (i) or (iii);
3. statement (iii) must not lead to (i) or (ii);
4. statement (iv) must not lead to (i), (ii), (iii), and another (iv) (Pshenichny et al., 2008).

Inference in the event bush is determined graphically and semantically.
Graphically, the simplest inference is that of one statement from another, e.g., “IF Father stays at home THEN Father stays at home”. Another example could be, IF “... boy gets wet” THEN “Boy... becomes sick”. The nature of omission dots will be explained below. This kind of inference implies four positional options itself,

IF (i) THEN (iii),
IF (iii) THEN (iii),
IF (iii) THEN (iv),
IF (i) THEN (iv).

Also, two statements may lead to one. This operation is called *combination* of premises (Pshenichny et al., 2008) and attributed the connective AND. An example of combination is IF ("... boy... follows his way in a wood..." AND "Fog occurs") THEN "... boy... follows his way in a wood in a fog" (Fig. 4). Only pair combination is adopted in the event bush. Aiming to combine three or more statements, one should ensure that a combination of at least two of them makes sense in the considered domain, and its result makes sense in combination with at least one of the remaining statements, and so forth, otherwise combination is not permitted (see example in Pshenichny et al., 2008). Moreover, in each combination, we distinguish the main and the modifying premises – e.g., "Boy follows his way in a wood" (or whatever formulation we choose) is the main member, and "Fog occurs", the modifying one. On a graph, combination is marked by a rounded crossing ("right turn") from the main premise to the conclusion. As is seen from the above, there is notable difference between combination and logical conjunction.

Besides, one statement may lead to a number of statements. This is the operation of *division* of corollaries, which is marked by a circle in Fig. 5. The circle implies that the resulting statement is compound and includes alternatives listed via OR (IF "... boy... follows his way in a wood in a fog" THEN ("... boy... looses his way in a wood in a fog" OR "... boy... walks slower in a wood in a fog" OR "... boy... admires the fog following his way in a wood"). Each alternative represents a separate node of the bush. Like in the probability theory (and unlike the disjunction in logic), the statements resulting from division must comply with additivity condition, i.e., be mutually incompatible and exhaust the whole variety of possible outcomes seen by an expert.

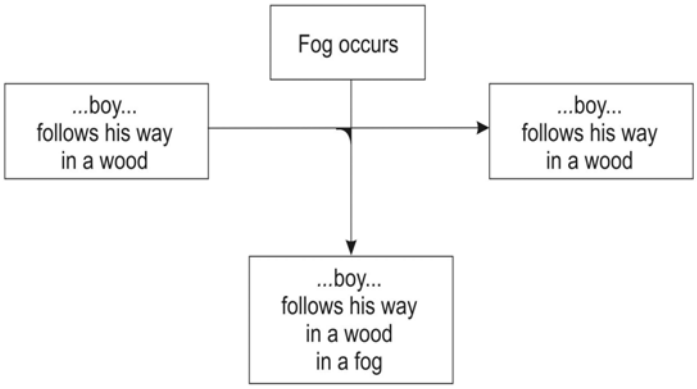


Fig. 4. Operation of combination in the event bush

Wishing to express a case that the options are compatible, one may put several arcs going out of one node (in Fig. 5, another arc leads from "... boy... follows his way in a wood in a fog" to "... boy... gets wet following his way in a wood"). This means that a statement can lead to any number of others, in combination with another statements or itself. However, drawing each next arc from a node, one should make sure that the newly introduced statement denotes a consequence really compatible with those denoted by the statements already derived from this node. Otherwise a new arc must begin not from the node but from the division mark (circle).

Two operations one after the other are not permitted in the event bush; a node must be in between. In other words, each result of each operation must be named. Thus the graphic issue comes to intersection with the semantic one.

Semantically, the event bush does not require strict definition of terms, nor does it insist on exhausting enumeration of all possible entities (i) and (ii). Still, to make maximum use of the event bush, one should very accurately manage each word introduced in any node. Strict and formal semantic rules are expected to be developed for the event bush method in the language of predicate logic. However, semi-intuitive guidelines, which notably increase the representative power of the event bush, can be introduced right now.

Regarding any statement (node) that is already present in the bush, we may specify, at least informally, what corollaries it may, or must, have. For instance, let us put that statements "Father/son... gets into a fog" must have two independent groups of corollaries – one, consisting at least of two statements describing whether the person is lost (these corollaries are listed via OR), and the other, consisting of one statement telling quite a different thing – that the person has got wet. Also, we may agree on that the statement "It is getting dark" may be combined in our bush (and hence, have corollaries) only with those statements describing the boy, which are, in turn, the corollaries of "(Something else) AND Fog occurs". This is reasonable if we consider a case that the fog, if happens, does in the

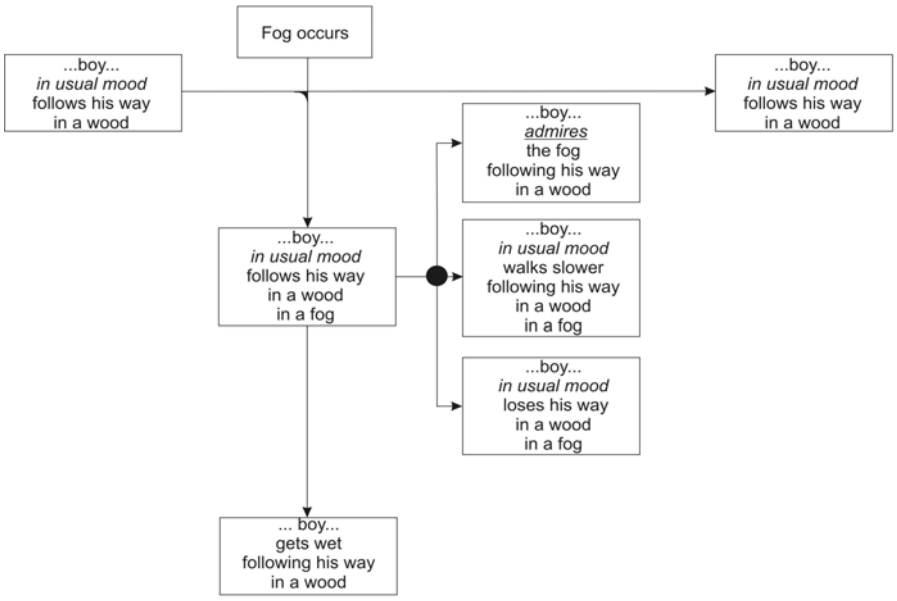


Fig. 5. Operation of division and multiple compatible corollaries in the event bush

afternoon, well before it begins to darken, and the boy, whose intention was to go for a walk to return before dark, is disciplined enough not to be late without a reason. Meanwhile, statements describing the father are not subject to this rule, because father may leave home and seek for the boy precisely because it has already become dark – certainly, if we believe that fog is the only circumstance that may impede the boy's return before dark.

It is noteworthy that we have made quite a few of additional remarks, or assumptions, at a pretty small step. These are,

- fog happens in the afternoon but well before dark,
- boy wanted to be back before dark,
- boy is disciplined and honest,
- fog is the only circumstance that may impede the boy's return before dark.

Such additional preconditions may arise at every step of construction of an event bush. Naturally, once arose, they remain at the following steps if are not refuted by new premises. We recommend to put these preconditions down in explicit form and store them with the bush. One should keep track of them at least not to loose them making new steps of inference and to avoid a contradiction between such assumptions introduced at different steps. Moreover, such assumptions can retrospectively refine the definition of the field of reality/knowledge the bush describes.

We may define the wording of the corollaries given the premise. For instance, we postulate that any corollary of statement “Fog occurs” must include “...gets into a fog”, and if we obtain a statement that includes “to be lost” (or “loose one's way”, or the like), at least one of its corollaries must include the property “get depressed spirits”. Moreover, rules can be suggested for a statement in a bush depending on whether it leads to a corollary alone or in combination, whether it is the main or modifying member in the combination, what the character of some preceding or neighboring statements is, and so forth. Importantly, the wording can be predetermined not only for direct corollaries but also for some farther consequences, up to the very end of a succession of inferred statements. Along with stored assumptions, this is often very helpful in retaining all the necessary information throughout the bush. Thus, concerning the property “in a wood” that we moved from top to left, we decide to keep it in all statements that describe only the behaviour of the boy, from primary (first-type) to the end results (tertiary statements) – to allow the case that this circumstance appears relevant at some step of inference. However, even if not, we may wish to have it explicitly pronounced throughout the bush. The same time, keeping such properties with the nodes means that simple statements describing the entities (i, iii, and iv) and associated with the corresponding nodes accumulate new properties and therefore become compound: “...boy is walking/follows his way/etc. in a wood” actually consists of the following

statements, "...boy is walking/follows his way/etc." AND "Walk takes place in the wood"/"the way lies in the wood"/etc.

Similarly, once we infer that "...father/boy... becomes sick", as we do not admit any immediate remedy in the wood, the said person would remain sick till the end of the story, and the sickness itself remains relevant, for it may lead to corollaries like "get depressed spirits" at any time. Therefore, all the concomitant consequences from the statement "...father/boy... becomes sick" till the end results of the bush will include the property "sick": "Sick father/boy ... (does something, feels something, becomes something, etc.)" can be rewritten as "Father/boy is sick" AND "Father/boy ... (does something, feels something, becomes something, etc.)". The connective AND here seems to be quite compliant with logical conjunction, in contrast to the AND we use in the operation of combination. However, "dragging" the words throughout the bush is not mandatory. On the contrary, the property "gets wet" can be mentioned only once in our bush relating each of the primary entities of the bush (I.e., boy and father), because, despite that the wet state may last long in a wood, it is relevant actually once (to our mind) – namely, in making a person sick. "Once" is not necessarily "instantly", we do not consider quantitative measures like duration here; this is a subject of other investigations (Nikolenko and Pshenichny, 2007; Yakovlev and Nikolenko, 2008). Speaking qualitatively, "once" means at one step of inference, whatever time it takes, seconds or ages.

An important feature of the event bush semantics is backward redefinition – an opportunity to change the formulation of statements introduced earlier (and thus to redefine the nodes) depending on new corollaries. For instance, as discussed above, a statement "... gets wet" implies "... becomes sick". Once someone became sick, our common sense implies that he/she was not sick before. Moreover, saying that he/she became sick under some condition and only thus, we mean that the property "to be sick" would not be attributed to those cases which have not met this condition, unless stated otherwise. Then, if there is no property, there must be its negation. Hence, we may automatically ascribe the negation of this property to all nodes of the bush that describe the same primary entities (for example, "son" and "father" but not "hungry wolf") and have no word "sick" in their formulation.

Furthermore, negation of a property is a property itself. For the sake of simplicity, we put "healthy" as the negation of "sick", though, generally speaking, one could define the physical states of a person as "healthy, wounded, sick, sick and wounded", and formulate the negation of "sick" per se as "healthy OR wounded OR sick and wounded", and multiply the nodes of an event bush, correspondingly. Phrases added by backward redefinition of nodes are given in italics in Figs. 5 and 6. In the example in Fig. 5, in addition, the added property ("admires") that made us redefine other nodes is italicized and underlined. The backward redefinition is the reason for numerous omission marks in the statements quoted above. Indeed, until the bush is completed, we cannot give a full formulation of any statement except an environmental one (put on the top) – but concerning the latter we are not sure it will "survive" at all. This implies that, with inferring new corollaries, the premises become more and more compound a posteriori.

Another semantically peculiar case is the «one from one» inference. This can be of two types, «IF A THEN A», e.g., "IF Father stays at home THEN Father stays at home", or «IF A THEN B», e.g., IF "... boy gets wet" THEN "Boy... becomes sick". The former case is trivial yet important. It reflects the situation when nothing happens. The latter one represents the option of "self-generation" of a consequence: it looks as "getting wet" produces "becoming sick" without any other premises. However, this can be the case if, and only if, we have assumed as a general rule for the whole bush that IF "... gets wet" THEN "... becomes sick". This is another semantic assumption that determines the wording of the corollary and, possibly, ultimate consequences. Moreover, like any other semantic assumption, it may be made not only for one, but also for some or even all entities "eligible" to take the place of omission dots in corresponding statements. Indeed, for the father, if he gets wet (in a fog or somehow else) the effect will be the same as for the boy – he will become sick. As was stated above, structurally, there are four opportunities for such inference, IF (i) THEN (iii), IF (iii) THEN (iii), IF (iii) THEN (iv), and IF (i) THEN (iv). Concerning the latter, it is very convenient to begin the inference in an event bush by drawing horizontal arcs across the entire plot from each primary statement to the right column, thus setting up a "skeleton" of the future graph (though these arcs may then disappear in the process of construction). In our experience, the inference «IF A THEN A» makes sense for a domain of knowledge only in this case, IF (i) THEN (iv). First, this provides a good semantic guidance for the process of event bush construction. Also, if we

come up with the same as that what we started with, this implies a beautiful, though still unexplored, opportunity of building semi-cyclic bushes. Although, to keep on the point better accepted by our common sense, we may slightly modify the formulation of the tertiary statement to make it merely grammatically different from its cause, e.g., IF "Father stays at home" THEN "Father at home". However, in the present case we would like not to do this and will formulate the primary statements, as well as and their corollaries, as follows.

Primary statement One: taking into the account that the first environmental statement "Fog occurs" is combined with it and implies at least one of the corollaries of this combination be "... loses his way", the primary statement must explicitly include the negation, that is "... follows his way". Then, at present, we put this statement as "...boy follows his way in the wood...". However, we do keep in mind that other properties will be added to it in the process of construction of the bush.

Primary statement Two is formulated yet simpler, "Father stays at home" - but with the same reservation.

Armed with these guidelines, we can go ahead composing the event bush to answer the question, what can happen if the father says "yes" and boy goes for a walk.

3. DISCUSSION

General answer to this question is given in Fig. 6. The graph represents only the part of the event bush that describes what can happen to a boy. It does not include the part describing scenarios of the father's and joint father and boy's destiny. Statements in that part may look like

"Healthy father, who left home searching for healthy boy
being late following his way in a wood in high spirits,
gets into a dark",

or

"Sick boy in depressed spirits, being late following his way in a wood,
and sick father, who left home searching for sick boy
being late following his way in a wood in depressed spirits,
get into a dark".

Full version of the bush does not fit the publication format and will be reported elsewhere. However, the link between the two parts of the bush can be easily seen in Fig. 6. In the considered area of reality, this link is determined by paternal love and anxiety. In the domain of knowledge, it is expressed by the statement "Love exists" added as environmental (ii) to the bush. It produces, in combination with those statements telling that the boy is late or lost, an array of nodes ("Love urges to take care of...") that are, in turn, all combined with "Father stays at home" and thus lead to statements describing father's potential responses to the situation, e.g.,

IF ("Father stays at home" AND "Love urges to take care of healthy boy
being late following his way in a wood in high spirits")
THEN "Father leaves home in search for healthy boy
being late following his way in a wood in high spirits";

then,

IF ("Healthy father leaves home in search for healthy boy being late
following his way in a wood in high spirits" AND "It is getting dark")
THEN "Healthy father, who left home searching for healthy boy
being late following his way in a wood in high spirits, gets into a dark".

Naturally, other additions are also possible to the bush. For instance, we may suggest another node, "… boy gets wounded", and drop an arrow to it from virtually every node of the bush describing the boy, including the primary one, "Healthy boy in usual mood follows his way in a wood not going to be late". This is so because the very property "in a wood" may be assumed leading to "… gets wounded", for it is much easier to get a trauma walking in a wood than on a meadow or in a town. Similar inference (by "one from one" inference) could be made from "…gets in a fog" and "… gets into a dark", for both fog and dark notably favor wounding anywhere.

However, one may argue that not only fog and dark, but wood itself can also favor loosing the way. Hence, this may happen to a boy yet before the fog. To express this opportunity,

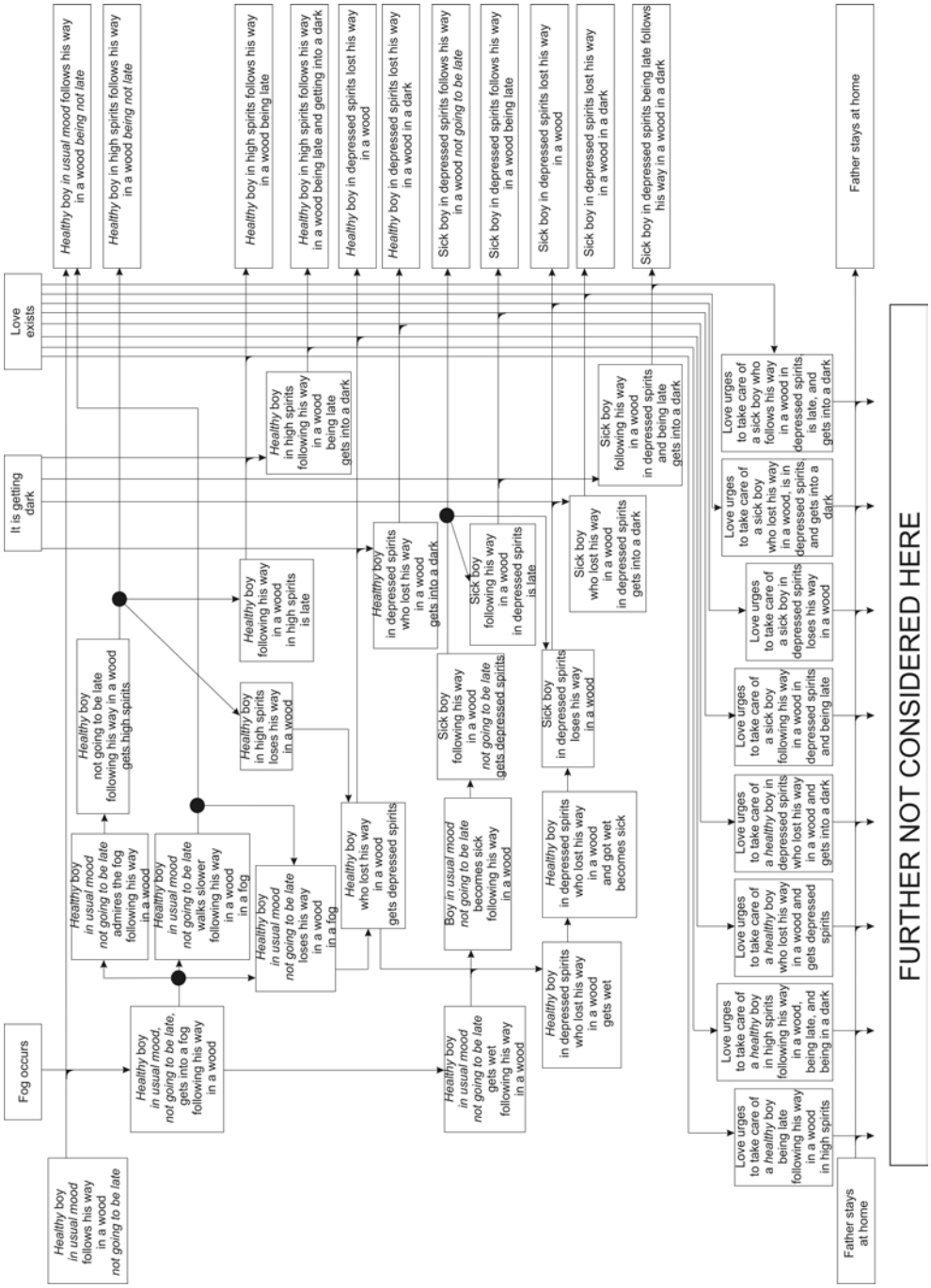


Fig. 6. A fragment of event bush describing what can happen if a boy goes for a walk in a wood for a couple of hours in the afternoon and a fog can occur.

we should rebuild our bush again. Then we would need to separate “walk” and “in a wood” and bring “in a wood” back to the top (Fig. 7). This example shows that “game is never over", and fixedness of terms and entities does not mean fixedness of concepts. Therefore, we can feel satisfied with one, however perfect, event bush, only for the time being, and then should look for updating and bettering it.

A natural question arises herewith; whether all these complications are necessary just to say “yes” or “no” to a boy who wants to go for a walk.

Certainly, not. For ages parents and other decision-makers have been saying “yes” or “no” without any decision-support tools. Moreover, perhaps in a case like the one described

above such tools are redundant indeed. But mind that billion-dollar decisions often answer very simple questions too and require just saying "yes" or "no". However, we see that even in this quite ordinary case the cost of decision could be boy's health and safety. Provided the father had a user-friendly decision-support software making recommendations based on an event bush, he may have used it – especially as the decision is actually more extended than just yes or no. If the father is wise enough, his most likely decision would consist of two parts, one, actually said (e.g., "yes, you may go"), and the other, unsaid ("but if fog occurs and he is not back, I will go seek for him, and take his dry clothes with me, and must do this before dark, because otherwise he may feel despair and become unpredictable"). To formulate this second, the most important, part, tools of decision support are highly welcome.

Though, one could argue that the decisions based on this event bush could be fairly well told by bare intuition and do not require such an intellectual effort. This might be true, but the weakness of intuition is its subjectivity, while event bush is a knowledge engineering tool that makes knowledge much more impartial. A domain of knowledge may start looking more simple or more complicated after using it, but in any case this tool provides a verification of intuitively made decisions.

Still, commonly the domains of knowledge in geoscience addressed by the event bush method appear far from simple, and the approach described in this paper allows one to handle this complicated and ambiguous knowledge and organize it to extract complete, well-substantiated, quite reasonably biased, and valuable conclusions.

"Traditional" domains of knowledge can be put in the event bush framework, as was demonstrated by Behncke and Pshenichny (submitted) by the example of description of volcanic eruptions of Etna, Sicily. The event bush suggested by these authors helped to better understand unusual eruptive activity that took place recently on this volcano, infer eruptive scenarios that were observed but had not been predicted, and foresee other possible scenarios to beware in the future. Consideration of physical models of geologic phenomena in the event bush framework will be discussed elsewhere (Pshenichny et al., in preparation). This is a perspective avenue for future research.

As for the relation of event bush to the most popular knowledge engineering and probability assessment tool in the geoscience, the event/probability tree, it can be concluded from Fig. 6 that a bush with several primary entities can be, with the loss of some valuable information, split into smaller bushes, each having only one such entity, and these, in turn, can be converted into event trees by ignoring all modifying members in combinations (and hence, omission of environmental statements), that means yet greater loss of relevant information. Naturally, a reverse operation ("assemblage" of event bush from event trees) is not possible.

Conversion of event and probability bushes into BBN and back was performed by Pshenichny et al. (2008) and Yakovlev and Nikolenko (2008). However, there existed problems in automatic conversion of statements of the bush into variables and states of BBN, which now seem to have been overcome even by those relatively loose semantic guidelines introduced in this paper.

Further perfection of these guidelines would lead, together with the structural rules introduced earlier, to thorough logical formalization of the event bush. This, in turn, must allow the construction of strict formal theories from event bushes and application of calculi of classical logic to prove or refute standpoints in given domain of knowledge like theorems in mathematics. This is seen as one of the main directions of ongoing research.

This research is fueled not only by academic interest, but also by an expected application to processing expert judgments. Any bush expresses, in fact, someone's opinion, or set of concepts someone has in mind in the given domain of knowledge. As has been shown in this section by the example of bush rebuilding (see Fig. 7), the same semantics can be used to generate a family of bushes expressing alternative opinions (or expert judgments) in similar terms. This opens a new opportunity for elicitation and reconciliation of expert judgments and computation of their logical probabilities seen as alternative to subjective probabilities used now (Pshenichny, 2004). Thus the conceptual modelling by means of event bush may lead to improvement of numerical assessment of expert judgments. However, to be able to quickly build and edit a number of bushes, a software is needed, which is being written now.

Other directions of development of event bush include quantitative applications of probability bush, particularly incorporation of geodata into the nodes and building the event bush-based geoinformation systems (GIS), development of temporal models and creation of

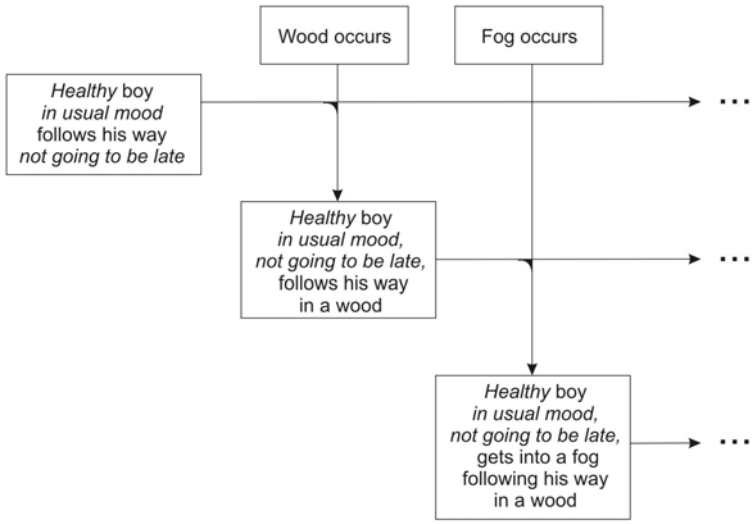


Fig. 7. Further transformation of the event bush to cope with new standpoints

learning algorithms for definition of temporal intervals associated with each node of the bush (Nikolenko et al., 2008).

Omitting the technicalities, the results of these studies show that the method of event bush is compatible with all the mentioned approaches and techniques. However, some of them lack congruity with one another. For instance, despite some interesting attempts (e.g., Gitis and Ermakov, 2004), formal theories per se are unlikely implementable in GIS. However, a formal theory can be recorded, with some theoretical development of the method, as an event bush. Based on the bush, a GIS can be built, where each node of the bush represents a layer or a theme. Thus a formal theory will become correctly exportable into a GIS. The same can be said about the combination of a formal theory and BBN. Another expected result is involvement of equations of multiple, competing or complementary, physical models into a BBN through the event bush. At present, to the authors' knowledge, these models are actually treated the same way as expert guesses in Bayesian computations in geoscience.

Generally speaking, nodes of an event bush can be attributed not only time and space co-ordinates but also synonymic terms, bibliography, comments, graphic and video information and so forth. This means that nodes may represent a kind of metadata, and the event bush can act as an ontology of the domain of knowledge/reality it describes. In contrast to most of ontologies, the event bush shows the ways of generation of specified classes of instances, which are precisely the steps of inference in the bush.

Many other promising interrelations of qualitative and quantitative approaches, and among the latter, the deterministic, probabilistic, fuzzy and, possibly, other ways of computation, are expected from the event bush.

The applications and extensions of this method discussed above must make it possible to express virtually any feature of an entity of the world addressed by our reasoning, be it a name of an entity, its explicit or implicit definition, consideration of what entities it resembles and what it looks like, what properties it has, what its sub/superclass and part—whole relations are, when and where it takes place, how it acts and what it is affected by, how this can be expressed in terms of quantitative (say, physical) models, what probability has its occurrence and what this probability is conditional on, and so forth. Moreover, suggesting a simple and intuitively clear language, the event bush is able to express different standpoints in similar terms and bring them in one framework. Moreover, the event bush suggests a unique approach to reconciliation of opinions – every step of its construction, from definition of the domain of knowledge/reality to making inference and attributing the pieces of additional information to the nodes, can be used to compare views and reveal discrepancies. Herewith, each possible disagreement can be appropriately placed. Scientists, decision-makers and other debating parties can query and figure out whether their misunderstanding sits in choice of primary/environmental entities, or in drawing a particular line of inference, or in formulation of the outcomes, or in considering

different time and space values, or in linking synonyms, etc. Not only the scientists but laymen able to reason in terms of event bush as demonstrated in Fig. 6 can pin down the contradictions and decide how they would like to cope with them. This makes the event bush highly amenable to decision-makers and gives us ground to consider it as a potential core methodology of organization and processing of information in the Earth sciences.
An important and intriguing issue is whether the event bush can be combined not only with BBNs but also with a large kit of existing knowledge engineering techniques – conceptual/existential graphs, semantic networks, Boolean networks, entity-relationship diagrams, causal loop diagrams, cognitive maps, qualitative probabilistic networks, influence diagrams, stock and flow diagrams, reasoning maps and others. Developed mainly by mathematicians and computer scientists with little or no account of the complexity of geoscientific information and wording, these techniques, however powerful, have found very limited application in the geosciences. Their interrelation with event bush needs to be explored, and we hope that a joint approach can be created incorporating the advantages of many methods.

4. CONCLUSIONS

Compared to other existing tools of knowledge engineering used in the geosciences (e.g., event trees, belief networks), the event bush provides a more detailed, structured, and accurate record of information. The principles of composition of event bush are simple enough to allow a geoscientist, who is not necessarily a specialist in knowledge engineering and artificial intelligence, to build and use event bushes for his own research needs as well as for everyday issues. Nevertheless, a bush is not as simple in composition as an event tree or BBN, and building and editing it requires some effort from a geoscientist. To make the best use of event bush, one should follow not only its graphic rules of composition and inference, but also the semantic guidelines that determine the formulation of statements. As a result, the language of event bush may differ from the language accepted in a modeled domain of knowledge and look somewhat clumsy and artificial; nevertheless, it provides much higher degree of objectivity and formalization.
The event bush has proved its principal compatibility with a wide range of methods and approaches used in the Earth sciences to organize information and compute values. This gives us ground to hope that this method may evolve into a complex methodology of treatment of geoscientific information.

ACKNOWLEDGMENTS

The research has been funded by Project V4 of Istituto Nazionale di Geofisica e Vulcanologia, Italy, “Development of quantitative tools to evaluate volcanic hazard”, a Human Capital Foundation grant, Information Technologies cluster project, 2006, and NATO-Russia collaborative linkage grant 982484.
Authors are deeply obliged to Jonah Utkin, Joan Marti, George Bardossy, and many other colleagues who have given very useful comments on this research, enlightened some weaknesses and fallacies of the approach and, hopefully, helped to cope with them.
Anonymous reviewers treated the manuscript rigorously but friendly and deserve special gratitude for making us rethink the practical issues of application of the event bush method.

REFERENCES

Aspinall, W.P., Woo, G., Voight, B., and Baxter, P.J., Evidence-based volcanology: application to eruption crises, Journal of Volcanology and Geothermal Research, 128, 273-285, 2003.
Gitis, V.G. and Ermakov, B.V., Osnovy prostranstvenno-vremennogo prognozirovaniya v geoinformatike (Fundamentals of spatio-temporal forecasting in geoinformatics), Fizmatlit, 256 pp, Moscow, 2004 (in Russian).
Jensen, F.V., Bayesian networks and decision graphs, Springer-Verlag, 268 pp, New York, 2001.

Newhall, C.G., and Hoblitt, R.P., Constructing event trees for volcanic crises: Bulletin of Volcanology, 64, 3-20, 2002.
Nikolenko, S.I., and Pshenichny, C.A., Temporal modeling by means of event bush. – European Geosciences Union (EGU) 2nd General Assembly, Vienna, Austria, April 2007, Geophysical Research Abstracts, Vol. 9, EGU2007-A-00497.
Pshenichny, C.A., Classical logic and the problem of uncertainty, Geological Prior Information, Curtis, A. and Wood, R. (Eds), Geological Society, London, Special Publications, 239, 111-126, 2004.
Pshenichny, C.A., and Khrabrykh, Z.V., Knowledge base of formation of subaerial eruption unit, Environmental Catastrophes and Recovery in the Holocene (Abstracts), Brunel University, London, September, 2002: http://atlas-conferences.com/cgi-bin/abstract/caiq-22
Pshenichny, C.A., Nikolenko, S.I., Carniel, R., Vaganov, P.A., Khrabrykh, Z.V., Moukhachov, V.P., Akimova-Shterkhun, V.L., Rezyapkin, A.A., The event bush as a semantic-based numerical approach to natural hazard assessment (exemplified by volcanology), Computers and Geosciences (in press), 2008.
Pshenichny, C.A., and Fedukov, R.A., Formal treatment of knowledge in water science by means of event bush. European Geosciences Union (EGU) 2nd General Assembly, Vienna (Austria), April 2007, Geophysical Research Abstracts, Vol. 9, EGU2007-A-01016.
Tutte W T , Graph theory as I have known it, Oxford University Press, 168 pp, New York, 1998.
Yakovlev, A.V., and Nikolenko, S.I., Primenenie kustov sobytii k zadacham geoinformatiki (Application of event bushes to solving geoinformation tasks). Proceedings of St. Petersburg State University of Information Technologies, Mechanics, and Optics (in press; in Russian), 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

Connecting Science and Decision-making: A Conceptual Framework through Organisation Knowledge Management

T.F. Siew[a]
[a] *Institute of Forestry Economics, University of Freiburg. Tennenbacherstr. 4, D-79106, Freiburg, Germany (tf.siew@ife.uni-freiburg.de)*

Abstract: There often exist controversies between scientists and practitioners with regard to the development and implementation of an applicable and transferable decision support systems (DSS) in the field of environmental or natural resources management. The challenge is to support decision-making through the implementation of a DSS across spatial scale as well as under different institutional and political conditions. Apart from the development and implementation of IT-based systems, the problem solutions also require holistic approaches that promote and enhance the cooperation and consensus between science and decision-making spheres. The objective of this paper is to develop a conceptual framework for integrating science and decision-making spheres through organisation knowledge management. Considering the multi-dimensional nature of the problems, the framework has been developed using multiple perspectives approach. It takes into account the elements that underlie the interfacing problems between science and decision-making spheres as well as knowledge construction and use from different perspectives. The understanding of the cognitive and socio-cultural elements are grounded by such concepts and theories as the paradigm lock, epistemic community and bounded rationality. Whilst, knowledge construction and use are elaborated through technical (scientific), organisational (societal) and personal perspectives. At this stage, the framework do not specify what should be done by whom. Instead, it provides an understanding about the issue through the analysis of the state-of-the-art. In addition, it also provides a setting for further study including micro studies on human decision-making and decision makers heuristics. Ultimately, the question of 'how' will be answered.

Keywords: Science-decision-making integration; paradigm lock; epistemic community; bounded rationality; organisation knowledge management.

1. INTRODUCTION

The cooperation or collaboration between science and decision-making spheres in environmental or natural resources management involves knowledge management through the implementation of supporting technologies. These spheres together form a socially constructed organisation with respective roles to play in the process of knowledge construction and use. Knowledge management in this organisation is not only about the concurrent management of content, culture, process and infrastructure [Chait, 1999], it also includes the management of social interaction between and within both spheres. Hence, organisation knowledge management could be used as a holistic approach to collaborate a wide spectrum of contributors and retrievers of knowledge resources ranging from people of different disciplines and portfolios to technologies that support the processes.

With regard to technologies, there often exist controversies between scientists and practitioners about the development and implementation of an applicable and transferable decision support systems (DSS) to support decision-making. On one hand, the application of complex environmental models has become an important part in decision- and policy-making processes [Verdenius and Broeze, 1999]. On the other hand, the implementation of DSS to facilitate decision-making remains a difficult issue [Giupponi et al., 2006]. The challenge concerns the generation of credible information and knowledge and the effective uptake of them (through the implementation of DSS under different institutional and political conditions) [Kolkman et al., 2005; Ballantine, 2005; Slob et al., 2007]. In this context, the difficulty from the construction of knowledge through the use of that knowledge is characterised as the interfacing problems between science and decision-making spheres.

In recent years, various approaches have been proposed to address the interfacing problems through the development of conceptual modelling. For example, the development of NetSyMoD methodological framework has been developed to facilitate the involvement of stakeholders or experts in policy- or decision-making processes [Giupponi et al., 2008]. The flexible and comprehensive framework uses a suite of ICT tools to tackle problems commonly encountered under integrated water management. Giupponi et al. [2008] concluded that more efforts are needed to strengthen the exchanges between research and policy spheres with regard to knowledge and technology transfer.

In this respect, the problem solutions require not only the implementation of new IT-based systems, but also the practice of knowledge management taking into account the important organisational aspects, particularly human and social issues [Kjærgaard and Kautz, 2008]. In other words, a systems thinking approach that account for multi-perspectives is needed to address the multi-dimensional interfacing problems. Based on this notions, the objective of this paper is to develop a conceptual framework of organisation knowledge management for integrating science and decision-making spheres. The framework has been developed incorporating the concept of multiple perspectives of the Unbounded Systems Thinking (UST) and the Knowing Organisation. The UST promotes the use of divergent thinking and perspectives of experts and non-experts in search of consensus on a solution and also on the pool of solution alternatives [Hall et al., 2005]. Whilst, the Knowing Organization proposed by Choo [1996] provides a unified view of the principal ways how an organisation can make use of information strategically. The ultimate aim of this approach is to strengthen the relationship among science, politics and public administration as well as interdisciplinary cooperation among scientific communities through collaborative construction and use of knowledge.

The development of the conceptual framework requires the understanding of the elements that underlie the interfacing problems as well as the understanding of how knowledge is constructed and used based on different perspectives. These understandings are elaborated in the following sections based on the review of published literature.

2. CONCEPTS AND THEORIES UNDERLYING INTERFACING PROBLEMS

A number of studies aiming at addressing the interfacing problems can be found in, *inter alia*, Mills and Clark [2001], Acreman [2005], Kolkman et al. [2005], and Slob et al. [2007]. They presented some insights about the problems based on technology innovations, organisational and/or individual dimension.

The systems approach requires multi-perspectives view about the issue. The following subsections provide an understanding about the cognitive and socio-cultural elements using such concepts and theories as the paradigm lock, epistemic community and bounded rationality. These elements are embedded in each of the dimensions.

2.1 The Paradigm Lock

The interfacing problems have been recognised by the UNESCO-HELP programme as the ‘Paradigm Lock’, in which science and decision-making spheres are locked into separate vicious circles [quoted in Acreman, 2005]. These circles are separated by a dramatic gap in

the knowledge, the aims and the way of thinking between these spheres. In addition, the language between those who analyse and provide disciplinary expertise (i.e. scientists) and those who decide (i.e. decision makers) also presents in the gap [Luiten, 1999]. Furthermore, scientists and water managers are driven by different forces, for instance legislation, transparency, consistency, funding and operation time scale.

According to Willems and de Lange [2007], the often limiting implementation of newly developed tools from the research community is the result of failing to take into account the needs of decision or policy makers at a sufficient level. Scientists seek the best theory to explain the data that is available; they are driven by innovation and understanding; they are concerned more with technical integrity; and their main performance indicator is publications that have been peer-reviewed by other scientists [Mills and Clark, 2001; Acreman, 2005]. As a result, the DSS developed by scientists have been classified somewhere between dilettantism and academic exercises [Biswas, 1975]. Whilst, water managers seek consistent methods and practical decision support tools to support decision-making.

Besides, science is often more comfortable in providing advice on what ought to be done and why, rather than practical advice on how it might be achieved [Boehmer-Christiansen, 1994]. Willems and de Lange [2007] argued that scientists view the end-user in the research project as the client for their research results. There is, however, a significant lack of transfer mechanisms that would allow passing the relevant information on to other stakeholders including policy makers and implementers on the ground.

The reasons affecting the uptake of scientific information as well as a lack of universal support for scientific input into policy making also include both contradictory science and uncertainty surrounding the available results. This has resulted in the lack of public confidence in scientific information. In this conjunction, decision makers also have difficulty in obtaining high-quality science at short notice [Slob et al., 2007].

On the other hand, Mills and Clark [2001] maintained that scientific information often is used in emotionally or politically laden natural resource management decisions. While some policy makers are unable to make use of highly technical advice, discrediting science and even the scientist is a strategy sometimes used by antagonists on both sides of the issues. This is because science applications to natural resource issues are usually done in the glare of public conflict and controversy [Mills and Clark, 2001]. In fact, DSS should be used as support systems but not decision makers [Courtney, 2001; Westmacott, 2001].

These debates indicate that there is a lack of coherent relationship between science and decision-making spheres. According to Willems and de Lange [2007], science-policy interrelationship is inefficient at this moment as it should/could be. Slob et al. [2007] also maintained that science and decision-making spheres are not well connected, although there is evidence that they sometimes are.

In the opinion of Acreman [2005], there is actually a continuum of expertise from basic to applied scientists through to water managers. Individual scientists producing research results along the spectrum from fundamental understanding to very applied. The continuum is however bound to complexity, which also introduces risk and uncertainty in the decision environment.

The ‘Paradigm Lock’ is closely related to the paradigm of the expertise community and the rationality of the decision makers. Their relation is illustrated in **Figure 1**. A deeper understanding about the paradigm of epistemic community and the bounded rationality of decision maker is provided in the following subsections. This understanding could provide a more founding explanation on the cognitive and socio-cultural aspect that exist in the science and decision-making spheres.

2.2 Epistemic Community

Decision-making process requires concerted efforts and combined expertise of a large number of specialists. These specialists include, for instance, economists, sociologists, ecologists, agriculturalists, foresters, wildlife biologists and planners. They all have a part to play in ensuring that the questions being answered are the appropriate ones, the widest possible range of options has been generated, and the likely consequences and necessary

contingencies have been predicted [Jeffers, 1988]. The formation of diverse epistemic communities may result in the emerging interfacing problems.

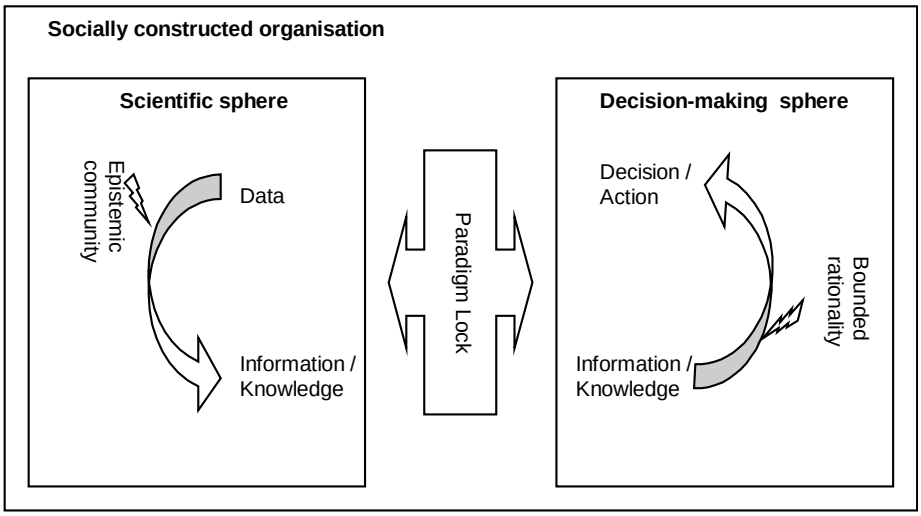


Figure 1. The connection between the 'Paradigm Lock', epistemic community and bounded rationality.

Epistemic community indicates a 'new' and in some aspects, atypical political actor. It constitutes of networks of experts coming with different experiences, from different backgrounds, a common interest, a shared task and diversity of knowledge [Cinquegrani, 2002]. An epistemic community as defined by Haas [1992] is a network of professionals from a variety of disciplines and backgrounds, who have (1) a shared set of normative and principled beliefs, which provide a value-based rationale for the social action of community members; (2) shared causal beliefs, which are derived from their analysis of practices leading or contributing to a central set of problems in their domain, and which then serve as the basis for elucidating the multiple linkages between possible policy actions and desired outcomes; and (3) shared notions of validity – that is a set of common practices associated with a set of problems to which their professional competence is directed, presumably out of the conviction that human welfare will be enhanced as a consequence.

Kolkman et al. [2005] maintained that the construction of knowledge within different paradigm groups leads to different interpretations of the problem situations. Each scientific discipline constructs its own models using its own paradigm. Consequently, this has also impeded true implementation of interdisciplinary methodologies and the development of generalised models [Norgaard, 1992]. Furthermore, Jakeman et al. [2006] noted that there are not only different paradigms and methods between biophysical scientists and social scientists, but also gaps in shared understanding between some of the major quantitative sciences.

The epistemic communities signify that individual community supports special interests better than collective ones [Norgaard, 1992]. They enable cohesion of a discourse and unite a community of their own followers. On the other hand, the uptake of scientific information provided by the paradigm groups is influenced by the rationality of a decision maker.

2.3 Bounded Rationality

Decision-making shares equivalent meaning with problem solving and management, which is the process of converting information into action. Although decision support tools can provide for rational information, the outcome of a decision is very much dependent on the rationality of the individual decision maker with regard to choice of information. According to Biswas [1975], management success depends on not only the quality and extent of the information available but also what information is selected for use and ultimately channelled into the decision-making process.

Hjorth and Bagheri [2006] maintained that any system in which humans are involved is characterized by the following essential system properties: bounded rationality, limited certainty, limited predictability, indeterminate causality, and evolutionary change. Bounded rationality as defined by Herbert Simon [quoted by Choo, 1996] is "the capacity of the human for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behaviour in the real world – or even for a reasonable approximation to such objective rationality".

Decision-making approaches could by and large be categorised into rational and bounded rational approach. Based on the rationalist approach, procedures for policy and decision-making usually require the collection of information to support the selection of a policy option. It assumes that a rational and therefore legitimate choice can be made (e.g. environmental impact assessment) [Slob et al., 2007]. In this approach, procedural uncertainty is managed through well-defined rules, routines and performances [Joshi, 2001]. On the other hand, the bounded rational approach suggests that the human way of thinking is not normative or rational but conditional. It means that humans use their whole life experiences to reach a decision in the real world. Moreover, individual taking this approach uses his power and influence to deal with conflict resolution, negotiation and compromise [Joshi, 2001].

The rationality of an individual and the resultant of his behaviour is also a crucial factor in affecting his relation with the science sphere. On one hand, individuals and organisations are forced to take standpoints and make choices based on uncertain knowledge and diverse views [Höijer et al., 2006]. On the other hand, each entity with capacity to make decisions and to carry out acts in a dynamic environment will face dilemmas between reconsidering the choice of action at each step based on newly perceived information. This process can be costly. However, unconditional commitment to chosen actions can lead to failure [Hall et al., 2005].

The divergent operational philosophies and socio-cognitive influence are the fundamental reasons for poor communication and interaction between scientists and decision makers. Norgaard [1992] maintained that the differences in the way different organisations transform data into information are the results of different assumptions, cultures, and paradigms within the disciplines. Whilst, the limitations of uptake of scientific information are attributed to bounded rationality of an individual. The embedded cognitive and socio-cultural elements are alike a shadow of an individual. Its visibility and the quantum of its presence (it could be quantifiable) is dependent highly on location and position where an individual is. The more precise understanding of these elements, coupled with the understanding of how knowledge is constructed and used, is imperative for the integration between science and decision-making spheres.

3. MULTIPLE PERSPECTIVES ON KNOWLEDGE CONSTRUCTION AND USE

Knowledge management has been vastly studied in business and information management fields. Various definitions about knowledge management can be found in Nevo and Chan [2007]. Knowledge management in an organisation is a process through which an organisation construct and use its institutional or collective knowledge. Choo [1996] noted that a solid understanding of how an organisation creates, transforms and uses information in an organisation is necessary. Failing of this understanding, an organisation would lack the coherent vision to manage and integrate its information processes, information resources and information technology.

According to Nemati et al. [2002], knowledge management is the practice of adding actionable value to information by capturing tacit knowledge and converting it to explicit knowledge; by filtering, storing, retrieving and disseminating explicit knowledge; and by creating and testing new knowledge. Sousa and Hendriks [2006] maintained that knowledge management addresses policies, strategies, and techniques aimed at supporting an organisation's competitiveness by optimising the conditions needed for efficiency improvement, innovation, and collaboration among employees. Knowledge management also involves organisational learning.

Knowledge construction may refer to knowledge creation and knowledge acquisition. Whilst, knowledge may be used or applied through a process of elaboration (the development of different interpretations), infusion (the identification of underlying issues), and thoroughness (the development of multiple understandings by different individuals or groups) [King et al., 2008].
Solving complex problems in a holistic system may require knowledge from any source and those knowledgeable in any discipline or profession with the support of technologies. In this conjunction, it is necessary to comprehend how knowledge is constructed and used based on technical, organisational and personal perspective, respectively.

3.1 Technical Perspective

Technical perspective of the UST relates that scientific technologies function with logic and rationality. In a rational environment, computer-based systems are used by scientists or paradigm groups to produce a series of possible rational (or right) problem solutions for analysing the situations [Hall et al., 2005].
The decision support systems (DSS) field has been recognised as dealing with such technologies for representing and processing knowledge in order to facilitate decision-making. It is believed that these technologies could provide useable knowledge at an appropriate point of decision-making process as well as at an appropriate level of precision [Giupponi et al., 2006]. Besides, DSS has also explicitly included decision evaluation in order to increase user satisfaction and better facilitate group discussion and compromise [Bell et al., 2003]. As a result of this functionality, DSS provide support to decision makers engaged in solving various semi- to ill-structured problems involving multiple attributes, objectives and goals.
In a complex environment, the variability, interdependency and uncertainty of factors affecting decision-making process are complex. DSS integrate data sources with modelling and analytical tools; facilitate development, analysis, and ranking of alternatives; assist in management of uncertainty; and enhance overall problem comprehension [Mowrer, 2000]. Hence, DSS functioning as an expert system can deal with the complexity of the decision problems through the enhancement of the limited capacity of the human mind. They may simulate or even replace human thinking and decision-making by preventing human shortcomings or the improvement of human characteristics [Bender and Manders, 1993].
DSS may support a right decision by providing rational information and knowledge. This includes information on possible outcomes of a decision as well as the values of the outcomes to the individual affected. This functions well in an ideal world, in which rational choice or rational decision-making could be made based on a complete set of available alternatives, reliable information about their consequences, and consistent preferences to evaluate these outcomes [Choo, 1996]. However, the types of information and knowledge used are often dependent on the cognitive level of an individual.

3.2 Personal Perspective

The types of data the actor perceives in the real world, as well as the types of knowledge the actor derives from the data are determined by the frames of perception of individual actor [Courtney, 2001; Kolkman et al., 2005]. These frames are recognised as mental model of an individual. The concept of a mental model is parallel to beliefs, i.e. they are continuously updated as the environment changes, yet the underlying foundations often remain unchanged over time [Hall et al., 2005].
In this respect, knowledge from the personal perspective concerns mainly about tacit knowledge that includes beliefs, perspectives, and mental models. According to Nemati et al. [2002], tacit knowledge consists of subjective expertise, insights and intuitions that a person develops from having been immersed in an activity or a profession for an extended period of time. The challenge of knowledge management is to integrate and implement tacit knowledge into the decision-making process. However, tacit knowledge is often so ingrained in an person's mind that they are taken for granted.

Acreman [2005] maintained that expert opinion is a form of 'best knowledge' possessed by an expert harnessed from their accumulated experience. On one hand, skills and knowledge of individual research scientist are desired to bear on controversial natural resource management policies [Mills and Clark, 2001]. On the other hand, the judgement of decision makers becomes a key element in decision-making with regard to the use of knowledge [Acreman, 2005]. Furthermore, individuals reacting on the UST's personal perspective may affect not only information seeking/sharing, but also problem reformulation and validation [Hall et al., 2005].

In order to ensure purposeful use, the knowledge resources need to be managed systematically and effectively in an organisation taking into account the elaborated perspectives as well as the cognitive and socio-cultural elements. The following subsection elaborates on how an organisation manages its knowledge resources.

3.3 Organisational Perspective

Knowledge has long been considered an important organisational resource [Nevo and Chan, 2007]. An organisation can be viewed as a distinct entity or an open system that manages knowledge resources. Knowledge of all types (i.e. technical and individual knowledge) must be supported in this environment. Courtney [2001] noted that all types of knowledge include tacit and explicit, deep and shallow, declarative and procedural, and exoteric and esoteric knowledge. In addition, the relationships and reciprocal influence between the organisation and the external environment also need to be considered during knowledge management process in an open system.

An organisation creates knowledge by developing new knowledge or replacing existing knowledge with new content through the implementation of technologies. It also acquires knowledge from individuals or through the search for, recognition of, and assimilation of potentially valuable knowledge from outside the organisation [King et al., 2008]. Choo [1996] maintained that knowledge creation is achieved through a recognition of the synergistic relationship between tacit and explicit knowledge in the organisation, and through the design of social process that create new knowledge by converting tacit knowledge into explicit knowledge.

In a decision-making environment, an organisational action is taken through the process of information interpretation, information conversion and information processing. This information is used by (an individual or) organisation to make sense of change in its environment, to create knowledge for innovation, and to make decisions about courses of action [Choo, 1996].

On the other hand, culture is also recognised as a knowledge resource of an organisation. According to Holsapple and Joshi [2001], cultural knowledge resource comprises basic assumptions and beliefs as well as an organisation's values, principles, norms, unwritten rules and procedure. The behaviours of the members of an organisation with regard to knowledge acquisition, sharing and internalisation are influenced by cultural knowledge. Therefore, it is important for researchers and practitioners to appreciate cultural knowledge resource.

The effective management of knowledge construction and use processes is important for creating and delivering relevant and useful information and knowledge by and to the right person at the right time. On the other hand, the extent to which an individual makes his knowledge available as an organisational resource depends heavily on managerial influences (e.g. leadership, reward systems, evaluation systems) [Holsapple and Joshi, 2001].

4. CONCEPTUAL FRAMEWORK OF ORGANISATION KNOWLEDGE MANAGEMENT

The understanding of the cognitive and socio-cultural elements that underlie the interfacing problems and the knowledge construction and use by different perspectives provide a setting for the development of the conceptual framework of organisation knowledge

management. The necessary understanding is important for an organisation to holistically manage its sense-making, knowledge building and decision-making processes [Choo, 1996]. The framework shown in **Figure 2** has been developed by incorporating the concept of multiple perspectives, i.e. technical, organisational, and personal perspective as proposed by the Unbounded Systems Thinking (UST) model and the Knowing Organisation. It illustrates the organisational structure of knowledge management in a decision-making system.

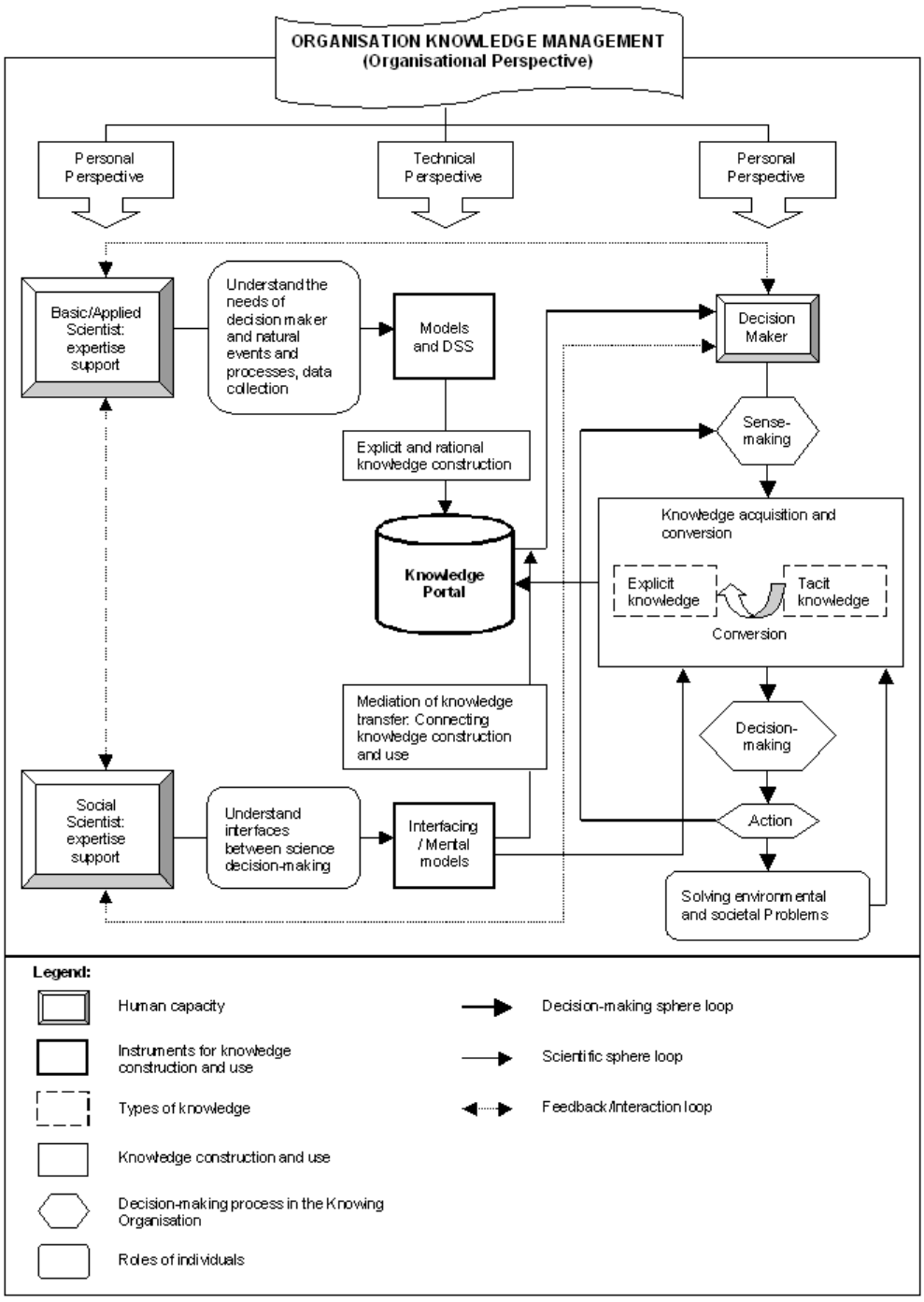


Figure 2. Conceptual framework of organisation knowledge management.

An organisation is framed as an knowledge management entity that manages the process of constructing and using the technical knowledge as well as its personal and institutional or collective knowledge. This organisation may be represented by an epistemic community,

which is a network of professionals and experts, who come from a variety of disciplines and backgrounds and who have a shared set of normative and principled beliefs [Haas, 1992]. In the ideal situation, the multi-person actor, who works within the framework of complexity and uncertainty, tries to re-define problems in broader context and attempts to comprehend 'change', and able to 'anticipate' using knowledge, various backgrounds and expertise [Cinquegrani, 2002].

The framework of organisation knowledge management emphasises on the relation and interaction between different perspectives and processes. According to Lovering [1999] and Lagendijk and Cornford [2000], it may be necessary to represent and clarify the relation between knowledge management, ICT usage and experts as mediators between the complexity of political decision and the tendency of institutions to become advanced learning organisation.

In this socially constructed organisation, the respective institutional functions and capacity of scientists and decision makers in a decision-making system are fundamentally determined by their respective roles in the decision-making system. The roles and capacity need to be clearly and effectively defined and communicated between science and decision-making spheres. In this respect, basic or applied scientists and social scientists provide expertise support to decision makers. Whilst, the decision makers have to make decision to come to an action and ultimately to solve environmental and societal problems.

The expertise support provided by scientists contribute to the decision-making process through facilitation of sense-making and knowledge creation stages. As explained by Cinquegrani [2002], the demand for the expert advice is a common phenomenon in policy-making processes at local, national and international level. Consequently, science has in fact undergone a major paradigm shift and moves from the traditional methods of production of scientific knowledge in the post-normal science era by taking into account its social and political context [Gibbons et al., 1994]. In this conjunction, science has been redefined as a social process, set in a social context, and involving actors and institutions and it is often called upon to provide solutions to societal problems [van den Hove, 2007]. In this context, social scientists should play an active role as mediators in the knowledge management process. They could strengthen the integration between science and decision-making spheres by facilitating the understanding of the interfacing problems as well as by implementing holistic approaches to address the problems. Social scientists contributes to the knowledge mediation process by developing conceptual models or problem structuring techniques to deal with the complex management problems.

Each personal in this organisation may seem to be driven by different forces in achieving respective aims and targets. However, they are interconnected by certain implicit forces as well as physical components in a multi-disciplinary environment as they attempt to solve common problems. In this framework, knowledge portal that stores explicit knowledge serves as a common platform of intellectual interaction. Models and decision support tools generate technical and rational knowledge that is reposited in the knowledge portal. In this regard, the needs of the decision makers should be considered sufficiently in developing models or DSS, as argued by Willems and de Lange [2007]. In order to improve the credibility of the rational knowledge, the technical tools that operationalise method into practice have also taken into account the issues of scales, uncertainty and risk.

The decision-making process incorporated in this framework is presented under the personal perspective of a decision maker. A decision maker makes use of the information and knowledge stored in the knowledge portal for sense-making and knowledge creation in order to make a decision or take an action [Choo, 1996]. During the decision-making process, tacit knowledge acquired by a decision maker should also be converted into explicit knowledge, which will continually enrich the content of the knowledge portal. Mental model is also used for this purpose, which could deal with the bounded rationality of an individual.

On the other hand, feedbacks or responses through communication among actors are represented by 'dotted-line arrow' in this conceptual framework. As maintained by Schwartz [2001], the need for suitable feedback is important in cognitive learning that deals with insights, reasoning and imagination. The process of feedback emphasises retrieval and extraction, association, repetition, recognition and the solution of problems. In addition,

Schwartz [2001] also noted that networks learn by changing the strengths of their interconnections in response to feedback and adaptive production systems.
The framework provides an overview and understanding about the organisational structure of knowledge construction and use process as well as the interrelation and interaction between different components and processes under technical and personal perspectives in a decision-making system. For addressing complex management problems using systems thinking, external implicit or explicit forces that have influences and impacts on each of the component and process in this organisation must also be considered. However, they are not illustrated in this figure.

5. CONCLUSIONS

The difficulties in the development and implementation of a transferable decision support systems could be addressed using systems thinking through the concept of multiple perspectives (i.e. technical, organisational and personal perspective). The conceptual framework of organisation knowledge management proposed in this paper illustrating the interaction between scientists and decision makers through collaborative knowledge construction and use processes. At this stage, the framework do not specify what should be done by whom. Instead, it provides an understanding about the issue through the analysis of the-state-of-the-art. Hence, the framework has been developed to provide a setting for further study on the cognitive and social-cultural aspects including micro studies on human decision-making and decision makers heuristics. Ultimately, the question of 'how' will be answered.
On the other hand, the successful implementation of the framework also requires further investigation and elaboration on the relevant tangible and intangible components and processes. The key components and processes may be changed, added, validated or improved.
Most importantly, it is necessary to instigate the awareness of the importance of forging and strengthening the relationship between science and decision-making spheres using holistic approach.

ACKNOWLEDGEMENTS

The author would like to thank DAAD for offering financial support for the purpose of pursuing a doctoral degree. The support provided by Prof Dr Gerhard Oesten and the constructive comments, inspiration and motivation given by Dr Frank Ebinger are also gratefully appreciated.

REFERENCES

Acreman, M., Linking science and decision-making: features and experience from environmental river flow setting, *Environmental Modelling & Software*, 20(2), 99-109, 2005.

Ballantine, B., Enhancing the role of science in the decision-making of the European Union, Working Paper, European Policy Centre, Brussels, Belgium, 2005.

Bell, M.L., B.F. Hobbs, and H. Ellis, The use of multi-criteria decision-making methods in the integrated assessment of climate change: implications for IA practitioners, *Socio-Economic Planning Sciences*, 37(4), 289-316, 2003.

Bender, J., and F. Manders, Expert systems and organizational decision-making, *Information & Management*, 25(4), 207-213, 1993.

Biswas, A.K., Mathematical modelling and environmental decision-making, *Ecological Modelling*, 1(1), 31-48, 1975.

Boehmer-Christiansen, S., Politics and environmental management, *Journal of Environmental Planning and Management*, 37(1), 69-85, 1994.

Chait, L.P., Creating a sucessful knowledge management system, *The Journal of Business Strategy*, 20(23, 1999.

Choo, C.W., The knowing organization: how organizations use information to construct meaning, create knowledge and make decisions, *International Journal of Information Management*, 16(5), 329-340, 1996.

Cinquegrani, R., Futurist networks: cases of epistemic community, *Futures*, 34(8), 779-783, 2002.

Courtney, J.F., Decision-making and knowledge management in inquiring organizations: toward a new decision-making paradigm for DSS, *Decision Support Systems*, 31(1), 17-38, 2001.

Gibbons, M., C. Limoges, H. Nowotny, S. Schwartzman, P. Scott, and M. Trow, The New Production of Knowledge, Sage Publications, London, 1994.

Giupponi, C., A. Fassio, J.F. Vàzquez, and J. Mysiak, Sustainable water management and decision-making, In: Giupponi, C., A.J. Jakeman, D. Karssenberg and M.P. Hare (Eds.), Sustainable Management of Water Resources: An integrated approach, Edward Elgar, Cheltenham, 2006.

Giupponi, C., A. Sgobbi, J. Myšiak, R. Camera, and A. Fassio, NETSYMOD - An integrated approach for water resources management, In: Meire, P.e.a. (Eds.), Integrated Water Management: Practical Experiences and Case Studies, Springer, 2008.

Haas, P., Introduction: epistemic communities and international policy co-ordination, *Int Org*, 46(1), 1-35, 1992.

Hall, D., Y. Guo, R.A. Davis, and C. Cegielski, Extending unbounded systems thinking with agent-oriented modeling: conceptualizing a multiple perspective decision-making support system, *Decision Support Systems*, 41(1), 279-295, 2005.

Hjorth, P., and A. Bagheri, Navigating towards sustainable development: A system dynamics approach, *Futures*, 38(1), 74-92, 2006.

Höijer, B., R. Lidskog, and Y. Uggla, Facing dilemmas: Sense-making and decision-making in late modernity, *Futures*, 38(3), 350-366, 2006.

Holsapple, C.W., and K.D. Joshi, Organizational knowledge resources, *Decision Support Systems*, 31(1), 39-54, 2001.

Jakeman, A.J., C. Giupponi, D. Karssenberg, M.P. Hare, A. Fassio, and R. Letcher, Integrated management of water resources: Concepts, Approaches and Challenges, In: Giupponi, C., A.J. Jakeman, D. Karssenberg and M.P. Hare (Eds.), Sustainable Management of Water Resources: An integrated approach, Edward Elgar, Cheltenham, 2006.

Jeffers, J., Decision-thinking about land use, *Land Use Policy*, 75-78, 1988.

Joshi, K.D., A framework to study knowledge management behaviors during decision making, the 34th Hawaii International Conference on system sciences, 2001.

King, W.R., T.R. Chung, and M.H. Haney, Knowledge management and organizational learning, *Omega*, 36(2), 167-172, 2008.

Kjærgaard, A., and K. Kautz, A process model of establishing knowledge management: Insights from a longitudinal field study, *Omega*, 36(2), 282-297, 2008.

Kolkman, M.J., M. Kok, and A. van der Veen, Mental model mapping as a new tool to analyse the use of information in decision-making in integrated water management, *Physics and Chemistry of the Earth*, 30(4-5), 317-332, 2005.

Lagendijk, A., and J. Cornford, Regional institutions and knowledge - tracking new forms of regional development policy, *Geoforum*, 31(2), 209-218, 2000.

Lovering, J., Theory lead by policy: the inadequacies of the new regionalism, *Int J. Urban Region Res.*, 23(2), 379-395, 1999.

Luiten, H., A legislative view on science and predictive models, *Environmental Pollution*, 100(1-3), 5-11, 1999.

Mills, T.J., and R.N. Clark, Roles of research scientists in natural resource decision-making, *Forest Ecology and Management*, 153(1-3), 189-198, 2001.

Mowrer, H.T., Uncertainty in natural resource decision support systems: sources, interpretation, and importance, *Computers and Electronics in Agriculture*, 27(1-3), 139-154, 2000.

Nemati, H.R., D.M. Steiger, L.S. Iyer, and R.T. Herschel, Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing, *Decision Support Systems*, 33(2), 143-161, 2002.

Nevo, D., and Y.E. Chan, A Delphi study of knowledge management systems: scope and requirements, *Information & Management*, 44(6), 583-597, 2007.

Norgaard, R.B., Coordinating disciplinary and organizational ways of knowing. Agriculture, *Ecosystems and Environment*, 42(1-2), 205-216, 1992.

Schwartz, H., Review: Herbert Simon and behavioral economics, *Journal of Socio-Economics*, 31(3), 181-189, 2001.

Slob, A.F.L., M. Rijnveld, A.S. Chapman, and P. Strosser, Challenges of linking scientific knowledge to river basin management policy: AquaTerra as a case study, *Environmental Pollution*, 148(3), 867-874, 2007.

Sousa, C.A.A., and P.H.J. Hendriks, The diving bell and the butterfly - the need for grounded theory in developing a knowledge based view of organizations, *Organizational Research Methods*, 9(3), 315-338, 2006.

van den Hove, S., A rational for science-policy interfaces, *Futures*, 39(7), 807-826, 2007.

Verdenius, F., and J. Broeze, Generalised and instance-specific modelling for biological systems, *Environmental Modelling and Software*, 14(5), 339-348, 1999.

Westmacott, S., Developing decision support systems for integrated coastal management in the tropics: Is the ICM decision-making environment too complex for the development of a usable and useful DSS?, *Journal of Environmental Management*, 62(1), 55-74, 2001.

Willems, P., and W.J. de Lange, Concept of technical support to science-policy interfacing with respect to the implementation of the European water framework directive, *Environmental Science & Policy*, 10(5), 464-473, 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 note on attitudes towards and expectation from the Decision Support Systems

J. Mysiak[a], C. Giupponi[a,b], Y. Depietri[a] and G. Colombini[a]

[a] *Fondazione Eni Enrico Mattei, Castelo 5252, Venice 30122, Italy (jaroslav.mysiak@feem.it, yaella.depietri@feem.it)*
[b] *Center for Environmental Economics and Management, Università Ca' Foscari di Venezia, S. Giobbe 873, 30121 Venezia, Italy (cgiupponi@unive.it)*

Abstract: For long time Decision Support Systems (DSS) have been believed convenient tools for transferring scientific knowledge, ready-to-use and free of political interference, onto policy-making agendas. The concept seemed to appeal to scientists and policy makers alike. The former embraced DSS as a way to get their voice heard, the later were believed to benefit from the easy-to-use interface to intricate models. The enthusiasm had not been shared by all scientists. Many complained about the limited uptake of the tools and their conditional suitability to address practical policy issues, messy and intractable as they often are. Different reasons were suspected for this drawback. In this paper we look into different pieces of evidence to find out what scientists, policy makers and users think of DSS, their limitation and hurdles which need to be overcome to put DSS at work. These sources include results of expert workshop, e-surveys targeted to DSS developers and users, interviews with scientific officers from the funding agencies, review of the EU FP Working Programmes and the databases of funded projects, and review of the scientific papers dealing with DSS.

Keywords: DSS; users' acceptance; implementation gap

1. INTRODUCTION

Assessing the extent to which diverse Decision Support System (DSS) lived up to the initial promises is a daunting task: many tools with little in common are referred to as DSS; there is little agreement on what constitutes the success or failure; and the importance of factors behind the success varies greatly from situation to situation.

Yet despite all these difficulties, it is important to study why and how decision support tools are used (or not used), by whom and for what purpose. Insights from these studies can help to assess the impacts and benefits to improved decision making, avoid mistakes and frustration, and restore reasonable expectations from scientists, scientific officers and policy makers.

In this paper we make sense of evidence from different sources, including results of an expert workshop, surveys targeted to DSS developers and users, in-dept interviews with scientific officers at the DG Research and the DG Environment, literature review and e-mails communication with users and developers. These sources are described in the annex. Our analysis is restricted to the DSS developed or used for water management. These systems typically integrate advanced modelling, simulation, optimization and knowledge-based tools with spatial data management functionality.

2. IN SEARCH OF DSS IDENTITY

There are many competing understandings of what DSS are or how they assist decision making. In Giupponi et al. (2007) we have explored this variety of meanings, suggesting that the common understanding is not strictly related to any of the proposed definitions. When asked to describe the tools in their own words, the DSS developers referred to expert involvement, supplementary techniques not translated into software code, tips and rules, and process management, in addition to different kinds of computerised tools.

The DSS occupy a niche at the border between science and policy; this niche is characterised by employing diverse software technology to support decision making. It is convenient to conceive DSS as a software instrument to transfer scientific knowledge, ready-to-use and unspoiled by political interference, onto policy-making agendas. The idea to wrap up up-to-date knowledge so that the complexity remains hidden appealed to scientists and policy makers alike. The former embraced DSS as a way to get their voice heard, the later were believed to benefit from the easy-to-use interface to intricate models.

However, there are several problems with this ambition: First, to justify efforts and means put into software development, the final product needs to be flexible enough to be applicable in a range of situations. This is at odds with traditional scope of DSS to address ill-defined problems which are hard to penetrate, complex to tackle and to a large extent unique [Cox, 1996]. Second, DSS as a nutshell of knowledge can work out only if the knowledge is well established and represent consensus among scientists. Not-yet-established knowledge, rife with uncertainty and thus contentious, although potentially indispensable for making sense of a policy situations, is not consistent with notion of credible software tools.

The issues of deep and unreducible (at least in short term) uncertainty and risk, along with role and accountability of scientists in situations in which these issues substantially affect policy making are at centrestage of science-policy dialog. Back in 1972, Alvin Weinberg called "trans-science" an area in which the essential issues seem to be resolvable by impartial knowledge, but which science for different reasons cannot unambiguously answer [Weinberg, 1972]. His and others' contributions fuelled a debate about the professional responsibility of scientists advising or taking part in policy debates for which they claim special expertise. That debate compelled the Operation Research Society of America to issue "Guidelines for the Practice of Operational Research" [Machol, 1972], suggesting ways how to reconcile scientific norms with adversary policy practices. More recently, similar debate had been held in environmental modelling disciplines, prompted by constructive critique of how mathematical simulation models were used to inform practical policy making [Edwards, 1996; Konikow and Bredehoeft, 1992; Oreskes et al., 1994; Pielke, 2003]. This debates prompted development of best practice guidance and help to sensitise modellers to issues and requirements of policy making.

Remarkably, these topics haven't been echoed in the DSS field, mostly concerned with usability of software and interfaces practical to conceal the complexity of scientific matter. The key topics discussed at the mainstream of the DSS field, although valuable for developing decision tools useful for practical purposes, featured dominantly software aspects and responsibility of scientists as tool developer, but not as scientific adviser. The major DSS topics include advanced software development methodologies, flexible enough to allow the intended users to constantly refine their requirements; assessment of usefulness and usability of the decision tools; critical success factors driving the diffusion and uptake of the tools for practical policy making; and advanced techniques for problem structuring and decision analysis, supporting different decision modes and styles.

To wrap up, many have tried to explain what the DSS are by listing what these tools should contain, or how they should inform the decision processes. We believe that the topics neglected in the DSS field but found important in other fields of science-policy interface are equally useful to explain the DSS and why they haven't been found easy to implement.

3. FUNDING OPPORTUNITIES

The past EU Framework Programmes (FP) funded many research projects which placed the DSS development prominently in their intents. But it wasn't until the FP5 (1998 - 2002) that the number of these projects reached a level never seen before. The FP5 Programme Energy, Environment and Sustainable Development (EESD) dedicated one of the six key actions to management and quality of water, under which the DSS development was explicitly encouraged. At that time the EU Water Framework Directive was yet to be adopted and the DSS were deemed useful tools to assist the water managers with the new requirements. Compared to three DSS projects funded under the FP4 (1994-1998), the FP5 supported twenty research projects explicitly meant to develop a DSS for practical water resource management. In the FP6 (2002 - 2006) the excitement dwindled away and applicants stopped placing the DSS as their foremost priorities. The funding instruments (integrated projects and networks of excellence) introduced in the FP6 encouraged large consortia and more ambitious objectives, making the decision support tools to take a back seat. Only four smaller projects (specific targeted research project STREP) placed the DSS more prominently.

The FP7 Working Programme 2007 mentioned DSS only once, asking for "decision support systems ... to allow stakeholders and decision makers to meet the often contradictory challenge of integrated resource planning without compromising natural resources of future generations (p. 31, FP7 ENV 2007). More generic terms such as decision support tools and decision support are mentioned a few more times, linking these tools to scenarios development, and prediction and adaptation to (expected) climate change impacts. The FP7 (2007 - 2013) puts emphasis on climate change research and impact assessment, i.e. topics which require examination of diverse bodies of knowledge which are rife with fundamental uncertainty and thus potentially contentious. Risk and hazard assessment, along with uncertainty analysis took the place once occupied by the decision support.

As the research priorities changed with time, so did the funding instruments and the research approach deemed most suitable. Along with these changes the FP7 encouraged engagement of civil society organisations (CSO) in research projects, creating so more and better opportunities to engage in a mutually beneficial dialog with scientists. The funding scheme "Research for the benefit of civil society organisations" introduces new participation modes and methods with higher potential in term of mobilisation and opening space for expressions. The projects funded under this scheme are composed by research organisations (RTD) and CSO who are who are equally involved in design, management and implementation of the research activities and their quality assurance.

The interviews with the scientific officers from the DG Environment and the DG Research shed light on the perceptions behind these changes. Operating at the science-policy interface, the interviewees appreciated the complexity of policy making processes more than scientists and DSS developers did. Because of this complexity, the DSS are believed more useful at lower (regional or local) levels of policy making, where the decisions are driven by technical considerations, rather then by political differences and positions. The appeal of DSS is related to the introduction of a coherent methodology and assessment techniques which yield comparable results across different regions. Many interconnected tools are preferred to monolithic, one-size-fits-all systems. The practical tools are expected to instigate dialog between the policy actor, to "simplify problems" and to test policy options. The tools' credibility is established by their transparency and ability to commence agreement and shared understanding. The low uptake of DSS and scientific advices is attributed to different cultures in science and politics/policy making. Any improvement would require greater scientific proficiency among policy makers and more openness to scientific information. The scientists are requested to better appreciate policy needs and tailor the scientific presentation for specific policy purposes. Most importantly, the scientists need to demonstrate suitability of scientific tools for any given policy requirement. Often, this demonstration is not sufficient and scientists are seen reluctant to engage in dialog with policy makers.

4. DSS DEVELOPERS PERSPECTIVE

Most of the approached developers appeared satisfied with their DSS related research and were not concerned about the fading prospects of DSS field. In contrary, most of the respondents were determined to do more research on the topic.

The respondents were convinced that existing DSS improved the communication between policy makers and scientists. The tools were also granted the ability to explore the issue from various standpoints, a feature which nearly everyone agreed is a defining attribute of DSS. But the respondent disagreed on whether these accomplishments are sufficient to improve decision making and to reconcile conflicts between the different legitimate views on what constitutes the problems and how it should be tackled.

The respondents apparently knew of or were actively involved in different cases in which DSS were applied successfully, and in others in which the DSS failed to meet the requirements of the intended users. The DSS failure were not necessary due to limited value of scientific knowledge to resolve policy issues, even if the most respondents acknowledged that the intricacy of policy processes were often neglected in scientific studies. Still, the respondents challenged the notion that the developed DSS remained mainly academic exercises with little impact on practical policy making.

Despite the mixed experiences, many respondents strongly believed that the academic research projects brought significant innovation and produced a host of benefits other than the software. Likewise, not all respondents believed that the implementation gap will diminish with greater IT proficiency of the policy makers or with technology advancement. Important achievements included assistance in problem exploring and facilitated exploration of different perspectives. On the other hand the past experiences of many have suggested that the DSS related efforts were not particularly useful to instigate interdisciplinary research.

Among the various DSS components, the data management, visualisation, model integration framework, algorithms to deal with qualitative data and uncertainty, as well as multicriteria assessment models received unambiguous support (see figure 1). The perceptions of other components' importance varied greatly among the respondents. Economic assessment techniques (e.g. cost benefit analysis), economic impact models and tools facilitating online discussion were encountered with strongest hesitation. Despite the fact that uncertainty analysis is ranked upon the most

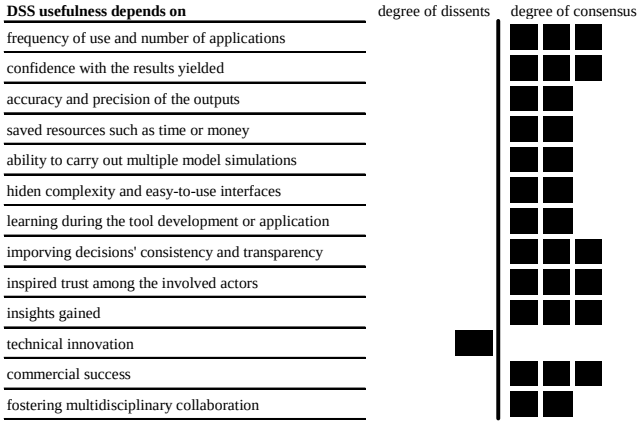


Figure 1. Respondents' perceptions of what constitutes success of DSS.

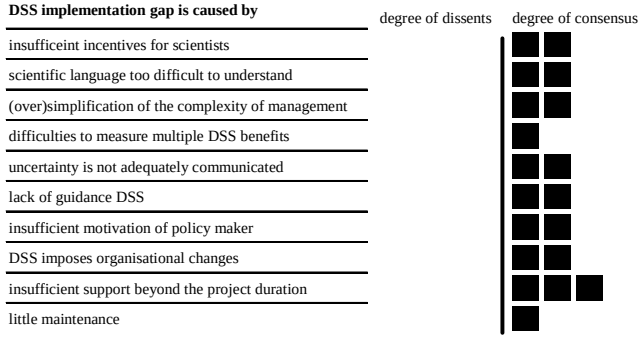


Figure 2. Reasons held responsible for low uptake of the

important components of DSS, the prevailing perception is that this topic is somehow neglected in practical applications.

Many factors have been found important for explaining the low uptake of existing tool for practical policy making (see figure 2). Among them, the quality of the developed tools divided the respondents, a fact which can perhaps be explained by different standards of quality and success. The academic institutional barriers and insufficient support beyond the project duration were among the factors which most respondents agreed on.

In their future research, most responders intended to address the implementation gap by paying attention to more user-friendly interface and by a closer collaboration with the end users. Only a few found it important to investigate the other barriers for limited diffusion of the available tools. The further progress in the field has however, not been linked to technology advances or increasing IT literacy among policy makers.

When reporting the achievement of DSS related projects, many scientists provided none or only constructed example. The paper which included description of real policy applications did not provide sufficient detail for an independent review and quality assessment. Stakeholder engagement was omitted completely or attached little importance. The evaluation of developed tools was hardly reported and only few articles included recommendations about what should be done to increase the uptake of the DSS.

5. USER PERSPECTIVE

Consistent with the expectations, our survey revealed a great variety of the DSS users and intermediaries. Among the users who responded to our survey were many agency scientists and scientific advisors with background in natural science. The other largest groups consisted of senior officials and high-level directors. Although most of them did not have prior experience with DSS, two quarters of the respondents interacted directly with the tools installed on their personal computers.

Their role in DSS application included most frequently the specification of the requirements, only half of the respondents were subsequently involved in the tools' practical application. Implicit expectation from many respondents was that DSS would facilitate stakeholders' engagement and help to identify policy options relevant for the given situations. Justifying decision taken in advance and satisfying regulatory obligations ranked fairly high. However, in the most cases the DSS application was an exercise or an exploratory study proposed by the academic partner. Less frequently, the efforts associated with the DSS development were justified by the new policy requirements and the limited skills of the previously used tools to support increasingly complex decisions. This explains the fact that the DSS, although dully maintained in the users' institutions, was used only in a few cases, often just in a single situation.

Contrary to what DSS developers believed, the users appeared convinced that the DSS improved the decision making compared to the approach which would have been applied otherwise. In a few cases the DSS were contested by those who were affected or involved in the decision. The satisfaction with different components of the systems appeared well balanced.

6. CONCLUSIONS

In this paper we have tried to combine diverse evidence to make sense of scientists' and users' opinions on DSS and motivations for their uptake or lack of thereof. More often then not the would-be-users were attracted into DSS projects by curiosity and interests to experiment new policy analysis tools. If the initial interests grew into durable commitment than the objective has been met. If this did not happen, for many possible reasons, than the exercise may still not have been useless.

While encouraging the DSS development, the research agencies intended to create institutional conditions for a greater engagement of scientists and policy makers in a mutually beneficial dialog. The decision tools had not been specified in detail in call for proposals, allowing the developers to design the content creatively. As the research

priorities shifted to areas where the knowledge is provisional and constantly evolving, as it is the case of climate change impact assessment, software tools appears less attractive as the instrument of knowledge transfer. The engagement of policy makers and wider stakeholders in designing research activities is, however, more imperative than anytime before. So is the scientists' contribution to policy debates, despite or perhaps because of the scientific uncertainty which underpins these debates.

The practical policy application of the developed DSS became unwritten demand and a way of practical corroboration of the developed tools. This is not wrong if the scientific articles provide sufficient detail for other scientists to review and assess the quality of the scientific advice given. Many reports however don't provide this detail. Innovation at the science-policy interface can have different forms: It can relate to improved content of the tools, their adaptation for specific purpose or study area, or their integration with other tools or modes of knowing. Alternatively, it can refer to how the tools have been used to inform policy making, how they facilitated dialog between the policy actors and how robust and lasting had been the agreement or consensus achieved. The latter requires new ways of engagement of scientists and policy makers/stakeholders, which is not possible without a greater understanding and appreciation of the policy process by scientists.

To facilitate these positive developments, we include a few recommendations that arose from the material analysed in this paper[1].

1) To facilitate independent review and reuse of tools developed; the scientists should cooperate more closely with open source software community. Software developed thanks to public research funds should be accessible for free and dully documented. Similar requests had been advanced by funding agencies (e.g. US National Health Institute) for scientific publications.

2) A repository of the software can help to keep track of the tools' further development and to facilitate experience sharing among users and developers.

3) Greater differentiation of funding opportunities and more incentives for greater involvement of stakeholders in planning and implementation of research activities. The new FP7 instruments (such as research for the benefits of civil society organisations) are a good example.

4) Better documentation of the results, including the stakeholder processes applied, and the critical scrutiny of the results yielded. The scientific community engaged in the development of decision tools should develop their own best practices and code of conduct in a similar way as this has been done in other policy oriented fields.

ANNEX: Sources on which the paper bases

1) A two-day expert workshop organised in Venice, October 2005 to review the experience from the DSS implementation and to examine the factors driving the success and failure. The workshop was attended by 15 distinguish researchers in DSS and environmental management field. Information about the workshop can be found under http://www.feem-web.it/dss-guide/

2) A survey targeting the DSS developers, consisting of 69 questions. Thirty eight respondents completed the questionnaire, making the return rate ca. 20%. The approached respondents included participants at the aforementioned workshop, lead authors of the DSS related articles published in the Journal of Environmental Modelling and Software and in the proceedings from the congresses of the International Society for Environmental Modelling and Software (iEMSs). A copy of the questionnaire is available at http://www.feem-web.it/dss-guide/survey.html.

[1] These recommendations are addressed more in detail in an extended version of this analysis prepared for a journal submission.

3) A survey targeting the DSS users. The questionnaire containing 26 questions was sent out by e-mail to DSS users suggested by the DSS developers. The questionnaire is described in Giupponi et al. (2007), a copy is available upon request.

3) In-dept review of selected papers published in the proceedings of the International Society for Environmental Modelling and Software (iEMSs), from the congresses in Lugano (2002), Osnabrueck (2004) and Burlington (2006). A sample of sixty papers had been drawn from a total of 141 papers referred to DSS.

4) Semi-structured interviews with the EU scientific officers from the Directorate General (DG) Environment and the DG Research. Further insights have been gained from less formal interviews and e-mail exchange with DSS developers and DSS users.

5) Analysis of the research projects funded under the FP4, FP5 and FP6, along with the review of the corresponding working programmes: The FP5 Programme "Energy, Environment and Sustainable Development" especially the key action line "Sustainable Management and Quality of Water"; FP6 Programme "Global change and ecosystems", and FP7 Programme Cooperation, thematic priority Environment (including climate change).

REFERENCES

Cox, P. G., Some issues in the design of agricultural decision support systems, *Agricultural Systems*, 52(2-3), 355-381, 1996.

Edwards, P. N., Global comprehensive models in politics and policymaking, *Climatic Change*, 32, 149-161, 1996.

Giupponi, C., Mysiak, J., and Depietri, Y., IWA Publishing. Harmoni-CA Report. Decision Support Systems for water resources management: current state and guidelines for tool development by. ISBN: 1234567890. Published by IWA Publishing, London UK, 2008.

Konikow, L. F. and Bredehoeft, J. D., Ground-water models cannot be validated, *Advances in Water Resources*, 15, 75-83, 1992.

Machol, R. E., The Obligations of Scientists as Counsellors, *Minerva*, 10(1), 107-157, 1972.

Oreskes, N., Shrader-Frechette, K. and Belitz, K., Verification, validation, and confirmation of numerical models in th earth sciences, *Science*, 263, 641-646, 1994.

Pielke, J. R. A., 2003, The role of models in prediction for decision, in Canham, C. and Lauenroth, W. eds.) Understanding Ecosystems: The Role of Quantitative Models in Observations, Synthesis, and Prediction, Princeton University Press, Princeton N.J., pp 113-137.

Weinberg, A. M., Science and Trans-Science, *Science*, 177(4045), 209-222, 1972.

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

Assessing the impact of environmental decision and information support tools

B.S. McIntosh and E. Diez

Centre for Water Science, Cranfield University, College Road, Cranfield, Bedfordshire MK43 0Al, United Kingdom (b.mcintosh@cranfield.ac.uk)

Abstract: A key aim of developing environmental models and software is to provide decision and information support to environmental policy and management. In developing such technologies we, as a community of scientists and computer specialists, hope to provide tools which exert a positive impact on policy and management processes, actions and outcomes. We want to contribute to a diverse range of objectives from better managing scarce resources, through mediating and avoiding conflict, to maintaining adaptivity and promoting sustainable development. As ever greater numbers of decision and information support tools (DISTs) are developed to support environmental decision making it is therefore important that we develop a clear understanding of the impacts we intend these technologies to exert, compared to the impacts they actually exert. So what impacts do environmental DISTs have on policy and management organisations and activities? The IS literature distinguishes between impacts based on scale (HCI – task – organisational) and life cycle stage. Within these categories internal (organisational) vs. external (action outcome) impacts, and perceived (e.g. perceived usefulness) vs. objective (e.g. reduction in task time) can be discerned. Environmental DIST developers tend to focus on external impacts to the neglect of internal impacts and the fact that organisations are typically the main means of policy and management action and therefore critical elements in the chain between information and outcome. Taking desertification as a case study area of environmental policy and management a questionnaire based study has shown that DISTs are viewed on the whole as yielding positive internal (efficiency) and external (effectiveness) benefits but that a range of negative internal impacts are of concern including training, cost and the need for organisational change. These are discussed.

Keywords: decision and information support tools; information systems; impact assessment; desertification.

1. INTRODUCTION

Key issues facing environmental policy and management organisations typically involve interactions between human and environmental processes which are spatially distributed (e.g. demography and river hydrology) and operate at different rates (e.g. urban expansion vs. flooding). Furthering the complexity facing such organisations, environmental issues may involve interactions between different processes at the same or different spatial (extent) and temporal (rate) scales located in hierarchical and nested structures [Holling *et al.* 2002].

Managing complex environmental issues therefore requires good quality information and the ability to manipulate such information to infer the potential outcomes of intervention *ex ante*. In this context decision and information support tools (DISTs) like GIS, computer models and decision support systems (DSS) can offer potential benefits to policy and management organisations through providing environmental information storage and analysis capabilities [Dale and English 1999]. GIS can be used to store spatial information on human and environmental processes and to derive useful indicators for management e.g. the calculation of flood risk from information on soil type, topography and river flow.

Computer models can be used to simulate human and environmental processes for the purpose of determining the likely outcomes of intervention *ex ante*. DSS can be used to evaluate options to identify optimal interventions e.g. the set of abstraction volumes along a river which yield maximum utility across all users.

However, concerns have been raised regarding the level of use of such tools and the possibility that a gap exists between the supply of information services from the academic research community in the form of DISTs and the actual information needs of environmental policy and management organisations [McIntosh *et al.* 2005, McIntosh *et al.* 2008]. It is clear that there is a strong need to interact closely with end-user organisations during development to ensure the usefulness of DISTs [Gottesdiener 2002, van Delden *et al.* 2007] But what impacts do DISTs have when they are used? Are they positive? Do they match with the expectations and ambitions of the designers? If not, what can be done?

To help answer these questions this paper will identify and characterise the kinds of impact that DISTs may have on environmental policy and management in terms of organisations and outcomes. This will be achieved through identifying and discussing key concepts and findings from the information systems (IS) literature along with some results from surveys of and interviews with organisations involved in developing and delivering desertification policy and management. These findings will then be discussed in relation to the kinds of impact which are typically expected by environmental DIST developers.

2. TYPES OF IMPACT IDENTIFIED IN THE IS LITERATURE

The IS literature provides a rich seam of theoretical and empirical knowledge regarding the use and impact of computer-based tools like DISTs. Indeed, we are classifying DISTs here as a type of IS – computer applications designed to provide information to groups of people acting purposefully [Checkland and Holwell 1999].

The IS literature has not been concerned with understanding how IS deliver external impacts e.g. resolving water shortages. Rather it has focussed on how IS lead to internal impacts within the context of different organisations e.g. cost reductions, productivity increases. As a consequence the kinds of impacts considered may not appear immediately relevant to the environmental context. However they help focus our attention on the means by which environmental policy and management outcomes are achieved – through organisation action.

In terms of characterising the impacts of using IS at the individual task level most of the characteristics are focussed on 'productivity change' (e.g. [Torkzadeh and Doll, 1999]) or improving 'performance' at the organisational level (e.g. [Tallon, 2000] [Seddon, 2002]) often in terms of efficiency [Andreu and Ciborra, 1996; Davis, 1989; Fiorito *et al.*, 2002; Nicolau *et al.*, 1995]. Efficiency is used here in the sense of the capacity of an IS to reduce the time and organisational costs or resources required to achieve aims and objectives.

However some studies go beyond the productivity factor at the individual level [Torkzadeh and Doll, 1999; Doll and Torkzadeh, 1998; Checkland and Holwell, 1999] to look at impacts including:

- Effectiveness: Ability of an application to help individuals achieve their aims.
- Task innovation: 'The extent to which an application helps users to create and try out new ideas in their work'.
- Management control: 'The extent to which an application helps to regulate work processes and performance'.
- Better understanding of the problem under analysis: 'The extent to which a computer-based information system is used to analyse cause-effect relationships'.

IS can also have an impact on individual perceptions and in doing so will significantly influence adoption outcomes. The two main perceived impacts that have been shown to be important to adoption are (i) perceived ease of use; and (ii) perceived usefulness, or

perceived relative improvement in users' job performance [Davis, 1989; Al-Gahtani and King, 1999].

At the organisational scale DeSanctis and Poole [1994] have framed the positive impacts of using IS in terms of improving decision outcomes - efficiency, quality, consensus and commitment.

3. DISTINGUISHING IMPACTS

Over the previous approximately four decades of IS research a number of important concepts and distinctions have been developed. Some of these are fundamental to understanding information use in organisations such as the distinctions made between data, information and knowledge by Checkland and Holwell [1999]. Others are more directly related to the nature and location of impacts exerted by IS such as the distinction between human-task-organisational levels made initially by Eason [1991] and the distinction between perceived benefits and objectively measurable benefits which lies at the core of major theories of IS adoption [Davis 1989]. We will focus on describing these concepts and distinctions here, first of all examining where impacts are exerted in terms of scale and location within an organisation, then in the context of IS (and by inheritance DIST) life cycle processes.

3.1. Scale of impact

It is necessary to have an appreciation of organisational structure and process when developing IS. There is a vast literature on organisational theory reaching back as far as the 1940s [Simon 1997] but only a few distinctions need be made to better appreciate IS impacts in terms of location.

It is useful to start by considering how work is undertaken by organisations. Checkland and Holwell [1999] distinguish between action (a set of activities linked together to pursue a purpose e.g. identifying areas of the world at risk of desertification) and activities (sets of tasks linked together to provide part of the transformation involved in a particular piece of action e.g. analysing data on vegetation state). The final level implicit in their conceptualisation is the task level – what individuals in organisations physically do (e.g. purchasing remote sensing data from a provider).

Following Eason [1991], Torkzadeh and Doll [1999] and Raz and Goldberg [2006] three main levels of analysis (or impact locations) can be distinguished when examining the organisational use of IS – (i) the **human-computer interaction (HCI) level**, which is concerned with how input and output device functionality including keyboard, mouse, graphical user interface layout etc. influence the performance of human interaction with computer systems generally (i.e. not in relation to particular tasks); (ii) **the task level**, which is concerned with how the IS attributes influence the performance of particular tasks by individuals, and; (iii) **the organisational level**, which is concerned with how IS and task attributes influence the performance of activities by groups, and across a whole organisation. IS may exert positive or negative impacts at each level in various ways. For example, at the HCI level an IS may be quicker or slower to use, with consequent impacts on the level of training required, and the efficiency of tasks and activities.

Improvements at the HCI level (e.g. controlling a PC through windows rather than DOS) don't however necessarily translate into positive impacts at the task level. Here impacts will relate to the fit between the structure of the task in terms of information inputs and outputs and the ability of the IS to satisfy inputs and handle outputs. Also the fit between the information demands of the task and the information quality attributes of the IS such as accuracy, trustworthiness, completeness, conciseness and objectivity [Nicolaou *et al.* 1995] will be important in determining the nature of the impact. Impacts at the task level might include changes to effectiveness or efficiency, output quality and reliability, learning, task restructuring and the generation of new ideas, which might also be termed learning.

However the information demands of tasks don't exist in isolation. Rather they exist within an organisational context – actions and activities (what Eason calls the 'organisational task'), a social system (other individuals) and a technical system (e.g. the IT infrastructure). Impacts at an organisational scale may be positive in terms of enhanced effectiveness or efficiency and cost reduction, they may involve costs arising from organisational change (restructuring actions, activities or tasks) or may involve costs arising from the need to build capacity (e.g. staff training or purchasing of new hardware).

It is worth distinguishing here between internal and external impacts. **Internal** impacts of DIST use are those such as have been just described – impacts within organisations on the structure and performance of tasks and activities by individuals and groups. **External** impacts can be distinguished as impacts on policy and management objectives and intervention outcomes – impacts felt externally from the organisation using the DIST. Citing work by Sterk *et al.* [2006], McIntosh *et al.* [2008] have argued that it is difficult to isolate the effects of DIST use on policy or management outcomes but that it is possible to isolate internal impacts such as learning within an organisation.

3.2. IS life cycle processes

In addition to level or scale, impacts can be distinguished in terms of where they lie within the overall life cycle of an IS. The IS literature distinguishes between a number of different processes across the life cycle of an IS from design to abandonment. The various processes identified can be grouped into three life cycle stages – pre-implementation, implementation and post-implementation. Figure 1 depicts the relations between these processes.

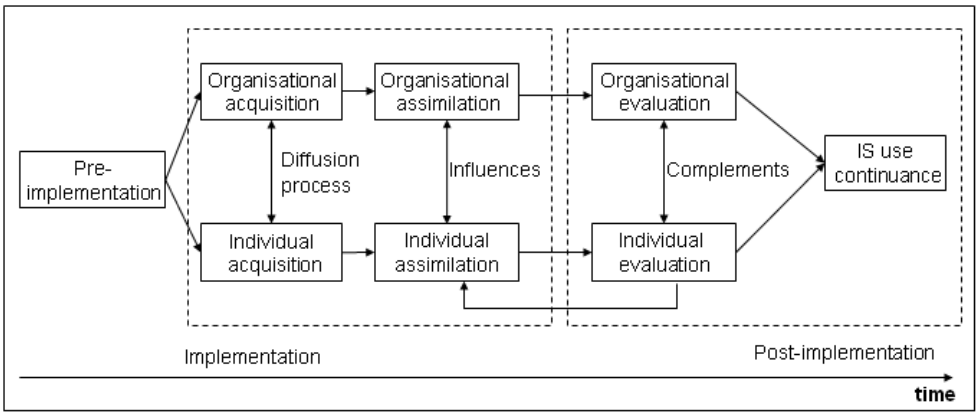


Figure 1. Processes involved in IS development and use

Pre-implementation is framed as the process of design and development prior to adoption by individuals or organisations [Hevner *et al.* 2004]. Implementation is framed as the process by which IS are taken up and used by organisations and may occur in various ways from bottom-up (driven by individual adoption behaviour to try out new technologies) or top-down (driven by organisational scale decisions to purchase software for example) involving processes of acquisition (influenced by diffusion) and assimilation into working practices [Al-Gahtani and King 1999, Fichman and Kemerer, 1997 Karahanna et al., 1999, McFarland and Hamilton 2006]. Post-implementation is used to describe the set of processes involved in continued use of a DIST and is viewed by some researchers as an extension of implementation [Liao *et al.* 2007]. It is framed in terms of individual scale evaluation of the benefits of using an IS, which may be largely informal, and organisational evaluation involving a more formal process of appraising the costs and benefits of IS use [Fitzgerald 1998, Kumar, 1990, Smith and Smith, 2007]. Table 1 shows how different IS life cycle processes can be categorised into these stages.

Table 1. Information system life cycle processes and their categorisation

Life cycle process	Process category

Design	Pre-implementation
Diffusion	Implementation
Adoption	Implementation
Implementation	Implementation
Acceptance	Implementation
Use	Implementation
Evaluation	Post-implementation
Assessment	Post-implementation
Continuous use	Post-implementation

For the outcomes of adoption to be successful (i.e. for an IS to be taken up and used by members of an organisation) it has been shown empirically that the members of that organisation involved must perceive that there are positive impacts (benefits) to themselves for example in terms of improving work efficiency or effectiveness [Davis 1989, Al-Gahtani and King 1999, Venkatesh *et al.* 2002]. The importance of perceived impacts within post-implementation has also been argued for [Liao *et al.* 2007].

It is therefore necessary when considering the impacts of DIST to separate out those which might be termed **perceived** (i.e. the impact that a DIST has on an individual's perceptions regarding its usefulness to that individual in his/her work) from those which might be termed **objective** and relate to measurable impacts such as productivity gains or greater achievement of policy objectives. Clearly most of the impacts that a DIST will exert will occur during the implementation and post-implementation stages, but important perceived impacts can be exerted during pre-implementation as a consequence of participatory and user oriented development approaches being employed.

4. CASE STUDY – DIST IMPACTS WITHIN DESERTIFICATION POLICY AND MANAGEMENT ORGANISATIONS

4.1. Introduction

Following the definition provided by UNEP in 1994, desertification is defined as 'land degradation in arid, semi-arid and dry sub-humid areas resulting from various factors, including climatic variations and human activities' causing the loss of soil productivity and poverty. It is a significant problem with FAO, UNESCO and WMO estimating that over 1.5 billion people are affected. It is also representative of the kinds of complex environmental issues that DISTs might help us understand and manage effectively.

However, like most areas of environmental policy and management there is little or no literature examining the kinds of DISTs which are used, the role that they play or the impacts that they have. To help shed light on these points a questionnaire was designed and delivered via the web to build up a picture of DIST use and impact within desertification policy and management organisations across the globe.

Purposive sampling was employed to identify organisations to send a request to complete the questionnaire. The key criterion employed was current involvement in desertification policy or management activities (e.g. conservation, rehabilitation, assessment, monitoring) in arid or semi-arid areas. Desertification policy and management involves a diverse range of organisations ranging from government departments and agencies through NGOs, Universities, research centres and businesses to United Nations agencies (e.g. UNCCD, FAO, GEF etc.). All were included if they satisfied the main selection criterion. Over 400 email requests to complete the questionnaire were sent out to identified individuals within each organisations spread across the organisational types and geographic areas as follows:

- Government (64%), UN (18%), universities (9%), NGO (4%), trade unions (3%), private (1%).
- Europe (40%), international (31%), Africa (12%), America (10%), Asia (7%)

90 responses were received, with 5 discarded due to their being incomplete. Respondents were spread as follows across organisational type and role, respondent position and geographic location (all declared by the respondent):

- Government (50.6%), universities (16.9%), NGOs (12.9%), UN agencies (11.8%)
- Research (52.3%), policy and management (39.5%), education (20%), technical assistance (7%), commercial activity (2.3%)
- Operational / technical (18.6%), senior management (16.3%), middle management (12.8%), directors (10.5%), co-ordinators and supervisors (9.4%), researchers (5.9%)
- Europe (46.5%), international (12.8%), Africa (11.6%), Asia (9.3%)

4.2. Results

The full results cannot be presented here due to space limitations so only those relevant to impact will be discussed. Table 2 shows which DISTs are used by desertification policy and management organisations. 76.5% of respondents indicated that their organisation use some form of DIST with GIS being the most commonly used tool and DSS the least commonly used. 23.5% of respondents did not use any DIST for a range of reasons which will not be covered here.

Respondents were asked to rate the benefits provided by the DISTs their organisation uses in terms of effectiveness (defined as the ability of the tool to help achieve organisational aims) using a 6 point likert scale (strongly agree – strongly disagree plus 'no answer') and efficiency (defined as the ability of the tool to help achieve aims more quickly) using a binary true / false. 72.3% of respondents strongly agreed that DISTs contribute to organisational effectiveness, with only 1.5% disagreeing. 95.4% agreed with the statement that the DIST they use contribute to improving organisational efficiency, with 4.6% disagreeing. Of course these responses only came from those respondents who use DISTs so do not reflect the views of those who do not.

Table 2. Use of DISTs by desertification policy and management organisations

Types of computer-based IS	Percentage
Geographical Information Systems	64.7
Remote Sensing	47.1
Statistical models	37.6
Simulation models	35.3
None	23.5
Decision Support Systems	20.0

The three most frequently mentioned benefits arising from the use of DISTs were speed of information processing (21.5%), data management capacity (16.9%) and improved understanding of (environmental) processes (12.3%). When asked to rate the benefits of DISTs over other sources of information respondents listed greater processing speed (20%), greater information reliability (13.3%), enhanced ability to convince others of output (13.3%) and enhanced effectiveness as a consequence of using them (13.3%).

For those respondents who do not use DISTs to handle information the most frequent three reasons why were a lack of skills within the organisation (35%), poor cost-effectiveness (26.7%) or the respondents being at a 'policy level' where information is provided by other organisations (26.7%). Other sources of information used were found to be external literature (65%), followed by expert consultation (60%), internal literature (60%), hard-copy maps (50%), spreadsheets (50%), mathematical or statistical analyses (35%) and laboratory experiments (10%), with the reasons why listed including good information reliability and detail, satisfaction of organisational information demand, contribution towards achieving organisational goals, cheap, easy to understand and no particular skills being required or skills already exist or no training required.

Respondents were asked to list management issues which reduce the level or extent of DIST usage. The responses are essentially concerned with the organisational changes required to implement new IS – need for training courses (46.7%), need for additional financing

(26.7%), the need for capacity building within the organisation (26.7%), a need to employ more staff (20%), additional time costs (13.3%) and the need to establish new protocols of work (6.7%).

5. CONCLUSIONS

From the desertification case study it can be seen that DIST impacts are largely viewed as positive, covering both internal (e.g. efficiency) and external (e.g. effectiveness) types. However it is also clear that a substantial number of organisations do not use DISTs, partly because other information sources are seen as adequate, and partly because of concerns about the impact of adopting DISTs in terms of organisational change. Adopting DISTs is seen by these people as involving additional cost burdens including training, hiring new staff, the financial cost of purchase, the need to restructure tasks and the potential that DISTs might make organisational performance slower.

It is here that we return to the point made earlier about the type of impacts that DISTs have on organisations. Typical environmental DIST developers are thinking about how to improve policy and management outcomes including for example better managing scarce resources, maintaining adaptivity and promoting sustainable development. Attention is generally not paid to the means by which such outcomes arise – and these means are, largely, organisations involved in formulating and delivering policy or management action.

It is clear that negative internal impacts, particularly around cost and training, are key organisational concerns within the context of desertification and need to be addressed. Concerns may be assuaged by measures such as incorporating and delivering good quality low cost training as part of the DIST 'package' sold to client organisations, or by calculating the costs involved in adapting the client organisation to the DIST and identifying ways in which the DIST design might be changed to minimise the need for organisational (activity and task level) change. Of course, communicating and convincing members of the client organisation of the benefits of DIST adoption will also be of use, but will require thought about what those benefits will be internally as well as externally, and some measurement of how they stack up against the necessary costs of implementation. Such analysis of costs and benefits is not typically done within the realm of environmental DIST development, and is an area for development.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support of the EC through the DeSurvey project (IP-003950).

REFERENCES

Al-Gahtani, S.S., King, M. Attitudes, satisfaction and usage: factors contributing to each in the acceptance of information technology. *Behaviour and Information Technology* 18 (4) 277-297, 1999.

Andreu R., Ciborra C., Organisational learning and core capabilities development: the role of IT. *Journal of Strategic Information Systems*, 5, 111-127, 1996.

Checkland, P. and Holwell, S., *Information, Systems and Information Systems*, Wiley, Chichester, 1999.

Dale, V.H. and English, M.R. (eds.), *Tools to aid environmental decision making*, Springer, New York, 1999.

Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 318-340, 1989.

DeSanctis, Geradine & Poole, Scott, M., Capturing the complexity in advanced technology use: Adaptive structuration theory. *Organization Science*, 5(2), 121-147, 1994.

Doll, W.J., Torkzadeh, G., Developing a multidimensional measure of system-use in an organizational context. *Information & Management*, 33(4), 171-185, 1998.

Eason, K.D.. Ergonomic perspectives on advances in human-computer interaction. *Ergonomics*, 34(6), 721-741, 1991.

Fichman, R.G., Kemerer, C.F., The assimilation of software process innovations: an organisational learning perspective. *Management Science* 43 (10) 1345-1363, 1997.

Fiorito, J., Jarley, P., Delaney, J., Information technology, US union organizing and union effectiveness. *British Journal of Industrial Relations*, 40(4), 2002.

Fitzgerald, G., Evaluating information systems projects: A multidimensional approach. Journal of Information Technology 13 (1) 15-27, 1998.

Gottesdiener, E., *Requirements by collaboration*. Addison-Wesley, New York, 2002.

Hevner, A., March, S., Park, J. Design Science in Information Systems Research. *MIS Quarterly* 28 (1) 75-106, 2004.

Holling, C.S., Gunderson, L.H. and Peterson, G.D., Sustainability and Panarchies, In: Gunderson, L.H. and Holling, C.S (eds.), *Panarchy, Understanding transformations in human and natural systems*, Island Press, New York, 2002.

Karahanna, E., Straub, D., Chervany, N., Information technology adoption across time: a cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly* 23 (2) 183-213, 1999.

Kumar, K., Post Implementation Evaluation of Computer-Based Information Systems: Current Practices. *Communication of the ACM* 33 (2) 203-212, 1990.

Liao, C., Chen, J.L. and Yen, D.C., Theory of planning behaviour (TPB) and customer satisfaction in the continued use of e-service: An integrated model. *Computers in Human Behaviour* 23(6), 2804-2822, 2007.

McFarland, D.J., Hamilton, D., Adding contextual specificity to the technology acceptance model. Computers in Human Behaviour 22 (3) 427-447, 2006.

McIntosh, B.S., Giupponi, C., Voinov, A.A., Smith, C., Matthews, K.B., Monticino, M., Kolkman, M.J., Crossman, N., van Ittersum, M., Haase, D., Haase, A., Mysiak, J., Groot, J.C.J., Sieber, S., Verweij, P., Quinn, N., Waeger, P., Gaber, N., Hepting, D., Scholten, H., Sulis, A., van Delden, H., Gaddis, E., Assaf, H., Bridging the gap: developing tools for environmental policy and management, In: Jakeman, T., Rizzoli, A., Voinov, A. & Chen (eds.). *State of the Art and Futures in Environmental Modelling and Software*, Elsevier, in press.

McIntosh, B.S., Jeffrey, P., Lemon, M. and Winder, N., On the Design of Computer-Models for Integrated Environmental Science, *Environmental Management* 35(6), 741-752, 2005.

Nicolaou, A., Masoner, M. M., Welker, R.B., Intent to enhance information systems as a function of system success. *Journal of Information Systems*, 9(2), 93-108, 1995.

Raz, O. and Goldberg, A., How to make IT work: Cognitive Perspectives for Better Information Technologies Performance. *European Management Journal*, 24(2-3), 199-205, 2006.

Seddon, P.B., Graeser, V., Willcocks, L.P., Measuring organisational IS effectiveness: An overview and update of senior management perspectives. *Database for Advances in Information Systems* 33(2), 11-28, 2002.

Simon, H.A., *Administrative Behaviour, A Study of Decision-Making Processes in Administrative Organizations*, The Free Press, New York, 4th Edition, 1997.

Smith, J. and Smith, P., *Environmental modelling: an introduction.* Oxford University Press: Oxford, 2007.

Sterk, B., van Ittersum, M.K., Leeuwis, C., Rossing, W.A.H., van Keulen, H. and van de Ven, G.W.J. Finding niches for whole-farm design models - contradictio in terminis? *Agricultural Systems*, 87: 211-228, 2006.

Tallon, P.P., Kraemer, K.L., Gurbaxani, V. Executives' perceptions of the business value of information technology: a process-oriented approach. *Journal of Management Information Systems* 16(4), 145-173, 2000.

Torkzadeh, T. and Doll, W.J., The development of a tool for measuring the perceived impact of information technology on work. *Omega*, 27(3), 327-339, 1999.

Van Delden, H., Luja, P. and Engelen, G., Integration of multi-scale dynamic spatial models of socio-economic and physical processes for river basin management, *Environmental Modelling and Software* 22(2), 223-238, 2007.

Venkatesh, V., Speier, C. and Morris, M.G. User acceptance enablers in individual decision making about technology: toward an integrated model. *Decision Sciences* 33(2) 297-316, 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

DSS Success Measures: Evaluating the SCaRPA DSS

N.F. Herron[a] **and S.M. Cuddy**[b,a]

[a] *Integrated Catchment Assessment and Management Centre, The Fenner School of Environment and Society, The Australian National University, Canberra ACT 0200, Australia (natasha.herron@anu.edu.au)*

[b] *CSIRO Land and Water, Black Mountain Laboratories, Clunies-Ross St, Canberra ACT 2600 Australia (susan.cuddy@csiro.au)*

Abstract: The evaluation of a decision support system (DSS) should take account of the intended scope of the system, which has usually been defined as a result of consultation and planning between the developer and client. Project governance and decision-making roles need to be clearly assigned, performance targets and measures established, together with an analysis of external factors that could affect DSS adoption. We propose that different criteria should be used for evaluating environmental DSS, specific to its intrinsic and its extrinsic values. The intrinsic value relates to utility or how the DSS facilitates process, and includes elements such as design, content and process. It is relatively easy to plan for success, where success equates to adoption of the DSS. The extrinsic value of a DSS relates to its usefulness or impact, which for an environmental DSS means the extent to which environmental outcomes are achieved more efficiently than without the DSS. Evaluation against this measure is difficult for the project team. The authors use the approach to evaluate the SCaRPA DSS, an environmental DSS for catchment planning and on-ground investment assessment.

Keywords: decision support system (DSS), software evaluation, adaptive management

1. INTRODUCTION

What constitutes a successful decision support system? Given the time and resources that are thrown at their development across a large range of disciplines and the very high potential for 'failure' [Mysiak, 2005], the question is an important one. General consensus in the literature is that the mark of success for a DSS is adoption by the client or intended end-user. Adoption is an attractive measure of DSS success because it acts as an umbrella indicator for a range of 'success' criteria and, assuming that adoption is voluntary, gives a positive signal that the DSS meets the client's needs. Moreover, adoption of a DSS is relatively easy to measure – at the most basic level, a DSS is either adopted and continues to be used, or is abandoned. But is adoption a meaningful metric for evaluating a DSS?

Suppose the main purpose of a DSS is to improve the efficiency with which a limited pool of resources is channelled into management activities to maximise environmental outcomes. Success, then, will reflect the extent to which decisions made with the DSS lead to a better environmental outcome per dollar spent than decisions made without the DSS. However, measuring success using this criterion is significantly more complex. Issues relating to the models/assumptions informing the DSS, the data used to inform the models/decision steps, what the outputs from the models represent (e.g. are they indicators or actual measures of an environmental value), the time-scales over which the environmental benefits are assumed to occur, etc. confound the evaluation process. And a deficiency in any one of these is not necessarily a failure of the DSS, although it could be.

To evaluate whether use of the DSS is leading to better environmental outcomes, a monitoring and evaluation (M&E) program is needed to assess the extent to which the magnitude and direction of environmental change is predicted correctly by the DSS. But is this enough? Could not the same decision, and hence the same environmental outcome, have been made without the DSS? To properly evaluate the DSS, the M&E program needs to determine whether use of the DSS improved the likelihood of achieving the environmental objectives relative to not using the DSS, or putting it the other way, that a worse outcome would have been achieved had the DSS not been used to inform the decision-making process. Neither the biophysical assessment, nor determining the efficacy of methods, is an easy task.

But suppose a particular DSS can be shown to improve decision-making for environmental outcomes, would we still consider it a good or successful DSS if it is not adopted? The answer is probably not, even though the failure might have little to do with the DSS itself, and a whole lot to do with the environment into which the DSS is delivered. How do we deal with externalities, those factors which neither the development team, nor the end-user can control, but which impact on adoption, when evaluating a DSS?

In this paper, we consider some of these questions in relation to the SCaRPA (Site and Catchment Resource Planning and Assessment) DSS. What criteria do we use to evaluate an environmental DSS (EDSS) and should these be the same for every EDSS? For the SCaRPA, the satisfaction of our clients, the Catchment Management Authorities (CMAs) of New South Wales (NSW), Australia, is paramount, but the system should also satisfy the requirements of the independent auditors of CMA natural resource management planning and investment. We suggest that some measures of success are relatively easy to assess and design for, others are harder, and that there is an element of risk over which the DSS developer has no control.

2. ELEMENTS OF ENVIRONMENTAL DECISION SUPPORT SYSTEMS EVALUATION

In evaluating an environmental DSS (EDSS), it is useful to differentiate between what we have chosen to define as the *intrinsic* and *extrinsic* value of the system (Table 1).

Table 1. Definition of value elements of a EDSS.

Values	Description	Analogous to*
intrinsic	utility – the extent to which the DSS is user-friendly, logical in its construction, fit for purpose, adaptable, enhances knowledge and facilitates a decision-making process	Usability Functionality
extrinsic	the extent to which the EDSS contributes to achieving environmental objectives more efficiently than without the EDSS	Impact

* Analogous software terminology: [ISO, 1998]

2.1 Intrinsic Value

An EDSS can be conceived of as having three intrinsic elements:

- a *design* element, which relates to the technical implementation of the system; such as the software architecture, the efficiency and logic of the code, run times and the aesthetic and logic of the interface;
- a *content* element, which relates to the data and models needed to use the EDSS to inform a decision-making process; and
- a *process* element, which translates content into a form for decision-making.

How well each of these elements is handled within an EDSS determines its utility, and ultimately its adoption and continued use. The intrinsic value reflects such things as the models used to explore what-if scenarios; the supporting tools for analysing and visualising

results, generating reports and storing supporting information, the implementation of a formal process for the making of transparent, repeatable, defensible decisions; and/or its data management capabilities. EDSS developers have significant control over the intrinsic elements. Through close consultation with the client and detailed planning at the outset of development, they can put into play a plan to maximise the likelihood of adoption.

2.2 Extrinsic Value

Adoption and ongoing use of an EDSS suggests that end-users believe it adds value to the decision-making process, even if they cannot confirm that it is contributing to better environmental outcomes. We define a useful EDSS as one that contributes to making better decisions for the environment, in terms of more efficient progress towards an objective or set of objectives than without the EDSS. It reflects the extent to which the underlying science, as conceptualised in the environmental assessment models, and the data informing the models are an adequate representation of the physical reality. If the data are incorrect or the assumptions about process reflected in the models are wrong, then the predictions about environmental impact will be wrong. EDSS developers might have some influence over model and data selection, but equally they might not. This will depend on the scope of the EDSS, specifically the extent to which it is to be a fully-contained system with hardwired models and prescribed data sets, or has a more flexible framework into which models and data are added by the end-user.

2.3 Externalities

Finally, there can be barriers to adoption that do not relate to intrinsic or extrinsic elements, but which can doom a DSS to an immediate archive, irrespective of the utility and usefulness of the DSS. We are thinking specifically of the political, institutional and funding environment into which a DSS is delivered, but there are no doubt other external forces which can also have this effect. Of course, some have the opposite effect and result in promotion of the product and increased uptake.

3. A DSS FOR NATURAL RESOURCE MANAGEMENT IN NEW SOUTH WALES, AUSTRALIA

3.1 Natural Resource Management in New South Wales

In New South Wales, Australia, 13 Catchment Management Authorities (CMAs) have responsibility for managing the natural resources in NSW to maintain and improve environmental outcomes. CMAs are required by legislation to develop Catchment Action Plans (CAPs) that set out region-specific targets for improving natural resource condition and help prioritise their investment in reaching these plans. Development and implementation of each CAP must be consistent with the *Standard for Quality Natural Resource Management* (the Standard) and contribute to the state-wide targets for natural resources included in the NSW Government's State Plan. The Standard aims to promote consistent, high-quality practice across the natural resource management (NRM) sector, and encourage adaptive management to achieve goals and targets for resource condition improvement. To implement their plans, CMAs have been given AUD$436M specifically for investment in on-ground works (to be undertaken by willing landholders) that will contribute to improved environmental outcomes. CMAs need to adopt procedures for investment decision-making which demonstrate a sound basis and are likely to result in the intended outcomes.

3.2 CMA Decision Support System Needs

The TOOLS2 project, which commenced in 2005, was established with the express purpose of developing decision support tools to assist CMAs in catchment planning and

on-ground investment decision-making. CMAs were extensively consulted to elicit details of their environmental issues, their assessment tools, their business processes, the influence of non-biophysical factors (e.g. cultural, risk, capacity building, etc.) on funding determinations and what enhancements were needed to their existing methods to facilitate planning and investment processes. In this paper, we restrict our discussion to the latter, i.e. their site-level investment processes.

For incentives assessments, most CMAs reported that they have satisfactory procedures for assessing the environmental benefit per unit cost of an investment proposal in terms of soil, salinity, riparian zone and native vegetation outcomes, but they generally lacked capacity to assess aquatic habitat and biodiversity outcomes. Many CMAs were quite attached to their existing suite of assessment methods, and some indicated reluctance to change. Generally, CMAs said they would consider alternatives, if they could be convinced that the alternatives improved on their existing methods. Improvements were not construed simply in terms of better science or more sophisticated models. The challenge for CMAs is finding a workable balance between the demands for good science, robust models and adequate data collection and the day-to-day imperative of processing large numbers of funding applications with limited time and resources. They also stressed that the level of investment in collecting data and assessing proposals should not be disproportionate to the proposed funding amount.

CMAs were questioned about their methods for calculating environmental benefits, their business processes and funding models. We needed to know how they used biophysical model results, together with considerations about costs and social values, to inform the decision to fund. While their methods differed in the details, their funding models could be broadly categorised as ranked (tender-based) and threshold-based approaches. All CMAs took non-biophysical values, such as cost of proposal, cultural heritage, risk factors, community capacity building, etc., into account when making a decision to fund.

CMAs saw enormous value in a DSS that had corporate support, including linking into a centralised data, financial, contract and reporting management systems, and ongoing training and product maintenance.

The key messages from consultation were that CMAs:

- see themselves as autonomous;
- have different issues;
- have considerable knowledge and experience;
- want a DSS that is consistent with the existing policy environment;
- want a DSS that has corporate support and links to corporate systems;
- want good science;
- want the level of assessment burden to be commensurate with the size of the investment;
- need a transparent, defensible, repeatable, auditable process.

There were many other must-haves and desirables, but the foregoing were fundamental in shaping the design of the DSS. To the extent that these criteria are catered for in the DSS, they are part and parcel of a major success marker: adoption of the DSS by CMAs.

3.3 SCaRPA – Site and Catchment Resource Planning and Assessment – EDSS

SCaRPA comprises two core modules, a catchment planning (SCaRPA-cp) module and an incentives assessment (SCaRPA-ia) one. The EDSS is supported by a configuration tool, which allows a CMA to populate the system with region-specific data, such as resource condition and management targets, biodiversity benchmark data and other reference data. Technical details of SCaRPA are reported in Murray *et al*. [2008].

SCaRPA-cp was developed to assist CMAs in developing Catchment Action Plans and Implementation Strategies through the provision of catchment models and priority mapping tools, spatial multi-criteria analysis and scenario building tools. It has been designed so

that, provided certain protocols are followed, CMAs can 'plug in' catchment-scale assessment models of their choosing. It contains tools for registering catchment models and data layers, generating priority maps, visualising multiple priority layers, building scenarios and evaluating the environmental impacts of land cover/management changes using the registered catchment models. SCaRPA-ia was designed to assist CMAs make funding determinations about proposals for environmental outcomes. It is also intended as a planning and education tool, where CMA staff can work with landholders to explore alternative management options and develop property plans. This module allows CMAs to design funding programs, evaluate the environmental benefit of landholder proposals against criteria specified in a funding program and rank or rate proposals by their cost-benefit ratio, leading to a determination to fund or not fund.

4. DEFINING EVALUATION CRITERIA FOR SCaRPA

Firstly, the selection of success metrics to evaluate an EDSS must be appropriate to its intended scope and use, i.e. the metrics cannot be pre-determined and need to be identified at project inception. The umbrella marker of success adopted by the project team was that SCaRPA would be adopted by the majority of CMAs in NSW. Thus our aim was to build a system that would perform well in an evaluation of intrinsic value. From a CMA perspective, we considered the extent to which SCaRPA delivered on content and process elements vital, but so too the design elements, relating to qualities such as the attractiveness, usability, efficiency and stability of the software. These design qualities very much determine a user's level of satisfaction with a piece of software [ISO,1998].

In terms of content, we determined that SCaRPA would include models that are based on good science and would, with the 'right' data, give reasonable predictions of the impacts of different management interventions on a range of environmental variables. As the project team was responsible for the development of the EDSS framework and component assessment models, our evaluation of success must necessarily include how well individual models meet the criteria of good science and rapidity and ease of use, as required by the CMAs. However, when it came to data considerations, we resolved that we would not construe as failure poor DSS performance resulting from the use of poor quality data sets. Responsibility for the selection of data to inform models within SCaRPA was devolved to the CMAs.

An evaluation of SCaRPA's extrinsic value is by definition a post-project exercise and thus out of scope for the project team. This highlights a disjunct between the short-term nature of EDSS development projects (typically 1-3 years) and principles of adaptive management that assume adjustments to an EDSS's utility based on ongoing evaluation of the environmental impact over the longer-term. However, an evaluation of SCaRPA's extrinsic value is important in terms of the long-term adoption and support of SCaRPA.

5. PLANNING FOR SUCCESS

5.1 Acknowledge Existing Capacity

Even though the CMAs are relatively new organisations, many staff members have come from the NRM agencies that previously managed those regions. They have extensive knowledge and experience of their region, and they have a very real sense of what is practicable. To improve the likelihood of adoption, and at the same time acknowledge CMAs' existing capacity, we incorporated their current practices into the design, i.e. CMAs could continue to do assessments they way they currently do.

With 12 CMAs engaged in the project, this had significant design implications as the EDSS needed to cater for 12 different methods. For site-scale incentives assessments, this has been achieved via a template builder. The Template Builder is where the CMA designs their funding program – it steps them through a series of data entry screens which capture

details about the program and from which they select the assessment criteria for evaluating incentives proposals from landholders:

- *Overview of program* – source of funds; how much funds; aims of investment; links to catchment, resource condition or other targets; closing date; other details;
- *Environmental model selection* – what environmental models will be run; what weights to give to them when combining model outputs; details on the assessment within individual environmental models;
- *Method for combining model outputs* – different methods for combining model outputs into a single environmental index are offered;
- *Funding model selection* – tender-based or threshold-based options; cost-sharing;
- *Social and other criteria selection* – select from existing criteria or create new assessment criteria regarding the social, cultural and other value of the proposal (e.g. protects an indigenous sacred site; encourages innovation; contributes to community capacity building)

This template is then used to create the assessment form for each application received within that funding program. An application will only need to provide data or answers to the criteria nominated in the template, and all applications within the funding program are subject to the same rules. This approach addresses itself to the first three key messages from consultation (s3.2), as well as establishing a transparent, repeatable, auditable process. By allowing them to populate SCaRPA with CMA specific lists of management and resource condition targets, the framework facilitates reporting against targets, as required by the auditors.

5.2 Enhance Capacity

CMAs tended to lack capacity to assess aquatic habitat and biodiversity values. Developing models for catchment and site level assessments of aquatic habitat resource condition and response to management changes was a critical component of the project. Work also continued on developing existing models for land and soil capability, salinity, carbon and terrestrial biodiversity for incorporation into SCaRPA. As much as possible, the model developers have tried to cater for different levels of assessment, consistent with the CMAs' desire to keep the burden of assessment commensurate with the proposed level of funding. This means that for low priority outcomes or proposals where funds are capped at relatively low amounts can be assessed using simpler methods than the methods for assessing environmental outcomes from proposals involving large sums of money.

By catering for all CMAs, we have enhanced individual CMA capacity by offering a wider range of assessment criteria to choose from, and we have included the option of adding new criteria. We also provide options for combining model outputs into single environmental benefit indices (EBI), visualisation and reporting tools and a transparent, repeatable, defensible, auditable decision-making framework. SCaRPA builds CMA capacity in catchment planning and the linking of catchment planning to on-ground investment (Herron et al, 2008).

5.3 Improve Environmental Outcomes

Planning an EDSS to ensure that the decisions arrived at from using it will lead to better environmental outcomes is challenging. Model developers can make every effort to ensure that the component assessment models are based on good science. Where the physical processes and relationships are well understood and are backed up by considerable data, there is a reasonable probability that assessments made using these models will provide the 'right' results. If the underlying data are wrong, or the model is wrongly applied, then its predictions will be wrong. Provided that we have documented the assumptions of our models within the SCaRPA, we contend that poor results from SCaRPA due to inappropriate use by the CMA do not constitute a failure of the DSS.

However, without a long-term monitoring and evaluation (M&E) program, it is difficult to evaluate the performance ('success') of component models. Here M&E serves two purposes: first, monitoring the physical changes that come from a management intervention allows progress towards environmental targets to be tracked over time, which leads to second, an evaluation of the accuracy and reliability of the models that predicted particular outcomes. For the SCaRPA project, which runs for 3 years, this is entirely out of scope.

M&E can pose further difficulties. To be of value, an M&E program needs to differentiate between changes in the environmental value(s) caused by the management intervention from those caused by external controls. An example of this is where a comparison of habitat condition metrics at two points in time shows a deterioration in condition, despite the implementation of conservation activities (informed by an EDSS) to improve site condition. Has the EDSS contributed to a perverse outcome or have external factors, such as drought, fire or plague, overwhelmed the management impact? An evaluation done at a single point in time is unlikely to be sufficient, unless the evaluation criterion is area or length of change. Monitoring data needs to span a sufficiently long time that the direction and magnitude of improvement can be discerned above the 'noise' caused by climate or other drivers. Environmental response times to management intervention do not conform neatly to political and management time frames.

This discussion highlights the difficulty in evaluating the contribution of an EDSS's component models to effective decision-making and the attainment of resource condition targets. As a consequence, measurement and reporting is more often focussed on achievements against management targets (e.g. area of land re-vegetated), which SCaRPA does support.

5.4 Reality Bites

Linking the SCaRPA system to corporate data management, contract reporting and financial management systems, identified during the consultation process as a key desirable of the CMAs, became an important goal of the project. The ability to access data from centralised databases, maintained by state agencies, had proved invaluable in the corporate rollout of a closely-aligned tool for assessing applications to clear native vegetation and CMAs wished to have the SCaRPA plugged in too. While this was included in the implementation specification, the hurdles associated with porting an external product into a corporate network were too high to be overcome within the life of the project. As a result the current version is stand-alone, but designed with enough forethought that when it is eventually linked, all the information regarding funding programs, implemented works and data collected as part of site assessments can be uploaded to the central system for permanent upkeep. We would argue that this inability to deliver on a key success measure cannot constitute 'failure' as it proved to be beyond the control of the project team, and thus became out of scope.

5. CONCLUSIONS

Evaluating an EDSS requires consideration of a range of success criteria. These will not be the same for every EDSS. We contend that the developers of an EDSS need to define their success criteria at the outset of the development process, having regard to its intrinsic (i.e. its utility) and extrinsic (i.e. its impact) values within the scope of the EDSS. This means the developers need to consider the experience of the end-user, the interest other stakeholders might have in the use of the EDSS for environmental decision-making, and finally success criteria which are meaningful to them as software architects. As part of this preliminary planning, it is useful to identify what aspects of the system the developers can own and plan for in the system design, and what components of their system they might need to consider, but which they have little influence over. External factors, pertaining to the environment into which the EDSS will be launched and which could threaten adoption, should be anticipated, and to the extent possible, a plan should be developed to address the threat. This process contributes to setting success criteria and the scope of the EDSS.

In the SCaRPA project, via a process of active consultation with our CMA clients throughout the life of the project, we have designed a catchment planning and incentives assessment system that performs well in terms of its intrinsic elements. By assuming responsibility for the component models that we have developed as part of the project, we have assumed some responsibility for performance against extrinsic outcomes, but how well our models perform against real physical outcomes has yet to be assessed. The metrics for evaluating the extrinsic value of an EDSS are tightly coupled to those for assessing achievement against resource condition targets. How to measure outcomes (condition targets) as opposed to outputs (management targets) is a challenge facing all resource managers and researchers. Experimentation, implementation, M&E and adjustment are the adaptive management stepping stones towards this goal.

While much has been written about the success or failure of DSSs, it has not yet been harnessed into developing an evaluation framework. We consider this paper, a first step in that direction.

ACKNOWLEDGEMENTS

SCaRPA is a product of the TOOLS2 project, a collaborative R&D project with the NSW Departments of Environment and Climate Change, and Primary Industries, and CSIRO (Land and Water). The authors would like to acknowledge the contributions of their colleagues towards the development of the assessment models and the implementation of the framework. The project operates under an Activity Agreement for use in the devolution of NAP/NHT funds under a Commonwealth and State Bilateral Agreement.

REFERENCES

ISO, *ISO 9241-11: Guidance on usability*, 1998. Retrieved February 5, 2007: http://www.usabilitynet.org/tools/r_international.htm

Herron, N.F., Barrett, T.W. and Cuddy, S.M. Linking catchment environmental planning to on-ground investment: the SCaRPA DSS, this proceedings, 2008.

Murray, N.I., Cuddy, S.M. and Herron, N.F. *SCaRPA User and Reference Manual*, CSIRO Land and Water, Science Report 08/xx, 2008.

Mysiak, J. *Success and Failure of Decision Support Systems for Integrated Water Resource Management*, Venice, Italy, 6-7 October, 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

Raising the Bar – Is evaluating the outcomes of decision and information support tools a bridge too far?

K.B. Matthews[a], M. Rivington[a], K. Blackstock[b], K. Buchan[a] and D.G. Miller[a]
[a]*Integrated Land Use Systems Group, Macaulay Institute, Craigiebuckler, Aberdeen, AB158QH, U.K. k.matthews@macaulay.ac.uk*
[b]*Socio-Economics Research Group, Macaulay Institute, Craigiebuckler, Aberdeen, AB158QH, U.K.*

Abstract: This paper continues a series of reflections on the challenges of developing and deploying decision and information systems (DIST) for environmental management. Our focus is on the additional challenges being posed by funders to evaluate the outcomes (changes in the world beyond the research institute). This is a significant raising of the bar for DIST, particularly when many still struggle to overcome the simple hurdle of being used at all. The paper reflects on the challenge of evaluating outcomes, placing it in the context of conventional analysis of DIST performance (for example validation, software testing and utility assessment). Several particular challenges are identified, the intangibility of many outcomes, the difficulty of receiving credit for post project outcomes, how to disentangle cause and effect for changes that occur, how to decide the relative importance of outcomes and finally a failing to recognise that despite best endeavours research may still be ignored. The paper then presents a simple evaluation process conducted as part of a transdisciplinary research project that uses model based indicators to communicate the consequences of climate change to land managers and generate discussion of likely adaptations to management. The outcomes of the evaluations were that the process provided new information and raised awareness of the specific research being carried out. The project was also successful in changing views on climate change for a majority of attendees; particularly where the existing levels of knowledge were limited. Yet despite the relative success of the evaluation process and the useful lessons learned, designing an evaluation process and interpreting the findings remains a serious challenge. The authors conclude by questioning whether outcome evaluation is a vital requirement for successful DIST development or does it generate expectations that cannot be met?

Keywords: DIST, climate change, evaluation, outcomes, transdisciplinary

1. INTRODUCTION

This paper continues the bridging-the-gap metaphor used in previous IEMSS and MODSIM conferences where the focus was on the design-use gap and asks is the evaluation of outcomes a bridge too far?

The desire for the evaluation of outcomes (changes in the world beyond the walls of the research institute) rather than outputs (in the form of peer reviewed articles or software) is increasingly seen within policy-relevant and policy-led strategic and applied research programmes in the U.K. This is partially driven by the perceived failure of research and particularly decision and information support tools (DISTs) to deliver against their "utopian" promises of the 1990's [Matthews et al., 2007]. The drive for outcome evaluation also reflects funders seeking to assess the cost-effectiveness of their research to provide a justification for future spend as part of the new managerialist paradigm embraced by UK government [Richards and Smith, 2003]. Researchers are locked into this search for evidence through competition for limited research funds. Finally, there is a desire for research to clarify and rationalise the messy confusion of pluralist democratic decision making [Solesbury, 2001] i.e. to provide evidence for which decision should be taken.

This context raises questions about the role of researchers and our acquiescence to such agendas. There is, however, a more positive side to the desire to evaluate outcomes. Recent evaluation literature has

highlighted the importance of understanding the relationship between context, process and outcomes [Blackstock et al., 2007; Patton, 1998]. Therefore, a focus on outcomes requires understanding *how* and *under what conditions* information is interpreted and used by stakeholders – moving away from an 'information deficit' model to a model of knowledge exchange [Ekboir, 2003]. So as researchers committed to improving standards of environmental management, planning and policy - how can we make sensible steps towards evaluating outcomes of DISTs whilst managing the expectations of funders of research, decision makers and other stakeholders? This paper argues that the evaluation of the outcomes of using DISTs presents conceptual, resource and institutional challenges. Lessons from social research design can help DIST development teams deal with these challenges of evaluating outcomes.

The paper first reflects on some of the alternative interpretations of evaluation and how these may relate to the particular case of evaluating outcomes from DISTs. The paper presents a case study that evaluates the outcomes of a process communicating to land management stakeholders the consequences of climate change for land use using the outputs from agro-meteorological modelling. The paper concludes by trying to draw some general conclusions for the process of using DISTs.

2. RELATED RESEARCH

Several reasonable forms of evaluation coexist in the DIST literature: model validation, software testing and utility assessment. These can be usefully considered in terms of their appropriateness of their methods for outcome evaluation judged.

Model validation represents the assessment of the performance of a model against an independently collected dataset. Within an experimentalist and/or reductionist paradigm such systems of validation are effective. Validations of environmental systems models when conducted with stakeholders can also significantly enhance the credibility of the DISTs based on them [Carberry et al., 2002]. Validation based evaluations are weaker where key data cannot be falsified or directly measured (the preferences of stakeholders), the systems of interest are very large and/or very complex (issues of pre-analytical choices, scale and equifinality in parameterisation) or where experimentation for validation would raise practical and ethical concerns and is not permitted [Giampietro, 2004]. The validation based evaluation can also fail to address issues of legitimacy and salience. That is the model may perform well in a given context but be overly specialised, too demanding of input information or fail to address key relationships [McNie, 2007] that are of interest to outcome evaluation.

Software testing in this context refers to processes of software quality control – either formal or informal. These range from simple debugging to larger structured processes of software testing with hierarchical, recursive breakdown and testing of software components, units and modules [Britton and Doake, 1996]. These can be undertaken by developers (in smaller projects) or by testing and change management teams for larger systems with ongoing development. Such software testing can be assisted by automated testing to ensure repeatability of results or to benchmark systems [Hutchins et al., 2006]. It represents a potentially significant overhead in DIST development. This can be caricatured as a luxury or as non-productive use of staff time by research managers under pressure to deliver more with less. Model testing in the era of: rapid and open source development; the increasing availability of model-bases and model-engines and the adoption of interoperability standards, has to deal with higher level meta-data and meta-model compatibility issues [Ittersum, 2006]. These issues also concern the credibility of DIST outputs as the technical ability to combine elements does not always infer that there is a justification for doing so. Such evaluations also typically have a quality assurance of outputs focus rather than assessing the outcomes of DIST use.

Utility Assessment. Both of the cases above have a credibility focus, which is necessary but not sufficient for the evaluation of DISTs. There remains for many DISTs the thorny issue of problem of implementation, or why does my perfectly implemented DIST not get used by anyone [McCown, 2002]? Utility testing has been extensively debated elsewhere [Matthews et al., 2006b] with the conclusion that issues of salience and legitimacy can undermine the utility of DIST's. Firstly there must exist an adequate and honest partnership between developers of DISTs and relevant stakeholders for developing the specification of DIST functionality but more importantly agreement on the process through with a DIST is used for example as part of deliberative inclusive processes. Processes of utility testing may be experimental testing [Matthews et al., 2006a], formal market research survey [Matthews et al., 2007] or based on soft-systems methods of participatory co-development [Matthews et al., 2006c] depending on

the particular circumstances and the resources available. Whatever the particularities utility testing is essential if the DIST is not to become shelfware.

Outcome Evaluation, however, needs to go beyond the utility, validity or structural integrity of a DIST and into assessment of the effects of using that tool within a decision making or management process. Evaluation of outcomes from the strategic rather than routine use of DISTs confronts several taxing issues. First is the intangibility of many outcomes that leads to just measuring what can be measured and calling that success when other outcomes that are more difficult or costly to measure are still very important. For example the value of the social capital bonds between key individuals in a community strengthened through by participation in a DIST development may enhance their adaptive capacity far more than the DIST itself. Secondly, it is also very difficult to attribute any outcomes in complex social systems to any particular interventions [Bellamy et al., 2001]. It is difficult to establish causality in real world research when processes and participants can not be directly replicated or controlled [Robson, 1993]. Attribution of belief that a change is stimulated by an intervention *is* possible, and this can be easier in newer fields where there are fewer interventions to complicate these attributions. Even where outcomes can be distinguished then is any success simply happenstance rather than predicated on the actions of the research team? Thirdly, the long term and cumulative nature of change may also be incompatible with short-term, project-based research and development funding schemes. The project team may have moved on to new research long before the outcomes can be measured [Blackstock et al., 2007]. Funders are also less keen to fund monitoring and evaluation of existing tools rather than being seen as the initiator of a new development. Fourthly even where outcomes can be distinguished in some way then there will be considerable disagreement on the relative importance of individual outcomes, particularly when these outcomes are non-commensurable. In such cases there is too often the resort to crude, *ad hoc*, indicator sets that simply pass on to a third party the difficulty of interpreting of intimately interconnected and highly nuanced systems-level outcomes. And the politics of knowledge suggest that outcomes other than those expected (or even desired) by the DIST funder will be discounted in favour of those that 'fit' [Kirk et al., 2007]. Finally there needs to be a recognition of the limits on science generated "expert" evidence within a plurality of expertise derived in different ways [Stilgoe et al., 2006]. Failure to recognise the legitimacy of alternative points of view and engage with deliberative processes can mean that otherwise sound science being ignored as an arrogant attempt at scientism. Assessment of the outcomes of research-based tools also need to recognise that research and research based tools such as DISTs are only one of the sources of influence and probably not one of the more important [Solesbury, 2001]. This is not to discourage DIST developers from contributing to policy and management decision making but expectations must be managed if we are to avoid a further "disappointment" crash as happened with DSS in the late 1990's.

For the authors there is no preconceived preference for evaluation methods since both quantitative and qualitative approaches can be equally rigorous or specious depending on implementation. Instead, our research team has tried to chart a pragmatic path to evaluation that addresses each of the issues of credibility, legitimacy and salience. Such a mixed-methods based approach to evaluation is presented in the following sections.

3. MATERIALS AND METHODS

3.1. Climate data and modelling

The underpinning research for the C4LU project has its origins in our research interest in the consequences of using alternative sources of data meteorological data in the cropping systems components of the Land Allocations Decision Support System (LADSS). Testing in this situation was conventional experimental analysis looking at the impacts of using off-site or modelled on-site data when measurements for particular variables were not available. The is process quantified the uncertainties introduced and the reliability of the DSS results using expert weighted performance metrics combined in a fuzzy logic framework [Rivington et al., 2005; Rivington et al., 2006]

Interest in using for climate change scenario data from UKCIP02 led to a desire to test the effectiveness of the Hadley Centre Regional Climate Model (HadRM3) in replicating historic conditions. Testing these hindcasts allowed the development of simple downscaling methods to correct for representational and other biases within the HadRM3 data for particular sites to increase the confidence in the data for use at specific case study locations. These monthly correction factors were also applied to the future scenarios on the assumption that the representational bias would still be present and other biases were likely to

remain. Here the testing was conventional statistical testing with the outcomes published in peer reviewed journals [Rivington et al., 2007].

3.2. Agro-meteorological metrics

In parallel with the climate data and modelling a project developing agro-meteorological metrics was undertaken with selected stakeholders. This project has been reported elsewhere [Matthews et al., 2008] but in this context the key features were that it tested the utility of existing agro-meteorological metrics as a basis for first for characterising future climate change scenarios and secondarily for stimulating deliberation on possible adaptive strategies that could be adopted by land managers.

The indicator set, the presentation methods and the format of the interactions with stakeholders were all piloted with leaders of the stakeholder organisations and then tested and refined (to a smaller set of key indicators) with stakeholder groups (total participants = 40). A small workshop format (8-16 participants lasting up to 3 hours) has been used with round-table presentation of outcomes from the modelling, deliberation on the significance of the results and the utility of the indicators combined with qualitative information gathering on adaptive strategies and a wider discussion or debate of related issues. Each workshop has thrown up additional refinements to either the process or the elements presented and have seen increasing demands for more sophisticated analyses. The latter result is particularly important since it indicated the research team are gaining skill in communicating the necessarily complex messages and when the means of presentation have been improved stakeholders seek more information.

3.3. Evaluating the outcomes

The team were asked to evaluate the outcomes of the most recent round of (4) workshops in February and March 2008 by the research funders. These are not a usual research funder but a science communication group within the Scottish Government tasked with increasing the interactions between active researchers and the wider publics. Their interests and those of the research team coincided in the communication of climate change consequences but their evaluation of success could not be in terms of peer reviewed papers. We were also anxious not to fall into the trap of judging success by the easily measured outputs, e.g. numbers of people attending. This would compromise our participatory research approaches by simply increasing the number of people at the expense of the quality of interactions, whilst not necessarily improving outcomes (increased awareness and understandings).

The attempt was made to evaluate the outcomes of the workshops at two levels. The first was adding a simple evaluation form to the materials used within the workshop. The second was to record the meetings to allow a subsequent analysis of how participants responded to the information in discussion. The former is presented within this paper while the latter has still to be undertaken (reflecting the balance of effort required for the two approaches).

The survey analysis asked five simple questions with tick box answers (options shown in brackets).

- Q1. Did you know about the Macaulay Institute's research before you attended the workshop? (Y/N**)**
- Q2. How much did you know about climate change before you attended the workshop? (Options)
- Q3. Has the workshop provided new information on the topic? (Y/N)
- Q4. Have you changed or adapted your views on climate change after the workshop? (Y/N)
- Q5. Where do you work? (Academic Institution, Government, NGO, Private Sector, Other)

These questions sought to assess how effective the deliberative workshop process is 1) to raise awareness, 2) to provide new information and 3) to influence attitudes. The research team did not attempt to measure changes in participants' practices, given that literature on barriers to changing farmers' behaviour [Burton et al., 2007; Kaljonen, 2006] suggests it would be impossible to rigorously assess the degree to which the DIST alone was responsible for change (or the lack there of). The results of the four workshops are presented below but represent just a first step into this issue for a DIST research team.

4. RESULTS AND DISCUSSION

4.1. Utility Analyses

The overall responses on the utility of the 15 indicators are shown in Figure 1. The previous processes have been successful in largely eliminating indicators that are not seen as relevant and the top two categories (Very and Quite) account for 58% of the responses. A middle category was deliberately not

provided to avoid "neutral responses" but some stakeholders felt the four categories were insufficiently fine grained. An enterprise specific category that allowed delegates to highlight indicators that were important but only for a narrower set of circumstances or activities was suggested.

The breakdown for each of the metrics is summarised in Figure 2, with the counts of the Very and Quite useful responses shown. This highlights the importance of growing season and access periods, but particularly the increasing view that metrics showing the monthly distributions of phenomena are useful (growing days, dry soil days, access days or their combinations). The continuing importance of climate metrics in the survey does not match well with the comments within the later discussions, suggesting the higher score for these indicators may be the result of their appearance first in the workshops (the delegates having nothing to compare the indicators with at that stage). This possible artefact highlights the need for extreme care in designing and interpreting evaluation outcomes.

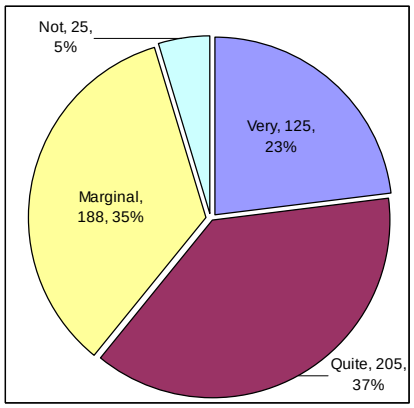


Figure 1 - Utility of the C4LU Indicators

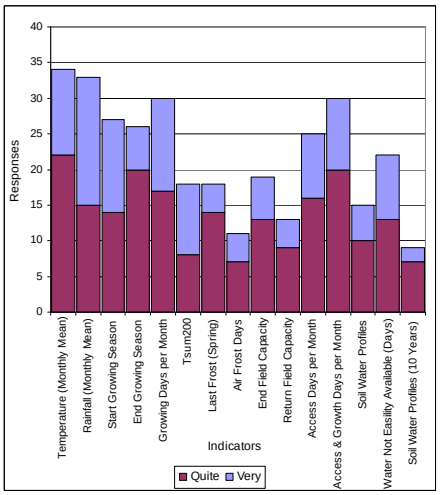


Figure 2 - Responses per Indicator

Figure 2 also shows the preference for the compound indicator access during the growing season. This indicator and the way in which it is presented was the outcome of earlier rounds of workshops. The benefits of processes that can encompass aspects of learning are thus evident. However, such flexibility and process evolution can conflict with the formalisms of conventional science where replication is required. Such methodological approaches may inadvertently constrain the ability to achieve the outcomes sought (increased awareness, knowledge and adaptation) by stifling adaptive participatory processes.

4.2. Outcome Evaluations

The responses from the C4LU workshops are summarised in Figure 3. These show that the participants were predominantly from the private sector with smaller numbers from the government, NGO and academic sectors. Of the private sector participants all were actively involved in land management decision making either in farming, forestry, fishing/hunting or in related activities such as tourism and the agri-food supply chain. In many cases the participants ran diversified businesses and thus had a broad view of potential climate change impacts. There was a significant representation of agri-environmental interests so the discourse was not dominated by productivism, though food security issues were frequently raised.

In terms of awareness of the Institute's research the workshops were marginally effective (5/40), but this reflects the Institute's previous extension role. The participants are aware that research is carried out but the topics of research (even at a headline level such as climate change) are poorly understood or are those of earlier programmes. The workshops thus represent a significant opportunity to do awareness raising for particular projects and the wider programmes of research, and this is vital if there is to be continued support for public-good funded research.

In terms of expertise all participants had at least a little knowledge (perhaps reflecting the efforts in the mass media and trade journals to inform on the issues of climate change). There were more participants classifying themselves as having a fair knowledge (and the lack of any classifying themselves as experts

may be due to modesty – considering how well informed many of the participants were). There was a fair degree of variation in the level of reported knowledge – with the Golspie case study (Figure 4) reporting lower levels of knowledge compared with for example Aviemore (Figure 5). This perhaps reflects the backgrounds and interests of the land managers in Golspie where conventional mixed-farming systems dominate rather than the Aviemore land managers who were strongly diversified enterprises and with a strong agri-environmental representation.

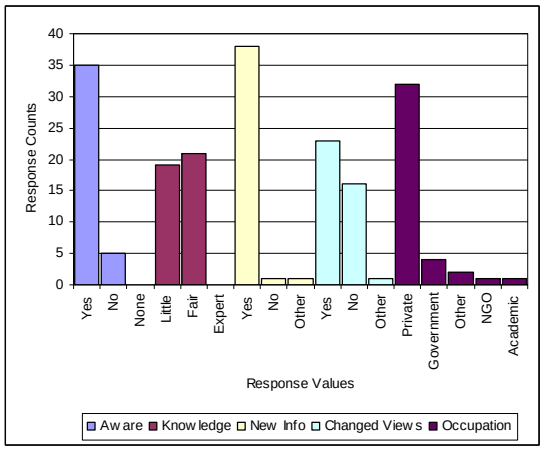


Figure 3 – Overall C4LU outcomes

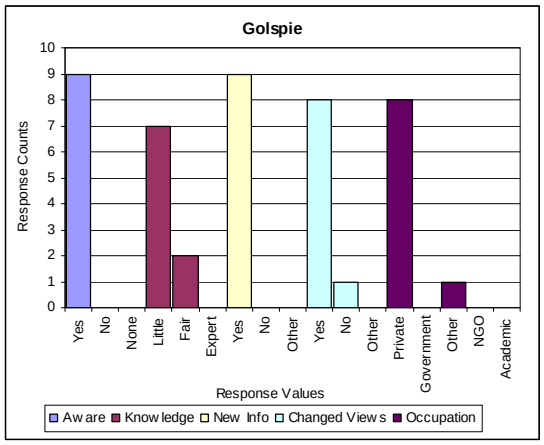


Figure 4 - Golspie C4LU responses

In almost all cases (38/40) the workshops were successful in providing new information. The two cases where this was not the case were government and academic colleagues who were participating in the workshops as participants but also providing expert interpretations of the implications of the metrics for particular sectors (e.g. forestry). The new information evaluation is a key baseline for the evaluation since this is the conventional role for scientific research – failure to be providing new information would raise serious questions of whether the research was relevant to the decision makers at all. The usefulness of the new information is assessed as outlined in Section 4.1 (above).

The acid test for the workshops was in changing or adapting views on climate change. In a majority of cases (23/40) the participants had changed their views on climate change. (Further analysis of the discussion may also indicate if those who have not changed their view already felt adapting to climate change was important). This was interesting as previous research had found that climate change was not seen by land managers as a priority for action or one for which they had much adaptive capacity [Scottish Executive, 2006]. The message of the workshops was that mitigation efforts will be important but that change in climate will have to be planned for and adapted to since a certain degree of change has already been entrained by previous emissions. The authors think this message was successfully delivered through the collective evaluation of the indicators than conventional research presentations.

The survey, however, only sought to evaluate if peoples views had changed, not *how* they had changed. This reflected a desire to keep the form as simple as possible and not to alienate participants by use of a complex questionnaire. The questionnaire is backed up by audio-recording of the workshop, which can be qualitatively analysed to bring out these more nuanced questions. The multi-method approach avoided a more explicit focus on changes in attitude within the workshop, which could be quite confronting in a group based process.

By combining the utility and evaluation datasets the authors attempted to develop a "profile" of the participants to explore if there was a strong link between the reported utility of the indictors and whether workshop participants has altered their views. Figure 6 shows each participant's judgement of the utility of each indicator, where the participants are sorted into those who did or did not change their views. There is no clear difference between the two profiles of the change and no-change groups. This suggests that individual actors have a complex set of values, interests and existing knowledge that strongly modify their responses to any research based intervention. It is interesting to note that the case-study with the greatest rate of change in view was Golspie (Figure 4) where the level of knowledge was lowest. This perhaps reflects the previous view (Section 2) that it is easiest to make the biggest impact in areas (in this case geographic but also by extension for subjects) where the knowledge base is less well developed and where perhaps actors have less entrenched views. The inference from this finding is perhaps that as researchers we need to be engaging with key actors and decision makers as early as possible rather than

trying to reshape a debate that has already "solidified" such that research outcomes are ignored rather than incorporated. Conversely, the failure of a DIST to change practices may be due to entrenched views not a poor tool.

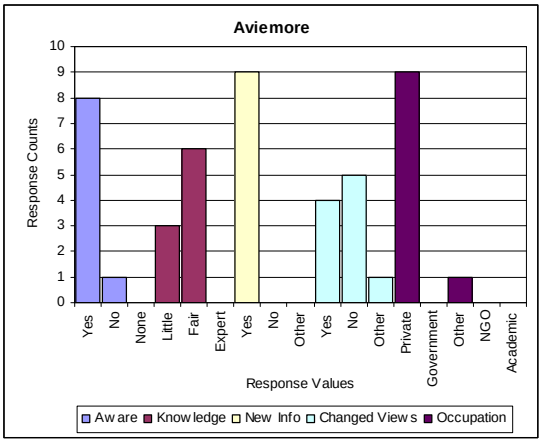


Figure 5 - Aviemore C4LU responses

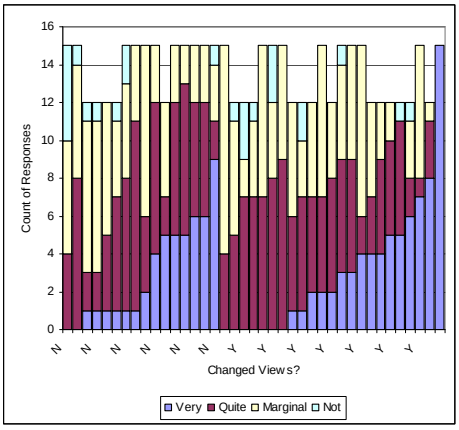


Figure 6 - Profile of utility of indicators responses

5. CONCLUSIONS

In this paper the authors have identified a further challenge for DIST developers: the evaluation of the outcomes of the use of their tools. The authors have reflected on the justification for such requirements and on the limited capacity of most DIST research team to deliver rigorous evaluations rather than easily measured yet largely meaningless output metrics. The authors have presented an example of evaluation used as part of a transdisciplinary research project communicating the climate change consequence for land managers and seeking to simulate deliberation on possible adaptive responses.

While it has been possible to generate simple but informative headline evaluations of the C4LU process outcomes, it is clear that the design and interpretation of such evaluation processes requires great care. These data open up discussions about potential changes in attitudes but illustrate the outcomes are hard to assess without exploring the participant's individual context and the influence of the deliberative process. While the research team developing the C4LU project are experienced in both model development and managing social learning processes, the additional requirements for evaluating the outcomes of the project still suggests this is a gap that may be too broad to bridge. Transdisciplinary research is already demanding of skills and resources within constrained budgets, yet this further demand is often introduced without recognition of the additional resource implications.

The evaluation of DIST outcomes is a challenge that needs to be met since evaluation and reflection in effect closes the loop that starts from initiation, and continues through design, build, and testing. How best to organise the evaluation of DIST? Should the teams developing them do the evaluation, since they are most familiar with the tools and can most quickly respond to feedback with improvements? Should the DISTs be evaluated by external teams with the skills to undertake the task? Is there the need for partnership, but if so how do we resolve the power relationships between a technologically skilled development team and evaluators with a different skill set and priorities? In part this requires engaging funders in a discussion about the meaning of evaluation metrics and debating what are the outcomes that DISTs can reasonably be expected to achieve. In the authors view a considered evaluation of the outcomes of DISTs could be the key to realising their potential in environmental management. Yet outcome evaluation may prove a bride too far if we fail to control inflated or utopian expectations of science based tools and interventions?

REFERENCES

Bellamy, J. A., Walker, D. H., McDonald, G. T., and Syme, G. J., A systems approach to the evaluation of natural resource management initiatives, *Journal of Environmental Management* 63(4), 407-423, 2001.

Blackstock, K. L., Kelly, G. J., and Horsey, B. L., Developing and applying a framework to evaluate participatory research for sustainability, *Ecological Economics* 60(4), 726-742, 2007.

Britton, C. and Doake, J., *Software System Development, A Gentle Introduction*, McGraw-Hill Publishing Company, London, 1996.

Burton, R. J. F., Dwyer, J., Blackstock, K. L., Ingram, J., Brown, K. M., Mills, J., Schwarz, G., Matthews, K. B., and Slee, B., *Influencing positive environmental behaviour among farmers and landowners - a literature review for DEFRA*, Report, Macaulay Institute, 2007.

Carberry, P. S., Hochman, Z., McCown, R. L., Dalgliesh, N. P., Foale, M. A., Poulton, J. N. G. H., Hargreaves, D. M. G., Cawthray, S., and Hillcoat, N., The FARMSCAPE approach to decision support: farmers', advisers, researchers' monitoring, simulation, communication and performance evaluation, *Agricultural Systems* 74, 141-177, 2002.

Ekboir, J., Why impact analysis should not be used for research evaluation and what the alternatives are, *Agricultural Systems* 78, 166-184, 2003.

Giampietro, M., *Multi-scale integrated analysis of agroecosystems*, CRC Press, Boca Raton, Florida., 2004.

Hutchins, M. G., Urama, K., Penning, E., Icke, J., Dilks, C., Bakken, T., Perrin, C., Saloranta, T., Candela, L., and Kamari, J., The model evaluation tool: guidance for applying benchmark criteria for models to be used in river basin management., *Archiv fur Hydrobiologie - Large Rivers Supplement* 2006.

Ittersum, M. K. v., Integrated assessment of agriculture and environmental policies: towards a computerised framework for the EU (SEAMLESS-IF), In: Voinov, A, (ed.), *3rd Biennial Meeting - Summit on Environmental Modelling and Software*, 2006.

Kaljonen, M., Co-construction of agency and environmental management. The case of agri-environmental policy implementation at Finnish farms, *Journal of Rural Studies* 22(2), 205-216, 2006.

Kirk, E. A., Reeves, A. D., and Blackstock, K. L., Understanding environmental regulation: disciplinary interactions and path dependency, *Environment and Planning C*. 25, 250-268, 2007.

Matthews, K. B., Buchan, K., Sibbald, A. R., and Craw, S., Combining deliberative and computer-based methods for multi-objective land-use planning, *Agricultural Systems* 87, 18-37, 2006a.

Matthews, K. B., Hutchins, M., and Hill, G., Bridging the design-use gap for DSS in environmental policy and practice., In: Voinov, A, (ed.), *Summit on Environmental Modelling and Software - 3rd Biennial meeting of the International Environmental Modelling and Software Society (iEMSs)*, 2006b.

Matthews, K. B., Rivington, M., Buchan, K., Miller, D. G., and Bellocchi, G., Characterising and communicating the agro-meteorological implications of climate change scenarios to land management stakeholders., *Climate Research* 2008.

Matthews, K. B., Schwarz, G., Buchan, K., Rivington, M., and Miller, D. G., Wither agricultural DSS?, *Computers and Electronics in Agriculture* 61, 149-159, 2007.

Matthews, K. B., Wright, I. A., Buchan, K., and Davies, D. A., Assessing the options for upland livestock systems under CAP reform: developing and applying a livestock systems model within whole-farm systems analysis, *Agricultural Systems* 90, 32-61, 2006c.

McCown, R. L., Changing systems for supporting farmers' decisions: problems, paradigms and prospects, *Agricultural Systems* 74(1), 179-220, 2002.

McNie, E. C., Reconciling the supply of scientific information with user demands: an analysis of the problem and a review of the literature., *Environmental Science and Policy* 10, 17-38, 2007.

Patton, M. Q., *Utilization-focussed evaluation*, Sage, Thousand Oaks, CA, 1998.

Richards, D. and Smith, M. J., *Governance and Public Policy in the UK*, Oxford University Press, 312 pp., Oxford, 2003.

Rivington, M., Bellocchi, G., Matthews, K. B., and Buchan, K., Evaluation of three model estimations of solar radiation at 24 UK stations, *Agricultural and Forest Meteorology* 132, 228-243, 2005.

Rivington, M., Matthews, K. B., and Bellocchi, G., Evaluating uncertainty introduced to process-based simulation model estimates by alternative sources of meteorological data, *Agricultural Systems* 88, 451-471, 2006.

Rivington, M., Miller, D. G., Matthews, K. B., Russell, G., and Buchan, K., Evaluating Regional Climate Model estimates against site-specific observed data in the UK., *Climatic Change* 2007.

Robson, C., *Real World Research: A Resource for Social Scientists and Practitioner-researchers*, Blackwell Publishers, Oxford, 1993.

Scottish Executive, *A Forward Strategy for Scottish Agriculture: Next Steps*, Report, Natural Scotland, 32pp., 2006.

Solesbury, W., *Evidence Based policy: Whence it came and where it's going*, Report, UK Centre for Evidence Based Policy and Practice, 2001.

Stilgoe, J., Irwin, A., and Jones, K., *The received wisdom. Opening up expert advice*, DEMOS, London, 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

User driven application and adaptation of an existing Policy Support System to a new region

H. van Delden[a], M. Ouessar[b], M. Sghaier[b], A. Ouled Belgacem[b], M. Mulligan[c], P. Luja[b], J. de Jong[a,d], M. Fetoui[b] and M. Ben Zaied[b]

[a] *Research Institute for Knowledge Systems, P.O. Box 463, 6200 AL Maastricht, The Netherlands (hvdelden@riks.nl, pluja@riks.nl, jdejong@riks.nl)*
[b] *Institut des Régions Arides, El Fjé 4119 Médenine, Tunisia (ouessar.mohamed@ira.rnrt.tn, s.mongi@ira.rnrt.tn, azaiez.ouledbelgacem@ira.rnrt.tn ,mondher_ga@yahoo.fr, benzaied_m@yahoo.fr)*
[c] *Environmental Monitoring and Modelling Research Group, Department of Geography, King's College London, Strand, London, WC2R 2LS (mark.mulligan@kcl.ac.uk)*
[d] *Centre for Geo-Information (CGI), Wageningen-UR, P.O. Box 47, NL-6700 AA Wageningen, The Netherlands (job.dejong@wur.nl)*

Abstract: The development of a decision or policy support system is a very time consuming and expensive process. Reusing an existing system provides therefore major benefits. This paper describes the process followed when applying the existing MedAction Policy Support System, developed for Northern Mediterranean regions, to the Oum Zessar watershed in south east Tunisia. The paper presents the differences and similarities of the problems and processes in the different regions and describes how they relate to the adaptation of the existing model and the place the system could take in the decision-making process. Furthermore, it discusses the likelihood of its actual uptake in the Tunisian context based on six criteria discussed in a participatory approach with potential users of the system.

Keywords: Policy Support System (PSS); User requirements; Linking science to policy-making; Desertification; Integrated spatial planning.

1. INTRODUCTION

In the past decades a large number of systems to support decision making in regional planning have been developed. Some for a very specific task in a specific region, others as generic tools that can be applied to a wide range of topics at different locations. Since the development process of a decision or policy support system is normally a very expensive and time consuming process, reusing the system for a broad range of problems and reapplying it to different regions after its development would give great benefits. For most systems, however, it is not clear to what extent they can be reused and what the required effort is to apply them to new regions and embed them in new organisations.
This paper will discuss to what extent the current version of the MedAction Policy Support System (PSS), developed for the European context, can be applied to Northern African countries using the watershed of Oum Zessar in south east Tunisia as a first test case.

For the research presented, the following questions have been formulated:

1. To what extent are regional development problems in Northern Africa similar to those in Southern Europe?
2. To what extent are socio-economic and bio-physical processes in Northern Africa similar to those in Southern Europe?
3. To what extent is the institutional context in Northern Africa similar to the one in Southern Europe?
4. What needs to be adapted to the system to make it applicable for regions in Northern Africa?

5. Based on this case study, can (a system like) the MedAction PSS provide support to policy-making in Northern Africa?

To answer the fifth question, this research pays specific attention to the six elements that help to understand the failure or success of a PSS in practice (based on Van Delden et al, 2007):

A. Strategic value of the system: to what extent does the system provide an added value to the current planning practice? How does it provide support to mitigate main problems in the region?
B. Availability of appropriate data and models: what is available at present or can easily be collected?
C. Credibility of the system: do the users have faith in underlying assumptions?
D. Language of the system: does it connect to the world of the end-users?
E. Institutional embedment: where will the system be based in the organisation? Who is actually going to use the system? Where in the policy process is it most beneficial?
F. Culture: are people willing to adopt and use the system? Is there commitment to give the system a place in the planning process?

To provide the context of the research, we provide a brief overview of the MedAction PSS in section two and a description of the region and its characteristics in section three. The paper will provide answers to the research questions posed, following the methodology described in section four. In section five, we present the outcomes of the interaction with local stakeholders and scientists. Section six aims to provide some preliminary answers to the six questions that help to understand the potential for actual use of the MedAction PSS in Northern Africa and the Oum Zessar watershed in Tunisia in particular. The last section discusses the conclusions and provides answers to the five research questions formulated above.

2. THE MEDACTION POLICY SUPPORT SYSTEM

The MedAction PSS has been developed as a generic system for Northern Mediterranean regions to support integrated decision making in the area of sustainable farming, water resources, land degradation and desertification. It allows the user to explore the impact of a wide range of external factors and policy options on policy-relevant indicators by simulating future developments in the region over a time span of 20 to 30 years.

Its goal is to assist policy makers in (1) understanding the important processes in the region and their interaction, (2) indicate current or future problems in the region or river basin, (3) assess the impact of possible policy measures to mitigate the problems, (4) evaluate the different alternatives and (5) stimulate discussion and improve communication between the different actors involved in the decision-making process.

The system comprises of a wide range of spatial dynamic sub-modules, integrated into a single computerised system (see Figure 1 for main components and relations). The climate and weather module produces rainfall, temperature and radiation figures based on historic datasets and IPCC climate scenarios. These are used for calculating run-off and evapotranspiration in the hydrology module and the growth of plants in the plant growth module. The hydrology and plant growth module are strongly interconnected for the calculation of evapotranspiration, biomass and soil moisture. Incorporated on the bio-physical side are also modules to calculate the erosion, sedimentation and salinisation of soils. These modules determine the suitability of the soil for different types of crops and natural vegetation and also produce several desertification indicators.
Economic and demographic growth is translated into the land use functions: residential, agriculture, industry and tourism. These functions are in competition with each other in the land use module. This competition is simulated with a constrained cellular automata model that allocates the demographic and economic activities on the land use map. The land use class natural vegetation takes in space that is not occupied by other functions or that is restricted for the other functions by zoning plans. The land use classes agriculture and natural vegetation are further specified through respectively the modules farmer's decisions

and natural vegetation which produce the crop types and the natural vegetation types that will cover the locations for agriculture and natural vegetation. The growth of the plants (crops as well as natural vegetation types) is subsequently calculated in the plant growth module mentioned before. The last module in the PSS is the water resources module. This module calculates the amount of available water for irrigation and drinking water, and determines the water used by the different functions based on the maximum available amount and the price of water from different sources (groundwater, reservoirs, and desalinised water from the sea).

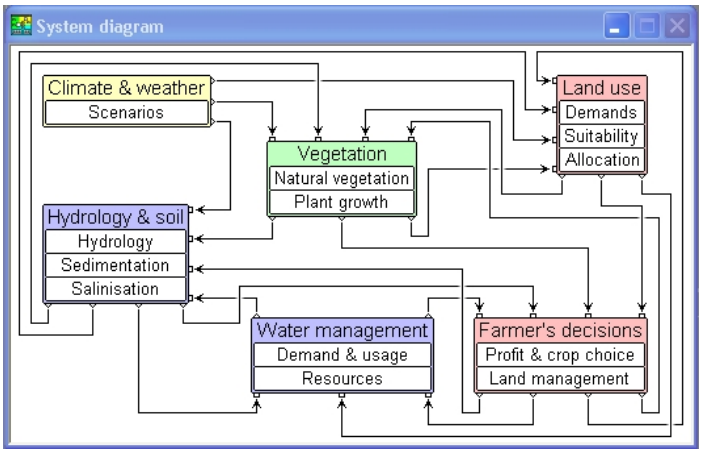


Figure 1. System diagram of the MedAction Policy Support System.

Each process is modelled at its appropriate geographical and temporal resolution. The finest spatial resolution is 100 by 100 meters; the land use module calculates on a yearly basis while the hydrology module has a variable time step in the order of minutes determined by the intensity and duration of the showers. To enable dynamic feedback loops between the different processes modelled, the MedAction PSS is developed with the **GEONAMICA®** software environment. This is an object-oriented application framework for building decision support systems based on spatial modelling and (geo)simulation (Hurkens et al, 2008). More information about the MedAction PSS can be found in Van Delden et al (2007).

3. THE OUM ZESSAR WATERSHED

The Oum Zessar watershed is chosen because this region can be considered representative of the arid southeastern Tunisia from an ecological, hydrological as well as a socio-economical point of view (Chahbani, 1984; De Graaff and Ouessar, 2002). The watershed covers a territory of 10 imadas (lowest administrative unit in Tunisia) belonging to three counties (Béni Khédachi, Médenine North and Sidi Makhlouf). The total population of the watershed is estimated, according to the population census of 1994, to 24188 inhabitants. The household number is 5758 with an average family size of 5.5. The area encompasses the mountains of Matmata (Béni Khédache) in the south-west, the Jeffara plain (around Koutine) and the saline depression (Sebkha) of Oum Zessar and ends in the Mediterranean sea (Gulf of Gabès).

Rangelands are the dominant land use in the study area. The vegetation is mostly steppe but the species composition is highly variable depending on relief and soil type. Crop cultivation takes up an important part of the land area. The climate of the pre Saharan Tunisia is characterized by hot and dry summers that last from May to August. Rainfall is very irregular with measured annual totals varying between 14 mm (min) and 590 mm (max) to average around 170 mm (Koutine station). This results in highly variable yearly yields and accompanying risks for the local population. To deal with these risks many households depend on several sources of income. Nonetheless, poverty remains a crucial problem.

4. METHODOLOGY

The process followed while applying the system to the new context and user community very much resembles that of the development process of integrated spatial decision support systems (ISDSS) – and Policy Support Systems – as described by Hurkens et al. (2008). This development process can best be described as an iterative process of communication and social learning amongst three involved parties – as shown in figure 2. First, there are *policy makers and analysts*, the users of the system. They provide the policy context and define the problems, functions and usage of an ISDSS. Second, there are *scientists* responsible for the main model processes and choices of scale, resolution and level of detail. Third, there are *IT-specialists* who design the system architecture and carry out the software implementation of the models and user interface. The interaction between the three groups involved is as important for the quality of the final product as the tasks carried out by each group individually and helps each group to gain a better understanding of the needs the others have and the possibilities they offer. For this understanding to take full effect, an iterative approach is best suited.

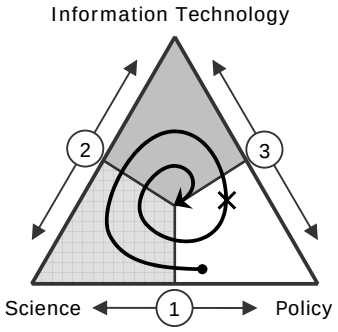


Figure 2. ISDSS development process

Where the original development process starts at the black dot in the figure, the application process described in this paper starts after a first iteration loop at the black cross. At this point a prototype is already available and discussions during a first interaction focus around the question if an ISDSS or PSS would provide a contribution to the decision making process and if so, if the system presented would be a good starting point.

To stimulate interaction and discussion between the different groups, we selected workshops as the appropriate format to answer the research questions. During the workshop sessions we discussed open questions and tried to reach consensus about the problems and processes in the region as well as the important elements for adaptation and application of the system. In these discussions local stakeholders contributed to the policy section of figure 2, local scientists to both the science and the policy section and the team that originally developed the MedAction PSS and applied it to Southern European regions, contributed to the science and IT sections.
Throughout the process both the local stakeholders and scientists were seen as (potential) users.

In the application process we take the following steps:

1. First meeting between developers and local scientists with the aim to:
   1.1. Understand the main problems of the region
   1.2. Present and discuss the existing Policy Support System
2. Application of the existing Policy Support System to the new region
3. Second meeting between developers and local scientists with the aim to:
   3.1. Understand the decision-making process in the region
   3.2. Become aware of data limitations
   3.3. Understand important processes in the region
   3.4. Discuss the extent to which the existing model is able to represent the important processes and if not, the effort to incorporate them
4. First user workshop with involvement of stakeholders, scientists and developers with the aim to:
   4.1. Review the findings of the previous meetings from the point of view of the stakeholders
   4.2. Discuss the main functions (e.g. analysis, communication) of the system

   4.3. Discuss the main external factors, policy options and indicators that should be incorporated in the system
   4.4. Discuss the institutional context in which the system could be used
5. Prioritisation of possible adaptations to the existing system
6. Adaptation of the system
7. Second user workshop with involvement of stakeholders, scientists and developers with the aim to:
   7.1. Review the adaptations
   7.2. Discuss the potential for its actual use
   7.3. Develop an implementation plan
8. Implementation of the system in the user organisation(s)

As described above, the development –and application– process is an iterative process. The eight steps described in the methodology should therefore not be followed in a very strict sense and time should be reserved for feedback loops between the different steps.

5. RESULTS FROM STAKEHOLDER AND SCIENTIST INTERACTIONS

For the application process we have selected as a first group of potential users the stakeholders who are currently involved in regional development at a strategic level in the Oum Zessar watershed:

- Commissariat Regional au Développement Agricole (CRDA) de Médenine, the representation of the Ministry of Agriculture and water Resources at the regional level (province of Médenine), in charge of the implementation of the agricultural development programs.
- Union Régionale de l'Agriculture et de la Pêche (UTAP) de Médenine, the regional representation of the national farmers union.
- Association des Jeunes de Zammour (AJZ), a local NGO based in the mountain locality of Zammour.

Furthermore, during the entire process, we collaborated with a local organization that takes the role of local scientist as well as potential user:

- Institut des Régions Arides (IRA), a leading research organisation in the field of arid areas and desertification, actively involved in communicating scientific knowledge to policy-makers and other stakeholders.

Besides the potential users defined above, scientists and IT-specialists from King's College London and RIKS were also involved in the application process.

At present, we have carried out the first four steps of the methodology described in the previous section. A summary of the information and recommendations obtained through the interactions with local stakeholders and scientists is provided below.

Through the workshops, the main issues in the region, the important drivers for change and the relevant indicators became apparent. The issues and related indicators are presented in table 1. The drivers cannot be directly linked to one issue or indicator in particular since they impact on the integrated system in its totality. The following main drivers are selected:

- Globalization on the prices of crops;
- Climate change and variability (droughts);
- Changes in EU migration policies;
- Introduction of new crop types ;
- Subsidies;
- Water resource allocation between sectors;
- Soil and water conservation strategy.

On the bio-physical side not many of the processes have to be adapted. The main difference is the cause of erosion. Where rainfall is driving most of the erosion in the current application, wind is a much more important cause in Northern Africa. Furthermore natural vegetation type groups have been adapted to local species and transition rules for succession and degradation have been adjusted.

Table 1. Main issues and indicators in the Oum Zessar watershed.

Main issue	**Indicator**
Sustainable economy	- long term profits in main economic sectors - household income
Water resources	- water shortage in the main sectors
Land degradation and desertification	- environmentally sensitive areas (ESA)

Larger adaptations are required to incorporate the socio-economic processes. Where in many European countries farmers often depend on their income from farming, in Northern Africa the household income in general depends on a combination of the following sources: crops, livestock, tourism or money sent over due to internal (large cities) or external (Europe) migration. Furthermore, household income and survival are much larger issues in Northern Africa than in Southern Europe.

There are also several drivers that are similar in nature, but have a different extent. Climate variation, for example, is much larger in Northern Africa than in Europe. In the model however, incorporating this variability doesn't need any special adjustment and changes can simply be incorporated by entering the correct data sets and parameter values.

At last, there are some proposed adaptations that would also improve the original model when applied to the European areas. Examples of these adaptations are: the possibility to have various crop cycles over the year, the possibility to have mixed land uses (tree and annual crops combined on one field) and the possibility to incorporate grazing densities (the current model deals with grazing as a yes/no component).

Besides the contents of the system, we also discussed its application domain and the way the system could be used in practice in Tunisia. A crucial element mentioned in this regard was the bottom up development of policies. Since local initiatives can be brought forward in the decision-making process at regional level, stakeholders have a high interest in using the system by themselves to investigate the impacts of the actions they propose and to discuss those proposed actions with other stakeholders.

6. POSSIBILITY FOR PRACTICAL USE

This paragraph discusses the findings of the six questions that help to understand the (potential) failure or success of a PSS in practice as posed in the introduction.

According to potential users, the MedAction system can provide an added value to current planning processes by providing some insight in the future development of the region and the interrelation of the different processes at stake. Also its explicitly spatial nature was mentioned as one of its main advantages. From discussions with stakeholders it became apparent that it could provide support to the main issues they are dealing with –regional economic development, water resources and desertification– although some of the details of the processes would need to be adapted to better fit the Northern African context in general and more in particular that of the Oum Zessar watershed in Tunisia.

Availability of data and models does not seem to be a limiting factor to use the system. Since the Institut des Régions Arides has collected a large amount of bio-physical as well as socio-economic data and has a very good understanding of the processes at stake in the region, it should be feasible to produce a first prototype. Based on this, focused data collection can be established to fill in the missing elements and improve the prototype.

At present, the credibility of the system has not been discussed to a vast extent. Since the application to the Oum Zessar watershed has only been set-up recently, calibration and validation are not yet finished. Assumptions made in the model for the Northern Mediterranean have been discussed and questioned and adaptations have been proposed to some parts of the system.

The language question can be taken literally and in a more abstract manner. Through discussions with stakeholders it became clear that it would be a great benefit to them if the language of the system would be French (or Arabic), rather than English. The current system already provides a number of possibilities to enter relevant scenarios and to provide the output in meaningful indicators: environmental sensitive areas, water shortage and long-term profits in the agricultural sector.

When discussing the use of a policy support system, often the distinction is made between the user, the person or organisation actually operating the system, and the end-user, the person or organisation using the results of the model. During preliminary discussions with the scientists in the region it was suggested that the model would be operated by the Institut des Régions Arides, a governmental research organisation that gives advice to local and regional government and other stakeholder groups operating in the region; the organisations that were seen as the potential end-users. After discussions with the stakeholders however, it became evident that they would like to operate the system by themselves as well. Since this requires an in-depth knowledge of the different processes that are incorporated in the system, it will be investigated during the next phase of the process if the human capacity is available at their organisations and if training them in the actual use would be the best way forward.

Both stakeholders and scientist have shown a large interest in using the system. The organizational culture in the different organizations seems very open to explore new initiatives and tools.

7. CONCLUSIONS AND RECOMMENDATIONS

The approach followed in applying the existing MedAction PSS to a new region provides promising results. It focuses on both the context of the system as well as the place it takes in the decision-making process. A first evaluation of the uptake of the MedAction system by regional policy makers and other stakeholders resulted in a positive response. However, given the early stages of the application process it is too preliminary to talk about a successful use of the system in Northern Africa.

Based on the discussions with local stakeholders and scientists as well as developers of the original model, we come to the following conclusions regarding the five research questions posed in section one:

1. In Northern Africa poverty reduction has a higher emphasis than in Southern Europe. Furthermore, climate variability is much higher and with a large part of the region depending on rainfall rather than irrigation, agricultural and natural vegetation are highly impacted by the yearly changes, enlarging the economic risk.
2. There are large similarities between processes related to desertification and regional development in Northern Africa and Southern Europe. In both regions a combination of bio-physical (climate, hydrology, erosion, salinisation, vegetation) and socio-economic (demographic, economic, land use and land management) processes play a crucial role in the dynamics of the region and as a result its future development. The details of the processes as well as the main drivers are however to some extent different, especially when looking at the socio-economics.
3. Compared to the institutional context in Southern Europe, bottom-up processes seem to play a larger role in Northern Africa. This might result in a different way the system will be used in the latter.
4. Main adaptations to the system to make it better applicable for regions in Northern Africa have to do with the way farmer's (or households) make there decisions. The current farmer's decision module assumes that the entire income of the farmer is dependent on agriculture. For the Northern African context it would be more appropriate if the household income depends on a number of factors like crop production, live stock, tourism and income from emigrants.
5. First interactions with local stakeholders and scientists show a large motivation to actually use the system. Added value of the system is seen in the recognition of problems and the integrated impact assessment of –a combination of– policy options on the economic, ecologic and social development of the region. Moreover (potential)

users feel that it can be used in an interactive planning process where actors from different disciplines come together to discuss the impact of different alternatives for improving sustainable regional development. However, the implementation process has just started and only a close interaction with the stakeholders during the coming period can ensure an actual uptake of the system.

The proposed methodology includes the presentation of an existing system to potential users. This approach has a number of advantages, since it shows what is possible and therefore offers a good starting point for focused discussions, but it also has the drawback that the stakeholder group already is confronted with possible solutions before going through a problem definition process. Since model and software development is a very time consuming and therefore expensive process, in our opinion, the main benefits of this approach –reusability of existing material– outweigh its potential limitations.

The methodology followed in the current research included open questions that were answered during workshop sessions. Since there was mutual respect between the different participants and a high degree of openness, this format worked very well. Most likely in the next phase, where certain elements have to be discussed in more detail with the different (science and policy) experts, a format that combines individual meetings and workshop sessions will be more appropriate.

A local (research) partner has been found essential in the application of the MedAction PSS to a new region. Not only has such an organisation insight in the relevant socio-economic and bio-physical processes in the region (or knows who has this insight), it is also able to raise interest with the local stakeholders and select the relevant stakeholders for the application process.

ACKNOWLEDGEMENTS

The authors wish to thank the European Commission for the financial support of the project DeSurvey: A Surveillance System for Assessing and Monitoring of Desertification (IP-003950).

REFERENCES

Chahbani, B., Contribution à l'étude de l'érosion hydrique des loess des Matmata et de la destruction des jessour: Bassin versant d'oued Demmer, Béni Khédache. M.Sc. thesis, Universite de Paris I, Paris, 1984.

De Graaff, J. and M. Ouessar (eds), Water harvesting in Mediterranean zones: an impact assessment and economic evaluation. TRMP paper n°40, Wageningen University, The Netherlands, 2002.

Hurkens, J., B.M. Hahn and H. van Delden, Using the GEONAMICA ® software environment for integrated dynamic spatial modeling, paper presented at this conference, 2008.

Labras. Analyse comparée de l'opportunité technico-économique de l'oléiculture en plaine et en montagne dans le sud-est tunisien. Graduation dissertation, ESA Mograne, 1996.

Mahdhi, N., Sghaier, M, Ouessar, M. 2000. Analyse d'impacts des aménagements de CES en zone aride: cas du bassin versant d'Oued Oum Zessar. Technical report, WAHIA project, IRA, Tunisia, 2000.

Ouessar M. Hydrological impacts of rainwater harvesting in wadi Oum Zessar watershed (southern Tunisia). Ph.D. thesis, Faculty of Bio-Engineering Sciences, Ghent University, Belgium, 154 pp., 2007.

Rahmoune, L. Dynamiques des systèmes agraires du sud est tunisien, cas de la presqu'île de Djorf. M.Sc. thesis, INAPG, Paris, 1998.

Van Delden, H., P. Luja and G. Engelen, Integration of multi-scale dynamic spatial models of socio-economic and physical processes for river basin management, Environmental Modelling and Software, 22(2), 223-238, 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

Usability assessment of an integrated land use model in Portugal

Cárin Alves[1], Lúcio do Rosário[2], Jasper van Vliet [1] and Hedwig van Delden [1]
[1] Research Institute for Knowledge Systems bv, The Netherlands [calves@riks.nl; jvanvliet@riks.nl; hvdelden@riks.nl]
[2] General Directorate of Forest Resources, Ministry of Agriculture, Rural Development and Fisheries, Portugal [lucio.rosario@dgrf.min-agricultura.pt]

Abstract: Regardless the possibilities of computer based tools for spatial planning, previous works have argued that planners have never fully used them. Some of the reasons that influence the acceptance and effective use of a new technology are the usability and the user's perception of its relevance to improve their tasks. This research aims to assess the usability and usefulness of a planning support tool by planning organizations in Portugal. The tool has proven to be easy to use and very useful to users; however, some technical improvements were suggested in order to improve it.

***Keywords*:** integrated planning approach; spatial decision support systems; usefulness; usability.

1. INTRODUCTION

Over the last decades spatial planning has evolved from a process with strong sectoral focus into a multi-disciplinary approach. This has increased the complexity of the spatial planning process and therefore the need for computer based tools to support decision making [Geertman and Stillwell, 2002]. Despite this progress, previous studies have concluded that in fact users consider these new technologies inappropriate for certain practices belonging to the planning process, such as forecasting, analysing and evaluating; and that their use is for the most part too general [Brail and Klosterman, 2001]. Reasons differ from technological to political context characteristics.

This paper focuses on the study of *usefulness* and *usability* of a spatial decision support system developed for the recognition, analysis and communication of problems related to the desertification processes. We believe these two aspects influence the success of a system and are both important to developers and end users. With this we hope to provide suggestions as to how we can improve such system.

This investigation is a case study in the Portuguese context and takes into account *ScenDes [Scenarios for Desertification]*, a spatial decision support system, developed in the scope of the DesertWatch project of the European Space Agency. The guiding questions for this research are:

- Is *ScenDes* useful/relevant for the assessment of desertification in the Portuguese context?
- Is *ScenDes* easy to use by planning technicians in Portugal?
- How can the current system be improved to be more effectively used in Portugal, in combating desertification?

2. CONCEPTUAL FRAMEWORK

2.1 The case of Desertification

More than one hundred formal definitions of desertification are believed to have been proposed so far [Diez-Cebollero and McIntosh, 2005], nevertheless the most cited definition nowadays is the one proposed by UNEP in 1990 and adopted by UNCCD in 1999. In this paper we adopt this same definition in which desertification is "land degradation in arid, semi-arid and dry sub-humid areas resulting from various factors including climatic variations and human activities" [UNCCD, 1999]. Land degradation is the "reduction and loss of the biological and economic productivity caused by land-use change, or by a physical process or a combination of the two" [Geeson *et al*, 2002]. It is considered one of the most alarming processes of environmental degradation and a threat to natural resources with consequences on food security, poverty and environmental and political stability [Sivakumar and Ndiang'ui, 2007].

Desertification is seen as a process, a sequence of casually linked changes. In the opinion of Diez-Cebollero and McIntosh [2005], "desertification might also be understood as a feedback process in which, as a consequence of lower productivity, farmers are forced to abandon infertile areas and exploit new land […]. Abandoned areas become then more prone to biological degradation". Consequently, we agree that desertification does not only affect biophysical processes, but also the socio-economic structure of affected areas.

2.2 Systems analysis for planning support

The problem of desertification, just like many other spatial planning-related problems, is complex and unstructured. By complex, we mean that there are large numbers of actors, factors and relations and phenomena occur at different temporal and spatial scales. By unstructured, we mean that there is uncertainty relative to the knowledge of solving the problem and that involved actors have often conflicting aims. Solutions require the competency of experts from different fields and a high level of public participation and inclusion of heterogeneous sources of knowledge [Voss *et al*, 2004, Sterk, 2007].

As mentioned in the previous paragraph, spatial problems consist of a diversity of elements linked together, interacting strongly. For planners, it's essential to understand how these elements act and react upon each other, to be able to point solutions and think of alternatives. Systems analysis is useful because it enables us to understand problems in its totality and dynamics. It rises above the enormous detail, at the same time it focuses on how individual parts interact with one another [Golley and Bellot, 1999]. The use of computer-animated systems – usually named models – can be very helpful for a systemic analysis. Planners can observe the effects of different actions and scenarios as they get a more logical and organised representation of the elements and their relationships. Decision makers can consequently make use of this capability to experiment policy actions [Grossmann and Bellot, 1999].

2.3 Computer-supported planning tools

Diez-Cebollero and McIntosh [2005] point out that computer models play useful functions in policy and management, mostly because they help to identify and set problems/issues on the political agenda, and help to visualize and explore future alternatives. They also facilitate political consensus and can be used as management tools.

There is different nomenclature to address these planning tools. Hereby we try to distinct and define some of the most common names used in this field.

As for Geertman [2006], he defines planning support as "all professional help in the form of dedicated information, knowledge, and instruments that people actively involved within formal spatial-planning practices can receive to enlighten their planning tasks and activities". In this perspective, planning support instruments refer to "computer-based tools,

dedicated to the support of specific professional spatial-planning tasks such as problem diagnosis, data collection, spatial and trend analysis, geo-data modelling, spatial scenario building, visualization and display, plan formulation, enhancing participation, and collaborative decision making". Planning Support Systems [PSS] can be understood as "geoinformation-technology-based instruments that incorporate a suite of components [theories, data, information, knowledge, methods, tools] which collectively support all or some part of a unique professional planning task". In addition, "these PSS aid the planning process by providing integrated environments […] in which three components are combined: 1] the specification of planning tasks and problems […], 2] the system models and methods that inform the planning process through analysis, prediction and prescription, 3] and the transformation of basic data into information which in turn provides the driving force for modelling and design" [Geertman, 2006].

As for Uran and Janssen [2003], tools that are explicitly designed to provide the user with a decision making environment that enables the analysis of geographical information to be carried out in a flexible manner are named Spatial Decision Support Systems [SDSS]. These are designed to assist the spatial planner with guidance in making land use decisions, and are sometimes referred to as Policy Support Systems [PoSS] [Kok and vanDelden, . The following components are taken in consideration: 1] a database management system, such as a Geographical Information System; 2] a library of potential models that can be used to forecast the possible outcomes of decisions; and 3] an interface to aid the users' interaction with the computer system and to assist in analysis of outcomes [Sprague and Watson, 1996]. It is our believe that both Geertman's Planning Support Systems and Uran and Janssen's Spatial Decision Support Systems are the same category of planning tools. In this paper we will address to *ScenDes* as a Spatial Decision Support System, but we agree that could also be named a Planning Support System, in Geertman's way.

2.4 The *ScenDes* SDSS

The *ScenDes [Scenarios for Desertification*] SDSS is a model that simulates land use change under different scenarios, and is a part of an integrated information system built in the scope of the DesertWatch project. In order to represent the processes that make and change the spatial configuration of the area, *ScenDes* is a layered model representing processes operating at three geographical levels.

At the scale of the entire area [1], the model uses economic, demographic or environmental growth scenarios, prepared by planning agencies or stakeholder groups. From these, growth figures for the global population, the economic activity and the expansion of particular natural land uses are derived and entered in the model as global trends. National growth will not evenly spread over the modelled area due to regional inequalities. Location and relocation of residents and economic activity will thus be influenced by regional characteristics [2]. At the local scale [3] detailed allocation of economic activities and people is modelled by means of a constrained Cellular Automata land use model. In principle, it is the relative attractiveness of a cell, together with local constraints and opportunities, which cause cells to change from one type of land use to another. Four elements determine if a cell is taken in by a particular land use function or not, the local neighbourhood, the accessibility to infrastructure, the physical suitability of the territory and its zoning status [RIKS, 2006].

2.5 Usability and Usefulness

As we mentioned, we consider usability and usefulness key elements for users' acceptance of new software or any other technology. We agree with Kraemer *et al* [1993] that *usefulness* is "the degree to which an individual believes that using computer-based information enhances his or her work". It is the case when users value the generated results higher than the investment [in time, data requirement, etc.]. Nevertheless, we believe that usefulness is not purely perceptual, but can also be measured objectively. Useful software does what it is built to do and actually helps to solve users' problems.

However, any software or technology can be very relevant and useful to users but may not be easy to use [and vice-versa]. *Usability* measures the quality of a users' experience when interacting with the software. In total, if it is easy to use, easy to learn, easy to remember, error tolerant, subjectively pleasing and performs the task for which it is being used. It normally incorporates factors such as design, functionality, structure and information architecture.

We consider both *usefulness* and *usability* important for users, developers and even managers. Taking in consideration that these tools are expected to serve specific user needs, the developers must have these in mind when constructing the product. On the other hand, users must recognize the possibility of solving their needs through such product. From the user perspective, easy-to-use softwares are important because it can make the difference between performing a task accurately or not and enjoying the process or feeling frustrated. From the developer's perspective, it is important because it can mean the difference between the success and failure of a system. And as for managers, software with poor usability can reduce the productivity of the workforce to a level of performance worse than without the system.

Despite the fact that in the past years planners and/or decision makers have used a sufficiently large number of methods and tools to support planning activities, it can be stated that they "have never fully embraced the diversity of the available methods, techniques, and models developed in the research laboratories to analyse spatial problems, to evaluate future options, or to project scenarios" [Geertman, 2006]. Throughout literature many reasons are pointed out as for this disenchantment with decision support technologies. We herewith try to outline those by differentiating problems existing internally in the technologies itself and problems which arise from an external context to the technologies, but may influence planners' opinions.

Table 1. Most common problems

Internal Problems		**External Problems**
Usefulness related	**Usability related**	
A. Planners think models are too *inflexible* to respond to always evolving needs [Walker, 2002]. B. Model outputs were seen as *uncertain* and their appropriateness doubted [Uran and Janssen, 2003]. C. Planners often find tools *irrelevant* regarding their needs [Walker, 2002]. D. Planners not always understood the technologies proposed and *its capabilities and limitations* [Walker, 2002].	A. Models are often considered *too complex*, *too detailed*, *time consuming* and there is *need for training* [Uran and Janssen, 2003]. B. Many tools *do not* fully *meet the original specifications* [Walker, 2002].	A. End-users are often *not involved* in the model development phase, causing wrong purposes [Uran and Janssen, 2003]. B. Planning problems are *wicked*; experts often do not agree about the best solution [Brail and Klosterman, 2001]. C. The rapid change in computer technology, and also in the modelling technology leads to a constant feeling of *computer illiteracy* [Landauer, 1995]. D. Failure in the implementation of such tools is often due to *inaccessibility*, or *lack of confidence* from the decision makers, or *institutional or political barriers* [Walker, 2002].

The identification of problems that we have considered to be related with usefulness and usability motivates us to carry out this research. Although we acknowledge the external problems in Table 1, they are not part of the current research.

3. METHODOLOGY

With the intention of studying the usefulness and usability of *ScenDes*, as well as to understand how we can improve the tool, we have contacted technicians, managers and researchers that are currently working on the theme of desertification in various organizations in Portugal. They have all participated previously in data gathering and field validation for the DesertWatch project, which *ScenDes* is a part of.

With the purpose of assessing the usability of the software we have targeted the technicians and researchers, as they are the ones who use the tools in the practical sense and have to learn its functionalities and requirements. And, in order to understand the usefulness of the software, we have focused on the managers/decision makers, as they are the ones who make the final decision regarding the implementation of the system, based on its relevance in obtaining the organisational goals. However, we have noted that, in less complex-structured organisations, managers/decision makers can also be technicians.

In this research we have performed semi-structured interviews and questionnaires. Also, we have visited the organisations involved and organized a workshop. First, we have booked visits to the organisations in order to meet the users and discuss their tasks and responsibilities in their organisations. In this first phase, we wanted to identify the two target groups. During these visits, we have conducted semi-structured interviews. In a next phase, we wanted to observe users while using the software, therefore we have organised a workshop, in which all users were invited to complete some exercises. During this workshop we were able to identify in which parts of the simulation procedure they had difficulties [meaning functionalities that needed to be further developed] and also, we could directly have their opinions and questions on the performance of certain tasks as well as the data requirements of the system. Finally, in the end of the workshop we have asked users to complete a questionnaire. Questions regarding usability were asked to technicians and researchers; and questions regarding the system's relevance/usefulness were given to managers. We have also included a space for comments, which gave further input to this research.

4. PRELIMINARY RESULTS

We have visited 8 different organizations, which consisted of national administration and regional administration bodies, civil society associations with a local interest and a group of researchers belonging to a public university. In total, we have consulted 22 people.

Tables 2 and 3. Participating Organisations and Use

Organisations:	No.
National scope	2
Regional scope	2
Local scope	3
Research driven	1

Users' position:	No. of People
Only Technicians or Researchers	15
Managers and Technicians	6
Only Managers	1

ScenDes was considered an easy-to-use tool and users have found it very relevant and useful for the organisations they belonged to. Nevertheless, *ScenDes* can be improved technically which, from our perception, would permit a more effective use. We herewith point out some positive and negative aspects considered as well as some suggestions from users.

Positive Aspects:

- *ScenDes* has an easy working environment and can be learnt easily and quickly.
- Users were generally satisfied about the possibilities of *ScenDes* and would consider using it in their daily work.
- The procedures were not too lengthy and users understood the steps. Error messages were not reported.
- Users identified *ScenDes* as a relevant tool for exploring planning options, not only for desertification, but also for land planning in general.

Negative Aspects:
- *ScenDes* implies GIS software and knowledge. Data treatment must be done *a priori*.
- *ScenDes* is not in compliance with the most used map formats in planning organizations – vector. It demands some extra work in converting and transforming data into raster data.
- *ScenDes* presents some absence of information [units on graphs and legends] and it is not careful on the management of cartographic data [no information on source and north arrow.

Besides the correction of mentioned problems, users suggested:
- A longer training time, since modelling tools have not yet been extensively used in planning organizations in Portugal.
- A clearer interface, meaning only parts that users can edit, should be visible.
- A more interactive tool, meaning more possibilities for users to add their numerical data, as well as regional related information.
- An easier to read output; not only a map, but also the highlight where major changes have occurred as well as variation indexes.

Some additional remarks:
- Users stated that *ScenDes* could not be used by regional or local organisations, because it was built for national scope analysis. Also, they have complained about the aggregation of some land use characteristics and of basic data. For some users, the land use map that was used did not include important national characteristics, such as the crop rotation system in agricultural areas. Although *ScenDes* is not suitable for very detailed analysis, for example city blocks, it can be set up for regional and city analysis. Pixel size, land uses categories and their level of aggregation can be chosen by users, thus overcoming the abovementioned remarks of the users.
- Users stated that introducing *ScenDes* in their organisations would not imply any organisational change. Our impression is that because users just briefly got to know about the possibilities of *ScenDes*, its use will be limited for now. We believe that as users start using *ScenDes* more effectively, they will understand the integrated approach that it demands as well as the work among different organisations that will be required. In our opinion both could very well lead to some organisational change.
- Users stated that they would not prefer a simpler tool. They prefer to have more control over data and changes, despite the complexity that it implies. This reflects users to be more interested on the process, than just the results. This is a positive aspect for an effective use of *ScenDes*, and SDSS tools comprising an integrated model in general.

5. CONCLUSIONS AND SUGGESTIONS

5.1 On *ScenDes* usefulness

Managers make a positive evaluation of *ScenDes* usefulness. It not only helps to study the desertification processes, but is also useful in land planning in general. Managers recognise that *ScenDes* enhances their organisation's work, and that it is particularly good to accompany the planning process, especially when setting goals and objectives, inventorying, evaluating alternatives of action and monitoring. Managers agree that *ScenDes* brings decision makers/technicians/civil society together, promoting communication and discussion.

5.2 On *ScenDes* usability

The majority of users have agreed *ScenDes* has an intuitive interface and is easy to use. In general, users were satisfied with the possibilities *ScenDes* offers. The tasks were not too long, easy to remember and no error messages were reported. Users understand the required investment in time, knowledge and data in order to fully work with the system and do not

consider it disproportional or excessive. Users claimed *ScenDes* was easy to learn, though training was needed.

5.3 On how *ScenDes* can be improved

Based on the results summarized in 5.1 and 5.2, we hereby conclude how *ScenDes* can be improved.

- Communication

Users do not fully understand the technical requirements for using *ScenDes*. Technical requirements such as GIS knowledge and the need for data treatment should be more clearly stated *a priori*. Users must know in advance if and how they can meet these requirements as we believe the opposite situation can jeopardize the use of the system.

- Training

Since modelling has not been extensively used as planning support, the provided training of half of a day was not sufficient for users to get realistic expectations on the possibilities and limitations of *ScenDes*. Training sessions should be longer and users' examples and interests/needs must be taken in consideration during training sessions.

- Technical Aspects

During the workshop, we have identified some missing information in the user interface that confused users. Some technical aspects can be improved such as introducing the missing information on charts, graphs, maps, as previously mentioned. By doing so it will be easier to understand what they can and cannot change. At the output level, in addition to the final map, variation analysis could be automatically calculated.

It is the authors' impression that users didn't get a full understanding of the possibilities and limitations of a SDSS such as *ScenDes*. Besides the comments already made regarding communication and training, this can be a symptom of the fact that most SDSS are born within research programmes with a strong scientific interest rather than in the spatial policy context with emerging needs. This aspect can nevertheless be minimized by increasing the communication between developers and users. An example of an iterative process with strong interactions between developers and users can be found in van Delden et al [2008].

Reflecting on the methodology used to assess usefulness and usability, we would like to make the following comments: On one hand, the interviews made it possible to understand the organisations' agenda and their managers. This was very positive because it enabled us to present *ScenDes* to them as an available tool for their problems and get them involved in this research. But because of their lack of time, only few managers tested the tool in the workshop. We believe that not having touched the tool could lead to a not very informed opinion about the relevance of the tool for their work.

As for the usability assessment, it was very positive to have adopted the workshop methodology. It permitted users to communicate more with the developer and discuss questions on difficulties in real-time. A less positive aspect was that we used theoretical examples, and not real cases. This was so because of the data requirements and shortage of time for this research. But we agree that the study of usability would be more reliable if users could work on a real planning question from their organisations. Therefore, to conclude this research, we will organize another workshop in which we will use real examples from the involved organisations.

ACKNOWLEDGEMENTS

The authors wish to thank the European Space Agency, the Mediterranean Agronomic Institute of Zaragoza as well as Regina Pimenta, Andreia Silva, Luísa Santos, Nelson Castro, Rita Farropas, Susana Leal, Carlos Conde, Gabriela Freitas and Margarida Nicolau [DGOTDU], Ana Maria Almeida [DGRF], Maria José Roxo, Bruno Neves, Magda Matias and Nuno Santos [Universidade Nova de Lisboa], Paula Vaz and André Mascarenhas [CCDR Algarve], José Faustino, Colatino Simplício, Lília Fidalgo and Paula Serra [CCDR Alentejo], Filipa Vidas [AFLOPS], Paula Bento [PrezaNatura], Hugo Pinto and Hugo Vieira [CUMEADAS]

REFERENCES

Brail, R. and Klosterman R. [eds.] *Planning Support Systems: Integrating Geographic Information Systems, Models and Visualization Tools*, ESRI Publications, NY, 2001.

Diez-Cebollero, E. and McIntosh, B. *DeSurvey – A Surveillance System for Assessing and Monitoring Desertification, WP1.8.1 – Deliverable D1.8.1.1 Report*, Cranfield University, Cranfield, 2005.

Geertman, S. and Stillwell, J., Planning Support Systems: An Introduction, Geertman, S. and Stillwell, J. [eds.] *Planning Support Systems in Practice [Advances in Spatial Science]*, Springer-Verlag, Berlin, 2002.

Geertman, S., Potentials for planning support: a planning-conceptual approach, *Environment and Planning B: Planning and Design*, 33, 863-880, Pion, London, 2006.

Geeson, N.A., Brandt, C.J. and Thornes, J.B. *Mediterranean Desertification – A mosaic of Processes and Responses*, John Wiley & Sons Ltd., Chichester, 2002.

Golley, F. and Bellot, J., Planning as a way of achieving sustainable development, Golley, F. and Bellot, J. [eds] *Rural Planning from an Environmental Systems Perspective*, Springer Verlag, New York, 1999.

Grossmann, W. and Bellot, J. Systems Analysis as a tool for Rural Planning, Golley, F. and Bellot, J. [eds] *Rural Planning from an Environmental Systems Perspective*, Springer Verlag, New York, 1999.

Kraemer, K., Danzinger, J., Dunkle, D. and King, J. *The usefulness of computer-based information to public managers*, CRITO-Center for Research on Information Technology and Organizations, University of California, 1993.

Kok, K. and van Delden, H. Combining two approaches of integrated scenario development to combat desertification in the Guadalentin watershed, *Environment and Planning B: Planning and Design*, 34, Pion, London, 2007.

Landauer, T., *The trouble with computers: usefulness, usability and productivity*, MIT Press, Cambridge, 1995.

RIKS, *ScenDes Users Manual*, RIKS, Maastricht, 2006.

Sivakumar, M. and Ndiang'ui, N. [eds.] *Climate and Land Degradation*, Springer-Verlag, Berlin, 2007.

Sprague, R.H. and Watson, H.J. *Decision support for management*, Prentice Hall, NJ, 1996.

Sterk, B., *A window of opportunities – The contributions of land use modelling to societal learning*, Wageningen University, Wageningen, 2007.

UNCCD, *UNCCD in those countries experiencing serious drought and/or desertification, particularly in Africa*, Secretariat of the Convention to Combat Desertification, 1999.

Uran, O. and Janssen, R., Why are spatial decision support systems not used? Some experiences from the Netherlands, *Computer, Environment and Urban Systems*, 27, 511-526, Elsevier, Amsterdam, 2003.

van Delden, H., Ouessar, M., Sghaier, M., Azaiez, O.B., Mulligan, M., Luja, P., de Jong, J., Fetoui, M. and Ben Zaied, M., *User driven application and adaptation of an existing Policy Support System to a new region*, paper presented at this conference, 2008.

Voss, A., Denisovich, I., Gatalsky, P., Gavouchidis, K., Klotz, A., Roeder, S. and Voss, H., Evolution of a participatory GIS, *Computer, Environment and Urban Systems*, 28, 635-651, Elsevier, Amsterdam, 2004.

Walker, D., Decision support, learning and rural resource management, *Agricultural Systems*, 73, 113-127, Elsevier, Amsterdam, 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

Participatory modelling of fire prevention and urbanisation in southern France: from co-constructing to playing with the model

M. Etienne[a], M. Bourgeois[a], V. Souchère[b]

[a] Ecodevelopment Unit, INRA, Site Agroparc, 84914 Avignon Cedex 9, France
(etienne@avignon.inra.fr)
[b] UMR Sadapt, INRA, , BP1, 78850 Thiverval-Grignon, France
(souchere@grignon.inra.fr)

Abstract: In 2006, officers from the Nîmes Métropole urban community and from the Agriculture and Forestry Service decided to apply a participatory modelling process to sensitise the representatives of the community to fire prevention issues at the interface between natural areas and urban zones. The approach went through 3 phases. The first consisted of co-constructing, according to the ARDI method, a conceptual model of the current situation, in a group of technicians and local policy makers. The process was facilitated by a researcher used to the companion modelling approach and in charge of encouraging the participants to follow precisely the 4 steps of the participatory modelling process. The second phase was to implement the conceptual model into an agent-based model. The last was devoted to playing with the model during series of role-playing game (RPG) sessions set-up to stimulate discussions between the representatives of the community, urban developers and local policy makers. After a short description of the time schedule of the participatory process, the agent-based model and the corresponding RPG set up are briefly described. Specific skills developed during the co-construction stage and the way the model was progressively adopted by the participants are discussed. The function of the model during the role-playing game phase is analysed and its capacity to enhance discussions on the interactions between forest fire prevention, cropland abandonment and urban development is discussed. Particularly, the way players behave during the game and argue during the debriefing of the RPG session is described and the way the impacts of the participatory modelling exercise were evaluated is described.

Keywords: Participatory modelling, agent-based model, role-playing game, fire prevention, urban development, cropland abandonment.

1. INTRODUCTION

In 2006, officers from the Nîmes Métropole urban community and from the Agriculture and Forestry Service were looking for an effective tool to sensitise the representatives of the local villages to fire prevention issues at the interface between natural areas and urban zones. Fascinated by the approach previously developed for fire prevention in French Mediterranean forests (Etienne, 2003), they decided to try a participatory modelling approach. The approach based on the ARDI method (Etienne, 2006; Etienne et al., 2008) was applied by co-constructing an agent-based model in a group composed of technicians covering the main activities developed in the area, together with local policy makers. The process was facilitated by a researcher used to the companion modelling approach (Collectif Commod, 2006) and in charge of encouraging the participants to follow precisely the 4 steps of the participatory modelling process.

2. THE APPROACH

The principle of companion modelling (Collectif Commod, 2006) is based on a dynamic perception of the decision making process. Decision making is therefore considered as the result of interactions between individual stakeholders and groups advocating contrasting representations of the world and having different level of power in the negotiation process (Weber, 1995). When applied to decision support in natural resources management, it is supposed to facilitate a negotiation process aimed at transforming the interactions between ecological and socio-economic dynamics. Modelling is a crucial aspect of the approach because it is used as an efficient mean of building a shared (but not unique) representation of a complex situation, accounting for the dynamics of the system and simulating management scenarios.

What makes companion modelling original is the way the models are designed and used, and the involvement of the stakeholders in the modelling process. The main goal is to help practitioners, experts, or policy makers to elicit and share their points of view on a given complex question. Companion modelling promotes a reflexive use of models by setting up participatory workshops where stakeholders will learn collectively about a complex system by constructing, adapting, manipulating or observing a model (Collectif Commod, 2008). Generally, the approach goes through six to nine phases (Figure 1):

- sensitising: to convince stakeholders concerned by the same question that companion modelling can be an efficient way of starting collective thinking on that question
- co-conceiving: to jointly construct a shared representation of the question and the corresponding complex system
- monitoring: gathering available information and getting relevant new data on the system
- implementing: developing the computer model (commonly an agent-based model)
- validating: making stakeholders feel comfortable with the model
- visioning: playing with the model (role-playing game) or simulating scenarios
- feeding back: getting stakeholders not involved in the participatory process aware of what happened and what are the main outputs
- evaluating: measuring the impact of the participatory modelling approach on stakeholders practices and decision making
- training: getting participants self-sufficient on applying the approach on other topics

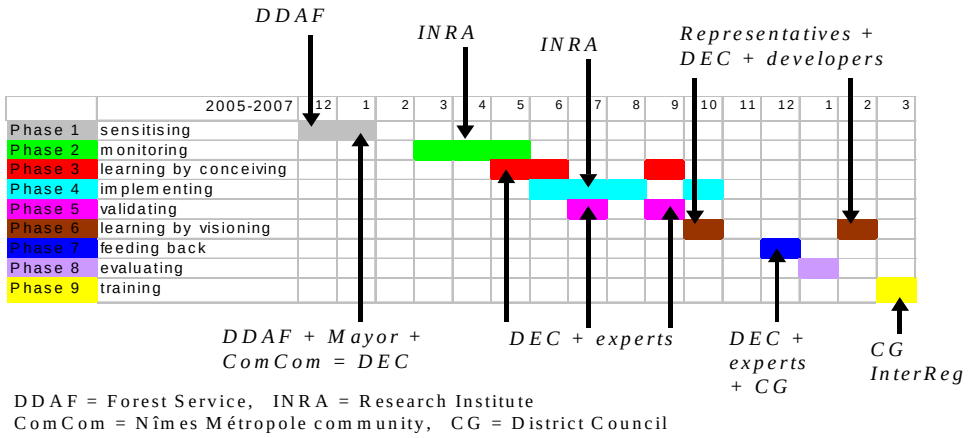


Figure 1: Time schedule and stakeholders involved in the companion modelling process of the Nîmes Métropole study case

In the Nîmes Métropole case study, participatory workshops were focused on phases 3, 5 and 6. But phase 2 was also used to gather and share spatial information on forest and urban dynamics, and knowledge of the main practices of local stakeholders (farmers, developers and foresters). Phase 3 was devoted, on the one hand, to elaborating and validating a virtual spatial representation of 3 contiguous villages taken to represent a typical organisation and structure of the northern fringe of Nîmes city; and, on the other

hand, to co-construct, a conceptual model representing the current functioning of the territory and the most likely trends in the next 15 years. Phase 6 was focused on learning by playing with the model during a series of role-playing game (RPG) sessions set up to stimulate discussion between the representatives of the 14 villages, urban developers and local policy makers.

3. CO-CONSTRUCTING A MODEL WITH STAKEHOLDERS

As mentioned in Figure 1 (DEC group), the core group leading the participatory process was composed of a mayor (vice-chairman of Nîmes Métropole), the officer in charge of environmental issues at Nîmes Métropole, the officer in charge of fire prevention at the Agriculture and Forestry Service and a researcher used to the companion modelling approach. At this stage, the core group paid special attention to the convocation of the working group, particularly on 4 aspects: choice of the participants, venues, scheduling of the workshops, mode of invitation (Etienne et al., 2008). The first point was discussed at length because the richness and relevance of the representations elicited during the co-construction exercise depend on the representativeness of the participants. Two options were considered: to work with real stakeholders or with experts from the extension services. The first of these was finally discarded due to the difficulty of selecting farmers to represent each dominant agricultural activity. Instead we chose officers from the extension services (agriculture, forestry, hunting, urban planning) who were thought to already have an overall view of the dominant practices of the main stakeholders. For the three other points, the workshops took place at Nîmes Métropole office, every 3 weeks, and participants were invited by the Nîmes Métropole officer to discuss the following question: how to reduce forest fire hazard at the urban fringe ?
The co-construction of an agreed representation of the question followed the ARDI method (Etienne et al., 2008), but a preliminary workshop was organized in order to define the environment and spatial scale to support the co-construction. According to the land use typology of the 14 villages surrounding Nîmes city (Bourgeois, 2006), 3 archetypes were defined and a virtual map gathering these 3 archetypes was proposed by the core group and designed on a GIS as a contiguous territory. This map was then modified and validated by the working group. After that, three participatory modelling workshops were organized in order to answer the 4 following questions:
Who are the main stakeholders involved in or with a duty to play a decisive part in fire prevention on this territory ? What are the principal resources of the territory and the key information needed to guarantee a sustainable fire prevention ? What are the main processes that drive huge changes in resource dynamics ? How does each stakeholder use the resources and modify the processes?

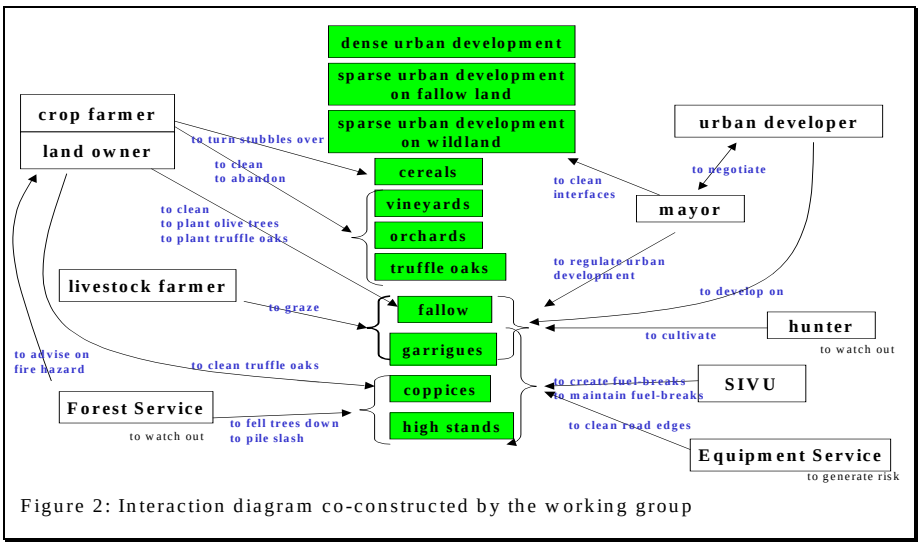


Figure 2: Interaction diagram co-constructed by the working group

This first participatory process led to a conceptual model of the current situation expressed by a diagram representing the key interactions between stakeholders and resources in relation to current trends in urban development and fire hazards (Figure 2).

4. THE AGENT-BASED MODEL

The co-constructed conceptual model was implemented by the research team into an agent-based model. The environment is modelled by way of a cellular automaton representing 3 virtual neighbouring villages covering the most common types of urban density and forest/cropland ratio measured around Nîmes city. The spatial grid is made of 83 x 69 cells, each of 1 ha. The vegetation viewpoint (Figure 3) provides 18 land use types with 7 types of croplands, 4 types of urban development and 7 types of native vegetation. This environment changes according to 3 ecological processes: fallow encroachment on abandoned croplands, natural regeneration after wildfires, natural succession from open garrigues towards Holm oak coppices or Aleppo pine woodlands.

Four agents play an active role in the model: urban developers, mayors, farmers and the local Forest Service. The developers propose to establish new developments near to current urban areas but with different levels of standing and price according to the past land use of the land: olive grove, fallow, garrigue or pine forest. To promote their project, they have to negotiate a building permit provided by the mayors according to their urbanising strategy (increasing the density of the present urban area, extending development on flat lands or on the hills). This strategy is defined in the local urban plan (PLU) that fixes for 10 years the areas where development will be permitted. When revising the PLU, new roads are planned giving access to new areas for development. Farmers take care of the crops mainly located in the flat lands and decide on key maintenance practices such as weeding the vineyards or ploughing the stubbles. When their crops are facing an economic crisis, they receive CAP subsidies to uproot vineyards or abandon cereal cropping on fields located near to urban areas. Finally, the Forest Service is gradually creating a strategic fuel-break perpendicular to the main wind direction and connected to well-maintained crop fields.

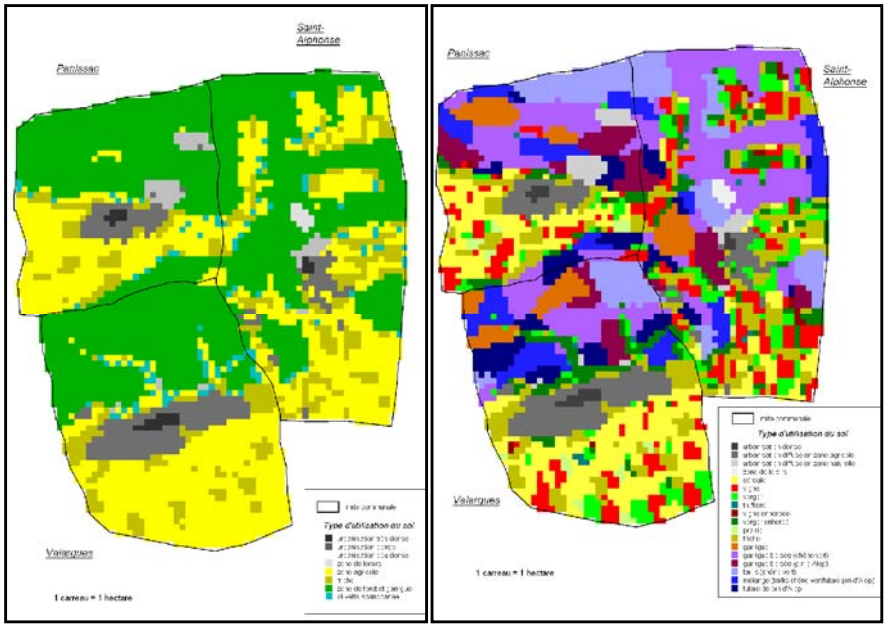


Figure 3: Vegetation viewpoint used in the model (right) and in the RPG (left)

The model runs on an annual time step, so the vegetation viewpoint is updated every year and provides a representation of land use at the beginning of summer. A wide range of indicators is calculated which can be plotted on graphs (i.e. fallow land area, urbanising ratio, burnt area, number of houses affected by fire) or charted on maps (tree cover and

dominant species, age of fallow lands, urban density, etc). Any indicator mentioned as a key data item for taking decisions or evaluating their activity by the participants during the co-construction workshops was implemented in the model.

5. PLAYING WITH THE MODEL

For discussing collectively on visions of the future, the role-playing game tool was selected because it allowed points of view to be shared and opinions challenged in a friendly way (Bousquet et al., 2002). This phase is very important because it encourages the participants to question each other's points of view, to discover the impact of their practices on the other participants' practices and resources, and to experience clear power relationships. It also reveals the difficulty of getting agreement on land management in complex situations. That is why, at this stage, it is newly necessary to pay special attention to the convocation process (Etienne, 2008): choice of the participants, location of the RPG sessions, mode of invitation. It was considered that Nîmes Métropole was entitled to urge the participants to discuss wildfire prevention, and the invitation signed by the vice-chairman presented the RPG as a mediation and collective thinking tool. The sessions were organised in one of the villages of Nîmes Métropole community, but inside a folk museum funded by the District Council. At each session 3 mayors, or their delegates for environmental or urbanisation issues, from 3 different villages, were invited as well as a developer, the officer of Nîmes Métropole in charge of environmental issues and one forest technician from the Forest Service.

The environment was similar to the environment of the agent-based model with the 3 neighbouring villages that were given a nickname like a local traditional name (i.e. Panissac). As mayors get confused with detailed vegetation maps, land use categories were extremely simplified with only 5 types: garrigues and forests, urban areas, olive groves, crop fields and fallow lands (Figure 3). Urban areas were divided into 3 levels of population density, and olive groves were identified as weeded or not. Special attention was paid to the spatial layout of the room (Figure 4) with 3 boxes for the 3 mayors, a big round table for Nîmes Métropole, 2 small tables for the Forest Service and the developer, and an interactive white board used both for representing the cadastral survey and land use map. The computer and the model were apart in a corner of the room.

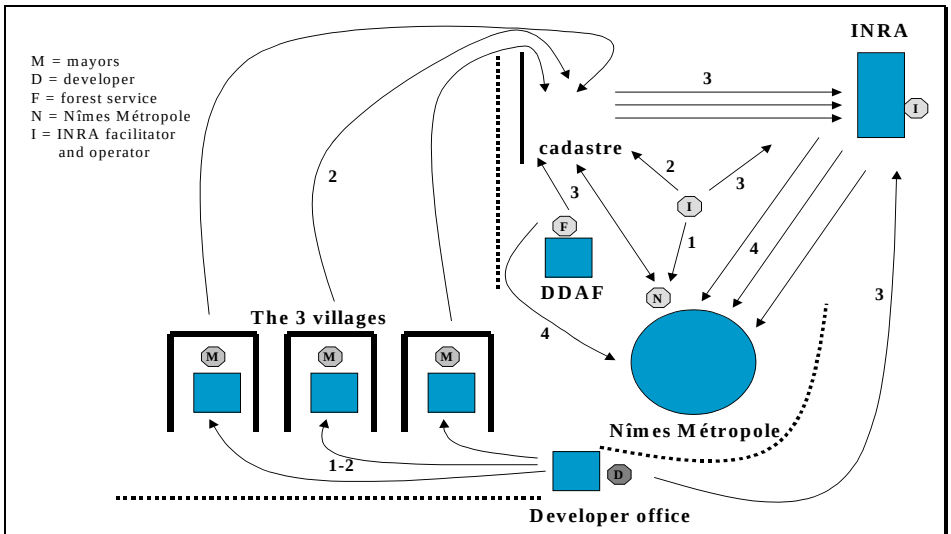


Figure 4: Spatial design of the RPG room and movements made by the participants during the 4 steps of one round

Each round of the game was made up of 4 steps. First, while the participants acting as mayors defined the limits of the urban zone (land suitable for development) and established the range of land prices according to the cadastral plot location, the participant acting as an urban developer threw 3 dice to sort the area he could develop over the next 3

years and thought about the characteristics (density, livelihood) of the corresponding developments. Secondly, a set of negotiations between the 3 mayors and the developer took place in order to decide the location and density of the developments to settle during the 3 following years. Simultaneously, the participant acting as the Forest Service designed a fuel-break aimed at reducing the fire hazard for the forest and advised the mayors on the development of his village with major sensitivity to fire. At the same time, the participant acting as Nîmes Métropole sorted an agriculture or recreation project and thought about the best place to establish it. Thirdly, the mayors went to the cadastral interactive map to identify the coordinates of the plots they agreed to develop and signed the corresponding building permit. Finally, the computer operator entered the decisions made by the participants, ran the model for 3 years (cropland abandonment, urban development, shrub and pine encroachment, wildfire), and printed the new land use maps of each village. Meanwhile, Nîmes Métropole invited the participants to negotiate about the sorted project. At the end of the negotiation, a decision has to be made on what to do, where and who pays. At the beginning of the following round, the facilitator made participants aware of the new value of their key indicators: cash, popularity rating, area burnt by wildfires, and land use maps. Depending on time availability of participants, after 3 to 5 rounds, the RPG is stopped, and after a short break, all the players are invited to discuss what happened during the session. The debriefing is mainly focused on 5 aspects: how participants feel, what do they think about the realism of the model, what was their individual strategy, did they become aware of vegetation dynamics and its impact on fire propagation, and what happened during the periods of negotiation.

6. EVALUATING PARTICIPATORY MODELLING

The Nîmes Métropole case study is part of a large-scale exercise in which about twenty companion modelling (ComMod) research projects were evaluated. The aim of the evaluation was to assess the impact of the ComMod approach in different ecological and socio-economic contexts around the world, and to get ideas about how to improve the ComMod approach in the future. Within this framework, a common designer questionnaire and participants' evaluation framework were elaborated and applied to Nîmes Métropole case study. The designers' questionnaire was completed by the project designer, in order to capture the designers' initial perceptions of the context and to record how it changed over the lifespan of the project. It also permitted the methods and associated artefacts used during the project to be identified and their impacts on the participatory and learning process to be analysed. The objective of the Participants' evaluation framework was to assess how the participants' experiences corresponded to the project team's perception of how the project was carried out.

For the Nîmes Métropole case study, it was decided to carry out a third-party evaluation conducted by an external researcher in order to minimize bias and get more confident results. The evaluator attended only the five sessions of the role-playing game as a simple observer, and was totally unaware of the contents of the project before beginning the evaluation process. Thus, although she had personally applied a ComMod approach in another context (runoff in a crop watershed), she can be considered as neutral with no vested interest in the outcome of this research project. She began the evaluation process in April 2007 by interviewing the project designer, with who she completed the designers' questionnaire. On this occasion, they discussed the implementation of the participants' evaluation framework. Among the 37 people who participated in one or several workshops (participatory modelling and/or role-playing game session), it was necessary to select someone to be interviewed. The 3 stakeholders belonging to the core group (Figure 1) were considered as compulsory members of the sample. Among the participants in the participatory modelling process, 6 people were selected so as to keep a representative of every institution (Table 1). When there were several people from the same institution, the those who attended the largest number of meetings were selected. Among the participants in the role-playing game sessions, the two urban developers, and all the mayors were interviewed, and in the case of their delegates, only the ones in charge of urbanisation were kept on the list. Finally, 3 additional interviews were carried out with officers from

the District Council who were interested in widespread use of the RPG with other villages concerned with the same issue.

Between April and June 2006, the evaluator conducted 23 individual interviews lasting 30 minutes to 2 hours depending on the participants. The full interviews were transcribed as well as the debriefings of the RPG sessions to supply complete raw material for evaluation. The evaluation was also based on the analysis of the documents provided by the project team: mainly the canvass describing the steps followed while applying the companion modelling approach and a log book detailing day by day the chronicle of what has been done (people involved, date and place, type of meeting, inputs, outputs). Publications, technical reports, master thesis and other background documents were also provided by the research team.

Table 1. Institutions involved in the participatory modelling process and its evaluation

	Participants in the process	Participants interviewed for the evaluation	Group
Research team (INRA)	2	1	Core group
Agriculture and Forestry Service	1	1	
Nîmes Métropole urban community	2	2	
Urban planner	2	1	Working group
Farmers' Association	2	1	
Private Forest Service	1	1	
Public Forest Service	3	1	
Land tenure Service	1	1	
Livestock Extension Service	1	0	
Fire Brigade	1	1	
Urban developer	2	2	RPG group
Mayors or their delegates	14	8	
District Council	5	3	Extra group
Total	**37**	**23**	

The interviews allowed to identify several decision-making related to the participation of mayors in the RPG sessions. One mayor decided to modify his urban planning project by integrating fallow lands nearby the urban area. Among the 5 mayors that the Agriculture and Forestry Service get in touch with to encourage them to develop a fire prevention plan, 4 get involved. They mentioned the RPG helped them to better understand the interest of such a plan and provided them useful information to discuss and argue its implementation.

7. DISCUSSION

Interviews carried out for the evaluation process showed that this new way of working was welcomed by most participants. All participants enjoyed the participatory, interactive and constructive aspects of this original way of working. During the participatory modelling workshops, they particularly enjoy the fact that everyone had the opportunity to give one's opinion and had time to understand the viewpoints of the other participants. The role-playing game sessions were considered by participants as catalysts having strengthened or accelerated the development of social relationships between them. For example, the officer from the regional Agriculture and Forestry Service realized during the sessions of role-playing that the urban developers were interested in fire prevention issues and felt really involved. That gave her the idea to include them as partners in new projects focused on this topic, such as the design of leaflets on fire prevention techniques at the interface between natural areas and urban zones.

Participants came to understand that it was necessary to manage in a genuinely cooperative way the constraints related to urbanisation and fire prevention, and that it was essential to pay more attention to vegetation dynamics on arable lands and natural areas. For example, some mayors discovered the possibility and the value of creating recreation areas as a buffer between the forest and urban zones. Role-playing game sessions were also an opportunity to initiate a discussion about the concept of "intercommunality" to sensitise the mayors or their delegates to the need to develop serious cooperation between villages for fire prevention issues. Urban planners and developers involved in the participatory process gave assurances that their practices in urban development will now take better account of fire hazards in their new projects of urban or mixed development zones.

In accordance with situated action principles (Suchman, 1987) and organizational learning theory (Argyris and Schon, 1996), participatory modelling led to a collective consciousness of the impact of fallow land encroachment near to urban zones on urban development dynamics and its impact on fire hazard dynamics. Discussions after playing with the model made clear the necessity of collective reflection on fire prevention procedures and urban development planning. But it also brought to light the need for specific financial support for integrating these costly devices to the new urban development projects, both in terms of initial investment and maintenance cost.

ACKNOWLEDGEMENTS

This work was carried out with the financial support of Nîmes Metropole Community and the Gard Forest and Agriculture Service, and the « ANR- Agence Nationale de la Recherche - The French National Research Agency » under the « Programme Agriculture et Développement Durable », project « ANR-05-PADD-007, Commod ». The authors also want to acknowledge all the participants to the role-playing game sessions.

REFERENCES

Argyris, C. and Schon, D., Organizational learning II. Theory, method, and practice. Reading Addison-Wesley, 1996.

Bourgeois, M., NîmetPasLeFeu: un outil d'aide à la decision pour la prevention des incendies dans le milieu périurbain nîmois. Master degree, AgroParisTech, 2006.

Bousquet, F., Barreteau, O., d'Aquino, P., Etienne, M., Boissau, S., Aubert, S., Le Page, C., Babin, D. and Castella, J.C., Multi-agent systems and role games: an approach for ecosystem co-management. In Janssen M (Ed), Complexity and ecosystem management: the theory and practice of multi-agent approaches, Elgar Publishers, Northampton, 248-285., 2002.

Collectif ComMod, Modélisation d'accompagnement. In *Modélisation et simulation multi-agents : applications aux sciences de l'homme et de la société*, Amblard, F. et Phan, D. (eds), Hermès Sciences, Londres, 217-228, 2006.

Collectif ComMod, La posture d'accompagnement des processus de prise de décision. Submitted to Nature, Science et Société, 2008.

Etienne, M., SYLVOPAST a multiple target role-playing game to assess negotiation processes in silvopastoral management planning. *Journal of Artificial Societies & Social Simulations 6(2)* http://jasss.soc.surrey.ac.uk/6/2/5.html, 2003.

Etienne, M., Companion modelling: a tool for dialogue and concertation in biosphere reserves. UNESCO-MAB, Paris, *Technical Notes 1, 44-52*, 2006.

Etienne, M., Du Toit, D., Pollard, S., ARDI: a co-construction method for participatory modelling in natural resources management. IEMSS Congress, Barcelona, 2008.

Mathevet, R., Le Page, C., Etienne, M., Lefebvre, G., Poulin, B., Gigot, G., Proréol, S. and Mauchamp, A., ButorStar: a role-playing game for collective awareness of wise reedbed use. *Simulation and Gaming 38(2)*, 233-262, 2007.

Suchman, L., Plans and situated actions: the problem of human-machine communication. Cambridge University Press, 1987.

Weber J. (1995). Gestion des ressources renouvelables : fondements théoriques, Cirad: 21 p. http://cormas.cirad.fr/pdf/green.pdf

Opportunities offered by the Design and the Use of a Gridded Multi-Scale European Soil Information System in Support of European Soil Policy

M. Van Liedekerke, P. Panagos
Joint Research Centre of the European Commission , Ispra, Italy
(marc.van-liedekerke@jrc.it)

Abstract: During the last years the need for a coherent approach to soil protection has come on the political agenda in Europe and in this context of a European Soil Thematic Strategy (STS), the European policy makers require access to European soil data to assess the state of soils at European level. The creation of European soil datasets is not new but has always been the result of a complex and time-consuming undertaking, the best example being the development of the European Soil Database. In the light of updates to existing soil data and of collection of new soil data simplification is needed through a suitable technical framework. Recently, the INSPIRE initiative (Infrastructure for Spatial Information in Europe) has embarked on a common framework for spatial data in the EU. One pillar is to conduct reporting and analysis of environmental information on the basis of a hierarchical system of grids. This system constitutes a suitable framework for the building of a nested system of soil data and could lay the basis for a Multi-scale European Soil Information System (MEUSIS). In order to achieve this, a common standard for the collection of harmonized soil information is to be developed and implemented. MEUSIS, aiming to be the harmonized soil information system for Europe which potentially could streamline better the flow of information from the data producer at a local or regional scale to the data users at wider scales, will offer a number of opportunities. The parties involved in the use of such a system are the data contributors, the data users (scientists and policy makers) and the data managers, who generally speaking will only benefit from such an approach.

Keywords: grid, European reference grid, INSPIRE, soil data, ESDAC

1. BACKGROUND

During the last years the need for a coherent approach to soil protection has come on the political agenda in Europe; it was introduced as one of the thematic strategies within the Community's 6th Environment Action Programme. In this context of a European Soil Thematic Strategy (STS), the European policy makers require access to European soil data and information of various types to assess the state of soils at European level. Also, as part of the newly proposed Soil Framework Directive (SFD), Member States would need to delineate and communicate to the European Commission so-called "risk areas" which are areas at risk to major soil threats (soil erosion, lack of organic matter, etc.). The way to do this practically would need to be discussed between Member States and EC once the Directive will be into force.

As part of this need to collect and assess soil data and information, the European Commission (EC) and the European Environment Agency (EEA) decided to establish a European Soil Data Centre (ESDAC), located at the EC's Joint Research Centre, as on of ten environmental data centres in Europe. Each environmental data centre acts as the primary data contact point for the EC's DG ENV in order to fulfill its information needs. ESDAC operates under the scheme illustrated by figure 1.

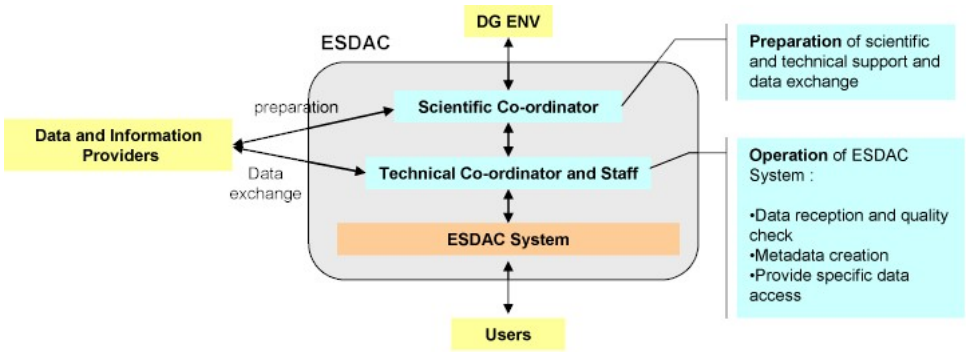


Figure 1. Operational ESDAC

The sources of soil information that currently reside at the ESDAC are JRC in-house and commissioned soil research activities, results from activities within the European Soil Bureau Network (the major scientific network for soil in Europe), results from EU funded soil related projects and results from collaborations with other organizations in the area of soils (e.g. EuroGeoSurveys). In the future, ESDAC would also need to prepare for the receipt, processing and making available reporting data coming from Member States in the context of the SFD.

In the past the JRC, mainly in its capacity of secretary to the ESBN, played a crucial role in the development of European soil datasets. The European Soil Database (ESDB), covering the EU27 has been developed jointly with European partners and is the only harmonized coverage of digital soil information for Europe [ESDB]. It is the result of a complex and time-consuming undertaking, due to the vast heterogeneity of soil data in countries. In the light of updates to such a database and of collecting other data in relation to the STS, simplification is needed through a suitable technical framework.

2. INSPIRE

The INSPIRE Directive [INSPIRE], aiming at the establishment of an Infrastructure for Spatial Information in the European Community (INSPIRE), entered into force in May 2007. This directive recognizes that the general situation on spatial information for environmental purposes in Europe is one of fragmentation of datasets and sources, gaps in availability, lack of harmonization between datasets at different geographical scales and duplication of information collection. The initiative intends to trigger the creation of a European spatial information infrastructure that delivers to the users integrated spatial information services. These services should allow the users to identify and access spatial or geographical information from a wide range of sources, from the local level to the global level, in an inter-operable way for a variety of uses. Policy-makers at European and national level are among the main targeted users who would need access to a number of services that include the visualization of information layers, overlay of information from different sources, spatial and temporal analysis, etc.

INSPIRE and soil data

Soil data are regarded as spatial data in the INSPIRE Directive (Annex III) and therefore Member States have to take them into account when setting up or adapting their national spatial data infrastructures. A common European data specification for soil will need to be set up in order to make data interoperability between soil data services possible. Given the heterogeneity of soil data models in Europe, as revealed by the reports [xxx] and [xxx], and more recently by work in the ENVASSO project [ENVASSO], together with the experiences within the European scientific soil data community during the set-up of the European Soil Database, it will be necessary to simplify significantly the way in which soil data are going to be represented.

INSPIRE and ESDAC

The INSPIRE Directive addresses Member States only and it is their task to come to a common position concerning soil data through collaborative work between so-called Legally Mandated Organizations (LMO) and the soil Spatial Data Interest Community (SDIC). The European Commission, and therefore the ESDAC, is not formally required to take part in this process. However, if ESDAC wants to consolidate its role as a compiler and provider of European datasets, it is only natural that it will interact with soil data players at Member State level and try to influence any decisions on European soil data representation.

INSPIRE and grids

During the preparation of the INSPIRE Directive, it was already realized that grids (or rasters) could offer tremendous opportunities to the representation of spatial data. Following the recommendations by the EEA and the INSPIRE Implementing Strategies Working Group, a European-wide reference grid (or raster) should be devised and adopted to facilitate the management and analyses of spatial information for a variety of applications; therefore, the JRC organized the "1st Workshop on European Reference Grids" [EUROGRID], inviting leading experts of different communities representing users of European grids. A grid for representing thematic information is a system of regular and geo-referenced cells, with a specified shape and size, and an associated property. A common position was reached and a "standard" technical solution was identified to be recommended for future adoption. It was also agreed to test this proposed solution in a number of initiatives. The selected grid would be used in spatial analyses, mapping and for the reporting of other sources of information to a common grid. The main recommendations from this Workshop to the European Commission were:

- to adopt a common geodetic datum (ETRS89) and coordinate reference system (ETRS89 Lambert Azimuthal Equal Area) for reporting, statistical analysis and display;
- to promote the wider use of these standards within all member states and internationally, by appropriate means (recommendations, official statement, …);
- to adopt a common European Grid Reference System for Reporting and Statistical Analysis. The system should be able to store regular grids and should be designed as reference for future Grids related to European territory; the System must satisfy a number of key principles, the most important being: easy to manipulate, hierarchical, be based on a Unified European Grid Coding System, based on units of equal area, adopt ETRS-LAEA;
- to encourage future European projects to make use of a standard European grid.

INSPIRE: the proposed grid.

The European Grid Reference System consists of hierarchically nested grids of regular and geo-referenced cells, with a specified shape and size. The system is based on ETRS-LAEA and its projection is centered on the common point (N 52°, E 10°); figure 2 illustrates the projection and extent; the coordinate system is metric and the sizes of the cells are 100km, 50km, 25 km, 1km, 500m, 250m 100m, etc.



Figure 2. Projection and extent of the ETRS_LAEA based grid system

At a specific level of the hierarchy, the cell is uniquely coded; e.g. the point with real world coordinates (5780354, 6436122) could be coded as 5780_6436 at the 1-km-level and as 578_643 at the 10-km-level. The nested nature of the grid system is illustrated in figure 3.

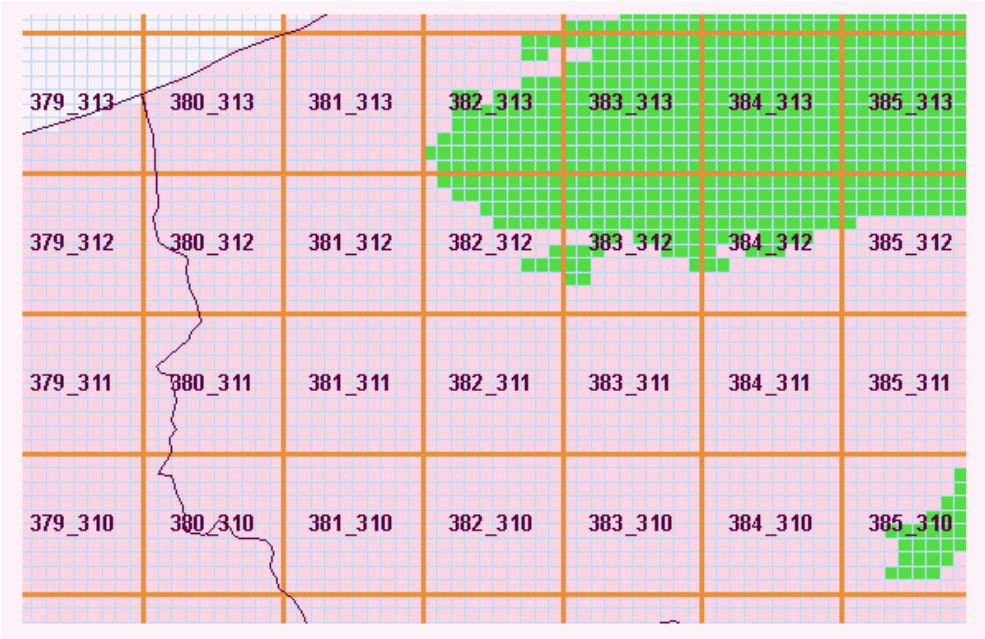


Figure 3. Nested grids.

Taking into consideration the recommendation to promote the wider use of such grids and in the context of dissemination of its soil data, the JRC European Soil Portal (eusoils.jrc.it) has taken the initiative of offering its European Soil Database as a library of rasters, one for each soil property [ESDB_RL]. Figure 4 illustrates one soil property, first as a European map, then as a detail for Spain, which reveals the 1km resolution.

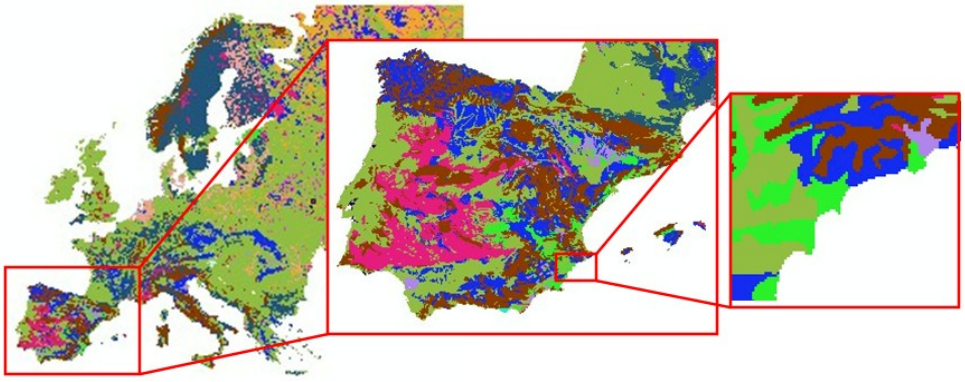


Figure 4. The “Limitation for Agricultural Use” soil property for Europe, as a raster

The idea of gridding has also been experimented with in a set of initiatives, linked to JRC: the ECALP project, the MEUSIS_SK project and the SIAS project.

3. ECALP: SOIL INFORMATION SYSTEM FOR THE ALPINE TERRITORY.

The objective of the ECALP project (2004-2006) [ECALP] was to create a system for gathering, and processing soil data for the Alpine Teritory, creating a so-called Alpine Ecopedological Information System to provide information to operational end users, land managers at different levels and soil scientists together with experts of other disciplines (in particular those using soil information for environmental and agricultural applications). The project participants came from the Alpine countries and formed a network of regional and national institutions which own and manage soil data. Particular attention was given to participants that would provide soil data close to national frontiers, as one could expect harmonization issues with trans-national boundary areas. It was decided to experiment with an INSPIRE-based 1km-cell (or pixel) grid approach as paradigm for data collection. A number of trans-boundary pilot areas of around 200 km2 (figure 5a) were defined, for which partners had to provide specific soil information in a common format for data exchange that was agreed beforehand. This format incorporated main soil and environmental features such as physiography, land use, topsoil and subsoil texture, drainage, rootable depth, parent material and soil type. Part of the exchange format was dedicated to metadata, i.e. information about data sources and collection and evaluation procedures used by each partner. Partners had to interpret and convert part of their traditional vector-based databases into the agreed raster format, as illustrated in fig 5b. This was not always a straightforward process due to the fixed spatial scale, potential loss of information and vector-to-raster issues. Also cross-boundary pixels needed special agreements on how to fill.

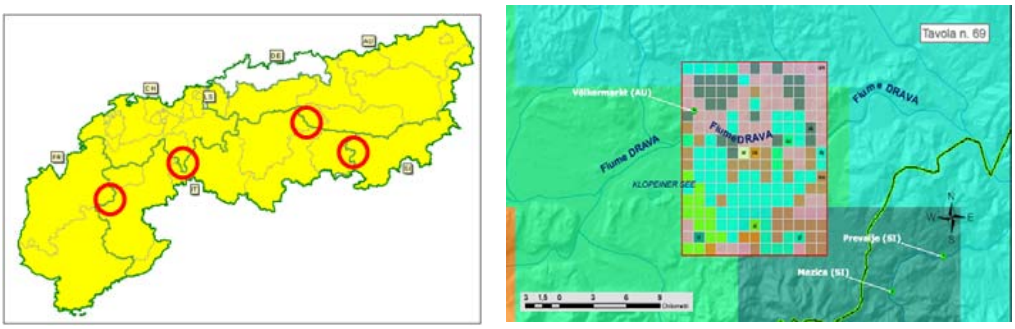


Figure 5.(a) ECALP Pilot areas, and (b)soil map example

4. MEUSIS_SK: A MEUSIS APPROACH FOR SLOVAKIA

During 2005-2006 the Soil Science and Conservation Research Institute from Bratislava conducted on behalf of the European Commission a study [MEUSIS_SK] that applied techniques which are foreseen in a Multi-scale European Soil Information System. In essence, the requirements were that the same soil data were to be provided for grid cells with different resolution, for the pilot areas of Slovakia indicated in figure 6. At 10km resolution for the whole of Slovakia (blue cells), at 5km for a part of Slovakia (green cells) and 1km for an even smaller area of 4000km2 (red cells). Freedom was given as to which data sources would be applied at a particular resolution.

Data sources for the pilot areas are represented by (i) primary geo- referenced data sources, used directly for pixel description (primary spatial information on the delineated soil bodies, e.g. soil map of Slovakia), (ii) secondary geo-referenced data, as a source of additional information for the primary geo-referenced data geographical (e.g. database of soil ecological units) and (iii) geo-referenced soil profile data used mostly for morphological and analytical information on the primary geo-referenced data
soil bodies. This project encountered no difficulties as only one party was involved in the set-up of the soil maps.

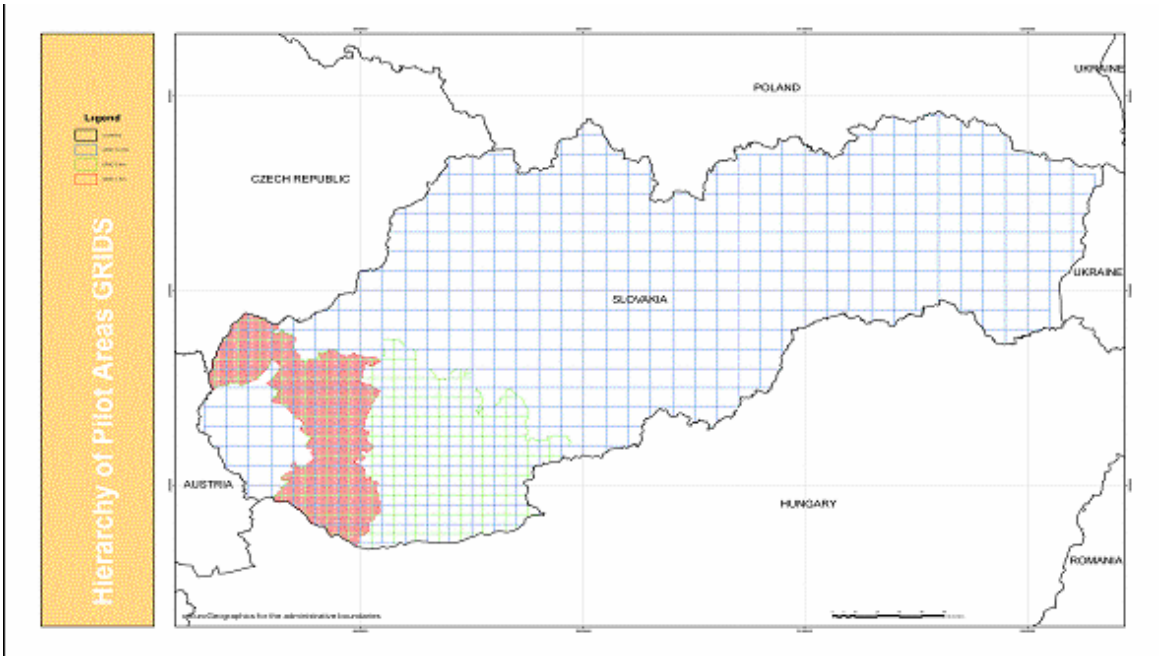


Figure 6. Overview of the MEUSIS_SK Slovakian pilot areas

5. MEUSIS IN ITALY (SIAS PROJECT)

The JRC is currently participating, in collaboration with the Italian Environment Agency (APAT), in a pilot project that seeks to develop an Italian Soil Database for soil erosion and soil organic matter according to INSPIRE principles and following a grid-based methodology with participatory contributions from Italian regional offices that are, according to Italian law, are the hold the soil [SIAS] . This approach exploits the possibility to transfer soil data and expertise which are available at local level to, in a first instance, the national level (APAT) and, in a second instance, to the European level (ESDAC). Again, as for ECALP, this happens according to agreed formats, but this time special attention is given to shared data quality indicators, both as quantitative indexes of data availability and specific confidence levels. The project involves all the 21 regions as data providers with different backgrounds and expertise; a priori, the data quality seemed to vary heavily from region to region and a minimum of harmonization was required. Therefore, it took a considerable amount of time to find a compromise for a suitable Data Exchange Format (available classes for specific properties, methodologies to derive properties, data typing, etc.). This project is still under development and the first results are expected by the end of 2008.

6. MEUSIS, THE CONCEPT

The INSPIRE grid system proposal could lay the basis for a Multi-scale European Soil Information System (MEUSIS), a system whereby soil data produced at a certain scale can easily be integrated or compared with soil data produced at another scale, provided that the rules for representation of the data are equal at all scales. From the local level (at high resolution, e.g. 100m), through the regional and country level (at medium resolution, e.g. 1km) up to the European level (at coars resolution, e.g. 5 or 10km).

In order to achieve this, a common standard for the collection of harmonized soil information must be developed. As some experience has shown, this is not an easy task and requires really the careful consideration and views of all players involved in the data contribution process.

In order to provide soil data in grid format, data providers will likely need to process their original soil data, held in traditional vector-based soil databases, in order to fit a grid, and again experience shows that this could lead to misinterpretations or even errors.

The MEUSIS concept offers a number of opportunities. Once data have been communicated in the form of grids, updates will be facilitated since it requires only the communication of data values for specific cells. MEUSIS also provides a suitable structure so that coherent and complementary data, available at various levels of a nested set of geographical scales, can fit together. Finally MEUSIS will also integrate into a comprehensive monitoring- and reporting system with different layers of various themes governed by INSPIRE principles.

MEUSIS aims to be the harmonized soil information system for Europe which will streamline better the flow of information from the data producer at a local or regional scale to the data users at wider scales (National, European and Global scales), all at the service of various levels of policy-makers, not the least environmental policy-makers at the European level.

MEUSIS: the way ahead
The SIAS project is to our knowledge the first in its kind in Europe that derives data for an INSPIRE theme at a country level from data at regional level, following a technical approach for reporting that was suggested by INSPIRE itself. In that respect, we are eagerly waiting for the results of this exercise which will be available by the end of 2008. It is the ambition of the ESDAC to set up a similar scheme which will let data flow from country level to European level in a similar fashion. For that, with or without European soil legislation that will oblige Member States to provide the Commission with soil "risk area" data, many financial and human resources are to be mobilized to act in a suitable technical framework. This is not the case yet, but soon when Member States will be faced with the transposition and implementation of the INSPIRE Directive into their national legislation, activities will speed up.

7. MEUSIS EVALUATION: DESIGN VERSUS USE

The parties involved in the use of a system such as MEUSIS are the data contributors, the data users (scientists and policy makers) and the data managers.

For the data contributors, there is an effort to be made in order to transpose their local data to an agreed format that structurally and semantically could be very different; this format could be imposed rigidly by the designers of the system but will receive a better acceptance if it is a trade-off between all involved parties. The partitioning of the geographical areas in fixed cells and the association of soil property values to such cells implies that data contributors, in order to deliver their geo-referenced soil data, only need to transmit data tables in which each record is linked to a cell identifier.
It is envisaged that the primary end-user of such a system will be the (European) policy-maker who needs, at all times, to be able to consult the state of European soils. The data of such a gridded system are easily presented and are difficult to misinterpret. Moreover, such gridded data can easily be overlaid with data from other themes, similarly gridded and often necessary to conduct integrated studies or to run environmental models. The other category of end-users, scientists, often at the service of policy-makers, benefit from the same advantages.
From the data management point of view, the grid approach offers many advantages, the most important being the utmost simplification of the geometrical model that needs to be dealt with; as a matter of fact, data management boils down to the proper organization of delivered data values within database tables.

On a whole, the grid approach will not contribute directly to the quality of final decision-making but needs to be seen as a tool that has a facilitating role in the overall decision-

making process that includes data contribution, data use and data management alke: the relatively simple underlying data structure allows all participating partners to interoperate and communicate efficiently and effectively.

Does the design and use of such a system fit the needs of its stakeholders? This is very difficult to answer a priori but there are many signs that such grid architecture is appropriate for all, relying on the limited but encouraging experience in the set-up of prototypes mentioned in this text.

8. CONCLUSIONS

In the context of various developments that take place at European level (i.c. the coming into force of the INSPIRE Directive with includes suggestions for data reporting, the establishment of a European Soil Thematic Strategy which includes the proposal for a Soil Framework Directive and the set-up of the European Soil Data Centre mainly as a vehicle for serving the European policy-maker with relevant soil data), the technical approach for setting up a Multi-scale European Soil Information System, based on the INSPIRE advocated grid approach for reporting, has a good potential to contribute to the efficiency and the effectiveness of the decision-making at European level in the field of soil, primarily because the proposed grid system is based on a relatively simple data structure that is easy to follow by all parties. A few examples from work at the JRC contribute to this statement and the outcome of other examples (like the Italian SIAS project) will hopefully confirm this.

REFERENCES

[ECALP]: http://eusoils.jrc.it/projects/alpsis/Ecalp_data.html

[ENVASSO]: http://www.envasso.com/

[ESDB]: “The European Soil Database distribution version 2.0, European Commission and the European Soil Bureau Network, CD-ROM, EUR 19945 EN, 2004"
(http://eusoils.jrc.it/ESDB_Archive/ESDB_Data_Distribution/ESDB_data.html)

[ESDB_RL]: “ESDBv2 Raster Library - a set of rasters derived from the European Soil Database distribution v2.0 (published by the European Commission and the European Soil Bureau Network, CD-ROM, EUR 19945 EN); Marc Van Liedekerke, Arwyn Jones, Panos Panagos ; 2006"
(http://eusoils.jrc.it/ESDB_Archive/ESDB_data_1k_raster_intro/ESDB_1k_raster_data_intro.html)

[EUROGRID]: “European Reference Grids Workshop, Ispra, October 2003, Proposal for a European Grid Coding System, Proceedings and Recommendations” (ed. A. Annoni) (eusoils.jrc.it/ESDB_Archive/ESDB_data_1k_raster_intro/annoni2005eurgrids.pdf)

[INSPIRE]: “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)”, published in the official Journal on the 25th April 2007, entering into force on the 15th May 2007. (http://www.ec-gis.org/inspire/)

[MEUSIS_1]: “MEUSIS: Multi-Scale European Soil Information System ”, Panagos, P., Van Liedekerke, M., Lado Rodriguez, L. and Montanarella, L. (2008). GEOconnexion International Magazine, Interoperable Geodata , Feb 2008 Volume 7· Issue 2: pp. 39-41.

[MEUSIS_2]: “MEUSIS: Towards a new Multi-scale European Soil Information System”. Panos Panagos, Marc Van Liedekerke, Nicola Filippi and Luca Montanarella. ECONGEO, 5th European Congress on Regional Geoscientific Cartography and Information Systems, Barcelona (Spain) June 13th-15th 2006. pp 175-177

[MEUSIS_SK] : http://eusoils.jrc.it/projects/Meusis/Slovakia.html

[SIAS] : http://eusoils.jrc.it/projects/Meusis/italy.html

[SFD]: http://ec.europa.eu/environment/soil/three_en.htm

[STS]: http://ec.europa.eu/environment/soil/index_en.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

Eururalis, a discussion support tool for rural Europe

W.A. Rienks[a], S.B. Hoek[a,] P.J.F.M. Verweij[a], R.M. Lokers[a] and B. Vanmeulebrouk[a]

[a] Alterra, Green World Research, Wageningen University and Research Centre, Wageningen, The Netherlands (willem.rienks@wur.nl)

Abstract: European rural areas have changed considerably in the past decades. Agriculture intensified whilst in other areas it marginalized. The rural landscape changed and it will continue to change in future. The forces driving that change are increasingly global by nature: demography, economic growth, climate change and international trade policies. The Eururalis 2.0 project assesses the combined impact of these driving forces. The results are published on a computer based tool to support discussions and to raise awareness of future changes in rural areas. The tool was made suitable for a wide audience, with policy makers as the main target group.

Keywords: Discussion support, globalization, rural area, impact assessment, land use change, tool, dissemination.

1. INTRODUCTION

What will happen to Europe's rural areas in the forthcoming years? What kind of threats as well as opportunities for socio-cultural, economic and ecological values can we expect? Can Europe's rural communities maintain their livelihood?

European rural areas have changed considerably in the past decades. Agriculture intensified whilst in other areas it marginalized. The rural landscape changed and it will continue to change in future driven by forces that are increasingly global by nature such as demography, economic growth, climate change and international trade policies. The Eururalis 2.0 project assesses the impacts of these combined driving forces and publishes them in a tool to support discussions and to raise awareness on the impacts of land use changes.

The discussion support tool developed within the Eururalis 2.0 project [Wageningen UR/MNP, 2007; Rienks, 2008, Eickhout, 2008] aims to help policy makers to obtain insight in the future development of rural Europe towards the year 2030. The longer term time horizon, the uncertainties about global trends, the huge spatial variety within Europe, let alone the various topics involved in discussions about rural areas make this a complex field. With the help of a series of multi scale models an extensive database was created. Only with the use of a computer based tool this database can be unlocked for the broader public.

This paper discusses how requirements for the tool have been gathered and managed: how end users and stakeholders were consulted and what happened with specific user feedback. Additionally it evaluates how the tool is disseminated on several international meetings with policy makers and scientists and in education programs.

2. COMPUTER BASED SUPPORT TOOLS

2.1. Introduction

In 1991, Sage established the concept of “the Decision Support System” or DSS as a source of advice for tackling management problems [McIntosh, 2007]. Around the same time, modellers started referring to the potential decision support functions of their product.

In the last few decades a large number and variety of decision support systems have been produced and published. The variety of the models can be classified according to the uncertainty and causality dimensions (Figure 1): ‘If uncertainty in the system and model is apparent, “what-if” type questions can be addressed. If uncertainties are small the probability of future events can be assessed. If causality of the model is prominent, a more systematic approach is possible. If causality is lacking, only regressive or deductive methods are available leading to projections or speculations of future events and conditions’ ([Becker, 1989] as quoted by [Van Ittersum, 1998]).

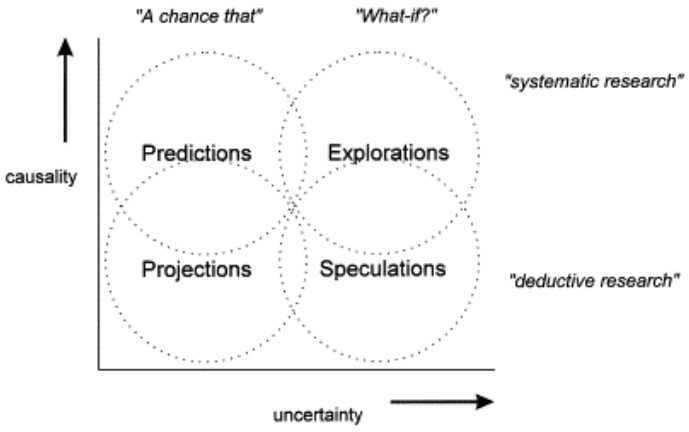


Figure 1: model classification (Becker, 1989)

2.2. Discussion support system

In the beginning of the nineties, the concept of policy life cycle was introduced: ‘Initially, there is no awareness of a problem. The first signs of a potential problem may come from the society (…). There is a period of discord and quarrel; people may even deny that there is a problem. After some time a common understanding eludes and governmental organisations begin to recognise that the problem exists. The political importance increases and policies and regulations are prepared’ ([Winsemius, 1990, Janssen, 2007] as quoted by Adriaanse [1997]). Subsequently, the policies and regulation have to be evaluated, eventually implemented and progress monitored at later stages.

The policy life cycle is also relevant for the policy making process aimed at changing land use. In each phase of such a process, different types of models are appropriate: projective, predictive and exploratory land use models [Van Ittersum, 1998]. In the first stage awareness is raised by exploring the problem. Within this stage - with people denying that there is a problem in the first place - there may be strongly biased discussions. In these discussions great uncertainty typically coincides with strong causality as well as with conflicting effects. E.g. in the field of rural areas discussions are about statements such as ‘biofuels contribute to a better environment’

or ‘trade liberalization is negative for rural areas and agricultural production in the EU’. Using the model classification from Figure 1, this type of problems is classified as explorative. Explorative tools are in particular useful in the phase during which awareness is raised, i.e. to facilitate the discussion.

2.3. Eururalis 2.0 tool

Eururalis 2.0 is dealing with the future of rural Europe and especially focuses on the way this future is influenced by global developments and strategic EU policies. The future of rural Europe is closely linked to current issues like:

- the enlargement of the EU internal market, since new Member States have joined
- ongoing liberalization of global trade
- reform of the European Common Agricultural Policy (CAP)
- climate change
- stimulation of bio-energy
- urbanization and infrastructural developments

As the future is uncertain and developments are increasingly hard to predict, Eururalis uses four contrasting future projections, or scenarios. They are based on the IPCC scenarios [Nakicenovic, 2000] and they represent uncertainties (1) in the way the world might develop and (2) in policy choices. Key driving forces are macro-economic growth, demographic- and political developments. The four future projections relate to different plausible developments defined by: (1) a global vs. a more regional approach to problems and strategies and (2) an open market vs. a higher level of intervention and regulation by governments.

Policy makers have a large variety of policy instruments which mostly aim at influencing the behaviour of relevant actors in society in relation to these issues. The Eururalis tool implements a subset of these instruments with a high effect on such behaviour: (i) market support; (ii) income support; (iii) ambition to stimulate bio fuel and (iv) stimulate less favoured areas.

The cause-effect relationship of both scenarios and policy instruments is implemented by a series of linked models in such a way that:

- the participants would be confirmed or rather confronted by the model results in their expectations;
- it would lead each of the participants to reflection; and
- it would fuel the debate on policy goals and instruments.

Scenario and policy effects are represented in the form of maps and graphs which show the variation of indicators within the EU at scales ranging from square kilometres to subregional level. Indicators are quantitative tools that synthesize or simplify relevant data relative to the state or evolution of certain phenomena. They are tools for communication, evaluation and decision making that can take quantitative as well as qualitative form depending on the purpose of the indicator [Gallopin, 1997]. Indicators concern the people, the planet and profit – the so-called 3Ps [Elkington, 1997]. A few additional indicators were categorized under the denominator “land use”.

3. DESIGN PROCESS

In this chapter the design process of Eururalis is explained. This process was a combined effort of scientific research institutes, software engineers as well as policy makers. They worked together to further develop Eururalis – from version 1.0 to version 2.0. In section 3.1 the start of the design process was described. Section 3.2 describes the interaction between policy makers, software engineers and scientists in the process of making choices about the content and the look and feel of the tool. Section 3.3 gives details on how the data were managed and on how the user interface of the tool was designed.

3.1. The start of the process

There are many ways to develop a software tool. The development process of Eururalis can be characterized as evolutionary development. Figure 2 shows the typical stages of such a process [McConnell, 1996]. The process was stakeholder-driven: the stakeholders involved were scientists, policy makers and other end-users.

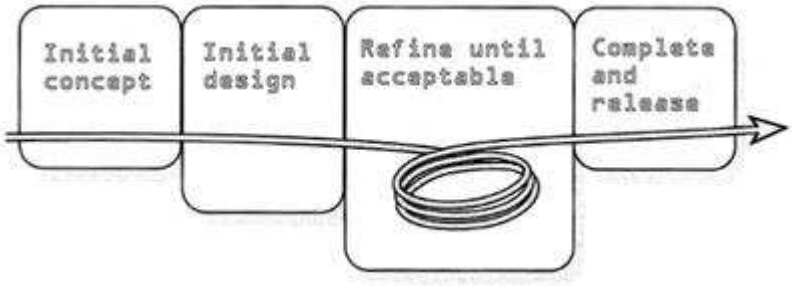


Figure 2: Evolutionary development model (derived from McConnell, 1996)

In figure 3 the design process of Eururalis is presented graphically; the typical stages described by McConell can be derived from this scheme.

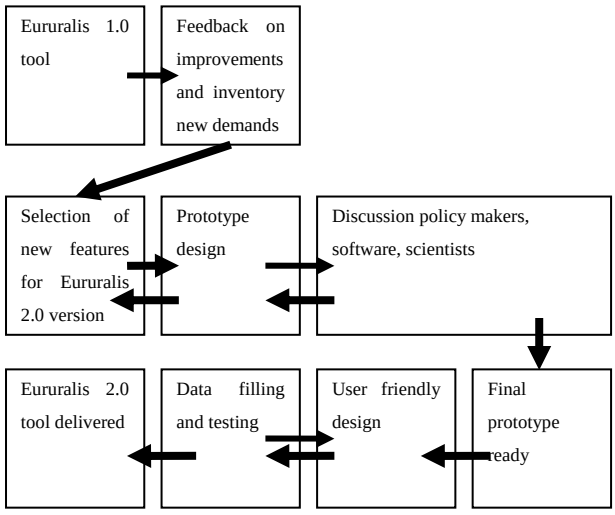


Figure 3: design process of the Eururalis 2.0 interactive tool

The initial phase of the design process started with the Eururalis 1.0 tool. The first Eururalis project was carried out in 2004. From the beginning, the tool was conceived as a discussion support tool [Klijn, 2005]. Since there was huge time pressure in the Eururalis 1.0 project, only little emphasis was given to the design and user friendliness of the tool. Potential end-

users were not involved in the design process, that was elaborated by software engineers in interaction with scientists only. Most effort in developing the Eururalis 1.0 tool was put in modelling and data gathering. Version 1.0 of the tool was released by the end of 2004 [WUR, 2004]. It was aimed to support discussions about the future of rural areas in Europe. This involved issues which go beyond the borders of single Member States.

The CD-ROM with the Eururalis 1.0 tool contained the mentioned scenarios, providing explorations as to how agriculture and the rural areas could develop towards the year 2030. Results were provided in the form of maps and graphs showing indicators. Users could indicate which indicator they wanted to study, for which time period and for which member state of the EU. This way, they could evaluate results by themselves. A summary of the conclusions and a number of fact sheets about important developments were added to the Eururalis 1.0 CD-ROM.

The Eururalis 1.0 tool was presented at various occasions, e.g. in the office of DG Agriculture in Brussels and during a meeting of the heads of the department responsible for rural areas (the directors) of the Member States organized during the Dutch presidency of the EU. During the latter meeting the Eururalis tool was installed on laptops of the directors of EU Member States to explore the model output during a two-day meeting. On the basis of this tool, a discussion was held about the future of rural Europe. The directors especially appreciated the possibility to employ the tool as a card index and visualize the output in land use maps. These features helped them to get an overview of the diversity in developments and interdependencies in rural areas both at national and EU level [Sterk, 2007].
After these meetings, the CD-ROM with the Eururalis 1.0 tool was widely distributed among a relative diverse audience by handing out copies to policy workers and scientists at congresses and other meetings. The project team of Eururalis received feedback from users – regarding both subject and technical matters.

3.2. Policy science interaction

According to Sterk [2007] Eururalis has not only played a heuristic role in creating awareness about future trends and threats concerning rural areas, but also an explicit community-creating role. The results of the tool served as a stepping stone for a discussion about the future of rural areas and functioned as a common language. Without the Eururalis tool the directors network for rural areas would fall apart. A third role Sterk mentions is the symbolic role for Eururalis (putting things on the agenda).

Halfway 2005, the Dutch Ministry of Agriculture, Nature and Food Quality commissioned the development of a new version Eururalis, i.e. version 2.0. Aim was to extend the existing model framework and the tool of Eururalis to support strategic policy discussions concerning the future of European rural areas (the EU-27). Target group were all those involved in policymaking on rural areas and agriculture at EU level as well as NGOs and scientists working in related fields.

The philosophy used for the further development of Eururalis was to preserve the good of Eururalis 1.0 and to improve the points which were indicated as weak. As strong points the following were mentioned: good balance between the 3Ps and land use, global context and common

language. A minor review of the four storylines was carried out. Integration of the used simulation models was improved.

Important points to elaborate on according to the consulted international policy makers and scientists (end-users) were:
- the implementation of interactive policy options
- development of a more balanced set of indicators
- the wish to look behind the results and know about the assumptions
- the wish to have results also at the level of EU regions.

The third point was often mentioned by scientists who put huge emphasis on the so called causal tracing of the results. The other points were raised by both policy makers and scientists.

The actual development of Eururalis 2.0 started by the end of 2005. The elaboration on the above-mentioned points were leading. Right in the beginning of the design process an interactive session was organized with both policy makers and scientists from several EU countries. Part of this session was to form smaller discussion groups and to make an inventory of wishes concerning interactivity, policy options and indicators in each of those groups. In this session it became more clear what kind of content policy makers wanted to have added to the tool. It also became clear which additional indicators they wished. In fact, a long list of wishes was obtained and it was immediately clear that not all could be realized given time constraints as well as technical constraints. The project team used these requirements to make an initial design of the content of the new Eururalis 2.0 tool.

It was clear that the requirements could not be presented in the framework of the Eururalis 1.0 tool. A complete redesign of the tool was needed to accommodate the extra demands. A first prototype of Eururalis 2.0 was developed and proposed to potential end users, i.e. policy makers. This group of policy makers from several Member States, the so called Policy Advisory Group, was consulted a number of times to give feedback on the consecutive prototypes. In those consultative meetings both content and technical design of the new tool were discussed. In this iterative process the Eururalis scientific project team, software engineers and policy makers gradually adjusted the tool and also adjusted their expectations of the tool. These meetings were organized because of the perceived importance of contextualization and network building among modelers, potential users and stakeholders for the acceptance of the tool and the models behind the data [Sterk, 2007].

3.3. Appearance and data management

The interactive use of results by the end-user was considered as very important. Results – i.e. maps and graphs - had to be available to the user instantly. Given the complexity and broadness of the field of rural areas and the complex multi model framework of Eururalis it was not feasible to calculate results on the fly within the tool. End-users specifically asked for an interactive tool. In order to guarantee needed real time results, a reference book approach (Wien, 2007) was opted for, with all modeled results built-in in a pre-cooked form. The data on the tool consists of 37 unique scenarios that show results for four time slices, for over 500 European regions that fully cover the 27 EU Member States, at 5 spatial scale levels and for 24 indicators in the people, planet, profit and land-use

domains. The results can be visualized in both maps and graphs. Fact sheets with interpretation of the results and general conclusions drawn by the project team are part of the tool.

The tool was designed in such a way that it has an appealing appearance with strong graphical content and an intuitive interface. The set-up of the tool is such, that the end-user is appealed to explore the impact of different scenarios and policies. The tool is equipped with controls by which the end-user can adjust scenarios and policy options and can view the effects on various indicators, in the form of zoomable maps and customizable graphs (Figure 4). Effects can be studied on different temporal and spatial scales and the users themselves can choose of which indicators they want to assess the impact.

With small user groups and amongst the research team, test sessions have been organized to remove flaws and to enhance the performance. Subsequently the dissemination of the tool was organized. The tool was packaged in such a way that it can be installed from a DVD as well as directly online (see http://www.eururalis.eu). It was made suitable for other operating systems than Windows alone.

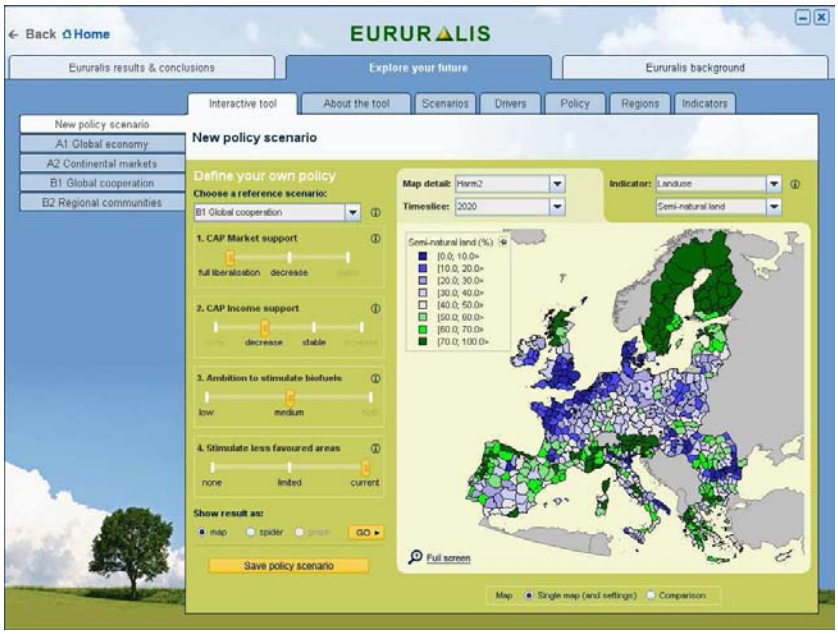


Figure 4: screen shot of the Eururalis application

4. IMPACTS OF EURURALIS 2.0

In this chapter, section 4.1 describes the dissemination process. Subsequently section 4.2 explains the spin-off of Eururalis 2.0.

4.1. Dissemination efforts

The Eururalis 2.0 tool is disseminated in various ways amongst the target groups of both the research community and strategic policy makers of the Member States, in Brussels and of NGOs.

Firstly at various places throughout Europe the results of Eururalis have been presented by the project team of Eururalis researchers. The audience varied from policy makers, consultants, NGO staff to scientific colleagues. Most presentations were structured in such a way that the main results of Eururalis were first presented orally and after which then the possibilities of the tool were exhibited. The audience and moderator interactively browsed the content of the tool. Sometimes a number of laptop computers were made available to let those present explore the results themselves; sometimes the audience was given the opportunity to ask questions for the moderator to answer interactively with the Eururalis tool.

Secondly some 3000 DVD copies have been printed and circulated at both scientific and policy oriented conferences and in the network of the Eururalis project team. The upcoming months further dissemination will take place. The DVD offers the user the possibility to explore the results of Eururalis.

Thirdly a number of sessions in education have been conducted especially about the future of rural areas and land-use change in Europe. Students were instructed how to use the tool and had to execute assignments such as analyses of changes of biodiversity or land-use in a specific scenario with help of the Eururalis tool.

Fourthly at the website www.eururalis.eu the interactive tool is made available. Potential users can download the tool there to their own computers. The first month of the launch of Eururalis 2.0 some 300 people used this opportunity. On the basis of e-mail contact resulting from the dissemination, it can be derived that users are mainly coming from research institutes.

4.2. Spin off

The goal of Eururalis 2.0 is formulated as to support discussions about the longer term future of rural areas in Europe. In chapter 2 it is mentioned that Eururalis has an explorative character focusing on issues with a possible conflicting causality and high uncertainty. The results of Eururalis should be regarded not as predictions of the future but as a means to provoke discussions. The following types of spin-off can be observed.

First, in the process of designing and filling the Eururalis tool, a network of scientists and policy makers already used Eururalis as a means of discussing with each other. The Eururalis 2.0 tool served as a common language to talk about and discuss complex and uncertain topics. Examples of such discussions are the meetings of the heads of the departments responsible for rural areas of the various Member States. The different scenarios of Eururalis provided flexibility in the sense that it questions ‘what are plausible changes in land-use’. It can accommodate diverse perceptions and can suit the explorative stage that the network of policy makers is in at the moment.

Second the results have been used by several scientific studies of third parties. Land-use results of Eururalis have for instance been overlayed with information of hotspots of agricultural birds and they have been used in workshops concerning the development of action plans for the Carpatian Mountain region. The results serve as a reference context for more targeted studies concerning nitrogen or forestry developments. For these

studies Eururalis unlocks the bandwidth of future trends at the global and EU level to which these studies can refer.

Third the results are used in vision development to explore what kind of opportunities and threats could happen in future. In Belgium the Eururalis land use results have been used in a forecast of spatial development. Another example is the use by the Standing Committee on Agricultural Research of the EU that used results from the Eururalis 2.0 to set out her vision for the future. Eururalis 2.0 is in the reference list of such projects. However given the explorative nature of the field Eururalis covers, the main contribution may be the impact it has in broadening the scope of thinking of the end-user and opening up discussions about the uncertainty of the future and the accommodation of different perceptions. This is something that is hard to measure, but needs further research.

5. CONCLUSIONS AND LESSONS LEARNT

The discussion support tool developed in the Eururalis project offers a flexible framework which proved to be very useful in discussions and in raising awareness on the possible future of rural Europe. A complex issue as the possible development of rural Europe involves many different dimensions, a lot of ways to observe the problem domain and various geographical and temporal scales. Discussion support systems like Eururalis are a powerful means to support discussions on such complex topics. The tool offers functionality to easily focus on the problem from different angles, like geography, policy option or type of indicator. As such it supports users in focusing on their specific interests, but also facilitates getting insight in related areas of interest.

The iterative development process used to design and implement the tool and the interaction between scientists and policy makers in this process has proven to be a valuable approach. It allowed the software designers to make the right design and clear implementation decisions, while at the same time increasing the focus on the specific requirements from the user groups. It can be concluded that the success of Eururalis and the use of Eururalis results is highly related to the efforts made on dissemination of the results. Dissemination was in fact already embedded in the development process. Already in this early stage a lot of effort was put into involving potential users into the development process and informing the outside world on the ongoing work in the project. These efforts were directed at both policy makers and scientists. These networking and content development activities were continued in the implementation phase, and thus the tool was successfully introduced into the user community

In practice it can be observed that the Eururalis tool and its results are used by policy makers, scientists and consultants. The results of Eururalis however, focus on the long term and are of an explorative nature. They generally do not lead to direct and concrete actions. It is therefore hard to objectively measure the actual use of results. On the other hand, it is clear that using the Eururalis tool supports discussion and thinking and increases the awareness on the possible future directions of rural Europe and the possibilities to influence this direction through policy.

A broad user community, ranging from policy maker to researcher, are a challenge for a project team to conceptualise and construct a usable product. The policy makers are concentrating on clearly distinguishable policy options and conclusions, while the researchers are mainly interested in the parameterisable modeling and methodologies behind the conclusions. The Eururalis tool is accompanied by booklets that each focus on the policy maker and researchers respectively: 1) an anthology based on the results of the Eururalis scenario study and 2) technical background and indicator documentation. Together with scientific and other publications, booklets, oral presentations and workshops the tool is a valuable product for disseminating Eururalis knowledge and raising awareness.

REFERENCES

Adriaanse, M., 1997. Tailor-made guidelines: a contradiction in terms? Proceedings of the second workshop Monitoring Tailor-made II, held in September 1996. Cf. http://www.mtm-conference.nl

Becker, H. A. and Dewulf, G., 1989. Terugkijken Op Toekomstonderzoek [Reviewing future research - in Dutch]. ISOR, University of Utrecht, The Netherlands.

Eickhout, B. and A.G. Prins, 2008. Eururalis 2.0. Technical background and indicator documentation. Wageningen University and Research and Netherlands Environmental Assessment Agency (MNP) Bilthoven. The Netherlands.

Elkington, J.: 1997, Cannibals with Forks - The Triple Bottom Line of 21st Century Business. Capstone Publishing Ltd., Oxford, UK.

Gallopin, G., 1997. Indicators and their use: information for decision-making. In: Moldan, B. and Billharz, S. (Editors). *Sustainability Indicators: Report of the Project on Indicators of Sustainable Development*. John Wiley and Sons Ltd.

Janssen,S.L.C., Verweij,P.J.F.M., Wien,J.J.F., 2007. Policy life cycle related tool development. In L.Oxley & D.Kulasari (eds.), MODSIM 2007 international congress on modeling and simulation, Australia and New Zealand

Klijn, J.A., Vullings, L.A.E., van de Berg, M., van Meijl, H., van Lammeren, R., van Rheenen, T., Tabeau, A.A., Veldkamp, A., Verburg, P.H., Westhoek, H., and Eickhout, B., 2005. The EURURALIS study: technical document. Alterra, Wageningen, Alterra report 1196. Cf. http://www.eururalis.eu/background.htm

McConnell, S., 1996. Rapid development - Taming Wild Software Schedules. Microsoft Press, Redmond WA, USA

McIntosh, B.F., Seaton, R.A.F., Jeffrey, P., 2007. Tools to think with? Towards understanding the use of computer-based support tools in policy relevant research. Environmental Modelling and Software 22, 640-648.

Nakicenovic et al., 2000. Special Report on Emission Scenarios. Intergovernmental Panel on Climate Change (IPCC), Geneva, Switzerland.

Nowicki, P., 2006. SCENAR 2020 – Scenario study on agriculture and the rural world. Directorate-General Agriculture and Rural Development. Cf. http://ec.europa.eu/agriculture/agrista/2006/scenar2020/final_report/scenar2020final.pdf

Petersen, J.E., Campling, P. and Gabrielsen, P., 2005. Agriculture and environment in EU-15 – the IRENA indicator report. European Environmental Agency, Copenhagen, Danmark. Cf. http://reports.eea.europa.eu/eea_report_2005_6/en

Rienks, W.A., 2008. The future of rural Europe. An anthology based on the results of the Eururalis 2.0 scenario study. Wageningen University Research and Netherlands Environmental Assessment Agency. Wageningen, The Netherlands.

Sterk, B., 2007. A window of opportunities – The contributions of land use modelling to societal learning. PhD thesis, Wageningen University, Wageningen, The Netherlands.

Van Ittersum, M.K., Rabbinge, R., Van Latesteijn, H.C., 1998. Exploratory land use studies and their role in strategic policy making. Agricultural Systems 58, 309-330.

Westhoek, H.J, van den Berg, M., Bakkes, J.A., 2006. Scenario development to explore the future of Europe's rural areas. Agriculture, ecosystems & environment, Volume 114, Issue 1, May 2006, 7-20.

Wien, J.J.F., Jansen, J.M.L., de Zeeuw, C.J. and Roosenschoon, O.R. Frameworks and interactive tools for Scientific Knowledge Systematization, MODSIM07 conference paper.

Winsemius, P., 1990. Guests in Our Own Home. McKinsey & Company, Amsterdam, 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

Multi criteria assessment – tool for integrated water management in Bereg landscape, Upper Tisza

Zs. Flachner[a]

[a] *Research Institute for Soil and Agrochemistry, Hungarian Academy of Sciences, H-1022, Budapest (flachner@rissac.hu)*

Abstract: In interdisciplinary projects – such as the Bereg landscape revitalization, Hungary Upper Tisza - one of the main challenge how to incorporate the different fields of interests, aspects. In the methodology - to reflect sustainability, as qualitative criteria - different domains (i.e. the economy, society and the population) need to be considered. The results from discourse-derived narratives and extensive modelling can be incorporated in the project alternatives' evaluation, reflecting long term scenarios generated (Scenes Project, 2007). Multi criteria assessment (MCA) is a methodology incorporates the different opinion of the interested parties to support decision making. The method allows to take into account more diverse public opinions and supports a better, transparent documentation of opinions, norms. In the Bereg project the improvement of the MCA method to reflect the specific aspects of areas with high flood risk and natural, social vulnerability took place with intention to improve the data collection and decision making process at regional and national level. Results were fed into the Bereg Interreg project (HUSKUA/05/01/139), MCA results has been accepted and inserted into the feasibility study of the flood reservoir alternatives. It allowed a better incorporation of the opinions of local SH's. Development of main pillars of criteria of the proposed assessment, data collection procedure and its difficulties and finally decision making results are presented.

Keywords: multi criteria assessment; integrated resource management, sustainable landscape development, stakeholder involvement, valuation of flood retention alternatives, decision support at different levels, Upper Tisza, Bereg landscape.

1. INTRODUCTION

Over the past 130 years dramatic changes had an effect on the land management of the Tisza river basin, Hungary due to massive canalisation [Flachner, 2006; Sendzimir et al, 2007]. To improve large scale grain production and navigation the river has been strengthened and surrounding areas (secondary floodplains) drained (90% of wetlands were lost, river shortened to 1/3). Due to several environmental, socio-economic reasons the revitalisation of these areas has been initiated at many micro regions along the Tisza river in the 1990's and followed by a governmental program called New Vásárhelyi Program (further on VTT) in 2000's [VTT, 2004] to introduce measures to reduce flood risk (such as flood polders in micro regions) and improve livelihood in the region. Key issues and relevant process in the Tisza valley are summarized in Table 1.

The Bereg landscape is a unique environment from several perspectives – as a border area to Ukraine it was not developed in the lat 50 years, traditional agriculture could remain for much longer time period than in other parts of the country. Its nature value has been recognized quite early – up to 40 % the area is under national protection; marches, peatbogs, special meadows with *Crex crex* are targets of national and international nature management work. In 2001 the flood destroyed 50% of the area, 6 out of 19 communities have been rebuilt, culture values restored (wooden churches from 17th century can be found in large numbers).

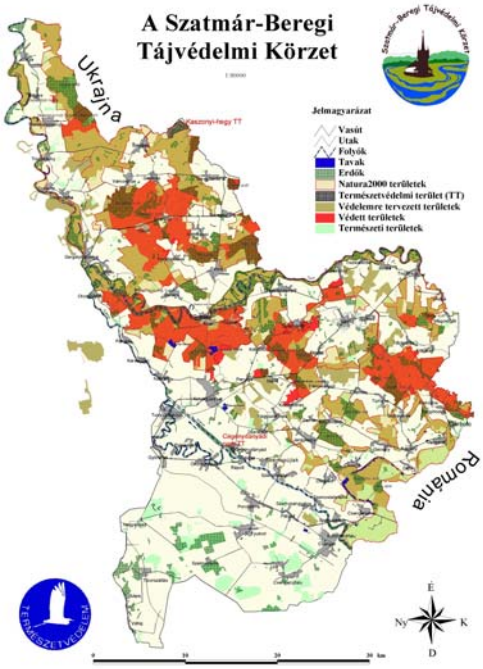
The Bereg landscape represents the Tisza catchments problems very well. It is part of the Szatmár –Bereg landscape protection area, situated in the north east part of the country (see Figure 1., red colour represents the nature protection sites). The issues listed in Table 1. cannot be solved with singe measure-based linear approaches on sustainable manner.

Table 1. The main issues at the Tisza catchments [Flachner, 2006]

Main issues	**Most relevant processes, problems**
Natural:	- Ecological decline, loss of biodiversity, fragmentation - increase of risks: flood, draught, invasive species, pests - soil degradation (texture, productivity) - stagnating water at large parcels - groundwater decline, pollution of resources(local, transboundary) - landscapes under different threats fragmentation, aesthetic loss, etc.
Social:	- ageing and migration from the region - increasing minority issues (gypsies) - high unemployment rate (avg. 30%. but up to 70%) - low education and awareness, loss of traditional knowledge - high values cultural values (built environment, traditions, local knowledge) under threat
Economic:	- poverty, segregation - land fragmentation, unclear ownership due to uncompleted land consolidation (LC), need for land use (LU) change - lack of financial capital, high cost of loans - lack of high quality, optimal scale machinery and technologies - lack of management capacity and co-operation

To find integrated, sustainable solutions to reduce or mitigate flood risk, contribute to nature protection and provide livelihood to the communities several methods and procedures should be incorporated, and complex criteria-system has to be applied in decision-making process.

Figure 1. Nature protection in Bereg - Szatmár landscape, Upper Tisza catchments [HNPI]



In Bereg landscape a community based strategy development and action setting process provided an excellent framework to collect ideas to build up the criteria system and define the long term perspectives [Bereg Strategy, FAO TCP project; Flachner, 2008]. Further on a cross border initiation resulted in an Interreg project for complex flood protection and floodplain revitalization in collaboration with Ukraine, where the objective was to develop authorized plans for polder developments in the light of the flood risk reduction program (VTT) objectives. The process of linking these elements and integrating it into an effective and efficient decision-making is described in the followings.

2. DEVELOPMENT OF CRITERIA SYSTEM

2. 1 Development of the Bereg strategy

To develop long term strategies a criteria-system reflecting the needs for long term, integrated rural-regional development need to be defined. The Bereg landscape strategy has been formulated in a participatory strategy development and regional planning process. In

three iterative cycles - involving different community groups, including minorities, policy makers and external advisors – the strategy and action plan have been developed in 1,5 years, but left opened for further improvement (so called 'living document' has been created). The strategy was incorporated into the LEADER (decentralized rural development under CAP) process and used as a background document for further discussion. Key elements of the strategy development process were:

- analysing the historic trends, processes (both quantitative and qualitative)
- understanding the casual links and select the important factors for common discussion for key stakeholder groups
- link the discussions to possible future scenarios (GEO4, IPCC) and derive key potential changes
- define principles and criteria and long term objectives
- select measures to implement in short term to gain benefits in short run.

In the process the most important identified floodplain development targets were:

- Increase the water storage capacity in the landscape (habitats, soil, deeper aquifers as well), and support the soil, nature, landscape rehabilitation processes, flood risk reduction and production safety at different locations.
- Contribute to lower risks of (external) threats at landscape, floodplain and region level – climate extremities, vulnerability in water, raw material, energy, food supply. Increase adaptive capacity of the system (social, economic, environmental).
- Decreasing GHG emission, and increasing carbon sequestration by appropriate land management (landuse and technologies).
- Support the rehabilitation and management of natural habitats of floodplains (oxbows, peat bogs, marshes).
- Contribute to maintain, preserve genetic variability of special indigenous species.
- Decrease, fix pollution load from point and non point sources.
- Support socio-economic development: efficient and economic land utilisation and production, high quality eco-products; new technology based alternative energy production and utilisation; tourism and recreation. Support market development both internally and externally.
- Support landscape and its heritage protection, which create a basis of improved life quality and livelihood in the miro-region and in the Upper Tisza river basin.
- Increase the population retentive capacity of the area and reduce poverty and social segregation.

2.1. Key procedural lessons of criteria settings

Figure 1 presents the applied process where the defined criteria system supports the selection of measures and incorporates differnet aspects.

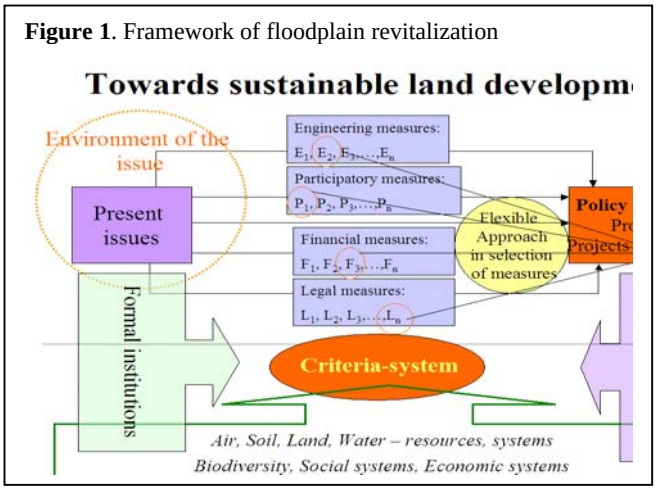


Figure 1. Framework of floodplain revitalization

The socio-economic systems have formal and informal structures which both have very high importance in the criteria setting in the Bereg landscape. Many communities had traditional roles, the flood defence and land management activities build up specific informal coalitions (eg. it worked very well during the evacuation in the 2001 flood, farmers were providing very efficient support to each other). Another very important informal criteria is the social acceptance of the proposed solutions by local society, which depends on e.g. the level of information provided concerning the VTT implementation; role and operation rules of the polders and expected benefits for the risk holders. Formal criteria is set by VTT law to designated flood-polder, where area

utilisation restrictions are set. The criteria-system can be clustered in the light of the main systems, such as:

- **Environmental**: efficiency, sustainable resource utilisation, revitalisation processes.
- **Social**: cultural, normative acceptance, embeddedness into individual, community aims. Legal basis of the system (connected to the norm structures). Institutional capacities, frameworks and procedures.
- **Economic**: planning, execution and maintenance cost; contribution to GDP, increase the income level and value added. Technical capacities define the system and solution boundaries.
- In the Interreg Bereg project these criteria were specified and detailed to represent the Bereg strategy and communicate the local needs to water managers responsible for authorized planning of the polder. (Besides national experts were incorporating the national expectations as well – it is presented in a table with bold.)

Table 2. Main clusters of criteria and methods used for Bereg revitalization

Cluster	Criteria	Method of criteria-setting
Environmental	**Water retention capacity of landscape** (**in flood risk**, in normal operation) - S **Landscape-river connectivity, structures -S** Soil conditions and agricultural productivity - S **Ecological value of habitats** (Natura 2000) S-M, revitalisation capacity and water dependence (water stress-index) S-M Duration and depth of water cover –S	Environmental impact assessment Field work, monitoring GIS assessments 1D-2D modelling (ARES) for flood risk, climate change impacts, operation options Landscape development history assessment (from 18th century)
Social	**Reduce risk and system vulnerability** (maximise regional, community, individual/private protection) -S Empowered land utilisation tradition, value added (floodplain management knowledge)- S-M-L **Inhibit not sustainable development strategies S** Joined responsibility and ownership, improved regional network and lobby power (S) -M Equity, transparency of system benefits S-M-L	Social discussion, forums Survey on values, expectations, knowledge Elaboration of alternatives of models, data and information gained from assessments Networking
Economic	**Lower implementation cost of constructions M** Lower long term maintenance cost L **Lower stock at risk** (e.g. changing locations of buildings, farms, infrastructures) M Maximised landscape production (re-parcelling, land use change, new products, services) S – M- L Realised environmental service cost (subsidies, payments, support) S-M	CBA of technical measures Land use structure assessment, Land Consolidation and Land Development planning Product cycle assessment of present and alternative LU CBA for services of LU

(with bold the national criteria is presented; L-long term, M-mid term, S- short term).

2.3. Need for quantifications – MCA in the polder development for Bereg landscape

The Interreg project main aim was to define and plan common water infrastructures in both Ukraine and Hungary to support the flood risk reduction and floodplain revitalization by shallow flooding [Sendzimir et al., 2008]. The planning process was build on different landscape simulation modelling as well as other complex assessment researches (e.g. soil development and degradation assessment; level and risk of inland water stagnation; ecological corridor development, historic land use change, terrain model (detailed digital elevation) and habitat development).

Finally three alternatives were developed for potential flood polder (up to 90 million m^3 water retention capacity) with different locations and measures to which decision-making process had to be defined, input information formulated. In October, 2007 the first study provided by the lead partner (FETIKÖVIZIG – Upper Tisza Water Authority) considered mainly the costs and flood risk reduction benefits; and none of the ecological, social concerns reflected in the *Bereg strategy* were incorporated. Since the project had to be finalized by 2008 February very short time left to implement a more complex assessment – building mostly on available data, maps, statistical information gathered in the Interreg project and in the FAO TCP project.

Due to the above described reasons multi criteria assessment (MCA) in combination with CBA was selected to use in the following steps:

1. Potential criteria from Bereg strategy were extracted.
2. Stakeholder assessment implemented to understand the role in the decisions (level of impact, level of decision making power)
3. Criteria were discussed and defined with key stakeholder-groups (farmers in the target area; water, forest, nature managers; mayors in the area, outside the area; regional, national decision makers).
4. First calculations were implemented to gain the indicators supporting the criteria for each alternatives (for simplification process 2 alternatives are selected)
5. Presenting the indicators for each alternatives and ask SH to provide values.
6. Statistical aggregation of values for each alternative, for each groups and special factoring for specific interest groups: most effected ones, highest decision power; highest competence.
7. Presenting the calculation results and to verify it with key SHs in the Bereg landscape.
8. Finalize the MCA, combine it with CBA and report to decision making panel for decision (based on the methodology suggested by Sijtsma).[Sijtsma, 2006]

Utilizing the knowledge gained in the mentioned projects and implementing (*most of the*) following steps the MCA assessment was performed (see Table 3). The process and its link to decision making is detailed in Chapter 4.

Table 3. Selected parameters for MCA assessment

Selected parameters for MCA assessment		
A. Safety	***B Socio - economic aspects***	***C. Environment and nature conservation (resource management)***
1. Village safety	1. Landscape management potential Conditions for (extensive) grazing	1. Soil management (structure, productivity)
2. Other infrastructures (roads, trains, channels)	2. Transparency (decisions, data)	2. Water management (balance, quality)
3. Human life protection	3. Networks (socio-economic)	3. Protection of landscape heritage
4. Safety of production (arable land, orchards)	4. Tourist potential (including secondary potentials)	4. Biodiversity, genetic diversity
5. Risk reduction effects on downstream and upstream	5. Game management	5. Water demand for ecology /habitats
6. Water storage capacity	6. Biomass production	6. Climate change effects – carbon sequestration, adaptation and mitigation

3. QUANTIFICATION OF CRITERIAS

The implementation of the process was not ideal, but as a first complex polder development in the frame of the VTT project has remarkable values and lessons to be shared. Due to lack of time the iteration with SHs were not fully comprehensive, but since experts worked in the area already for longer time (almost 2 years), the team could rely on the knowledge gained in indicator development and in participatory process before [Flachner, Németh, 2005]. The criteria development is performed by the author; debates and definition improvement has been conducted by key groups of different experts, local team members.

In the process of calculations several indicators had to be replaced by expert judgements, since data were not accessible (eg. agricultural subsidies per ha) or would be with huge delay (eg. number of game in the territory) available. In the followings some key quantification procedures are detailed to present the complexity of the decision making process.

3.1 Quantification of flood risk and polder alternatives

Dynamic modelling supported the MCA criteria definition, where expected climate change impacts on the landscape and the water regime were considered as well. A hydrologic - 1D-2D hydrodynamic model has been applied, the so-called ARES model [Koncsos, 2006].

The model was incorporating the specific inputs from the actual soil survey (e.g. soil water retention capacity), land use survey took into account specific nature conservation needs and restrictions as well. The modelling process was performed in 3 iterations – first round the potential locations, storage capacity and its impacts on local, regional flood risk reduction was calculated, several alternatives described.

Water management experts and local key stakeholders reduced these alternatives. The remaining 3 alternatives were further specified (max.- min. water level, total storage capacity, flood risk reduction capacity, max. – min. water flow and speed, water retention time, option for drainage of the territory) and modelled for base line conditions and in light of potential changes in the region: introduction of other polders in Ukraine and in Hungary; climate change impacts; sedimentation of the riverbed. The final iteration was based on specific local demands, replacement of infrastructures, protection of important infrastructures.

The model calculations provided very important indicators for the MCA assessment, such as:

- Size of effected area (potential for sustainable landscape management area) /km2/
- Water supply to soil moisture and to groundwater recharge (in non-flooding case, in flood risk case) m3/m2
- Water purification /N removal potential/m2/
- Support for biomass growth at the secondary floodplain /m3 timber/km2/
- Nutrient supply for the agricultural fields
- Size of build area, infrastructures under threat in flood risk and in shallow flooding case. /m2/km2/
- Water storage capacity /m3/
- Risk reduction in the region /cm/cm flood level/

The calculation of parameters is transparent, for different sub-alternatives (variants) slight differences could be exactly derived. Maps and tables were generated and these information shared in public consultation.

3.2 Quantification of soft indicators

The Berge strategy has identified the landscape (aesthetics) as a very important cultural heritage and resource for future development, especially taking tourism and hunting income opportunities into account. The water management agents (regional authority and planners) had difficulties to incorporate these aspects into the plans, especially these issues were not considered as a selection criteria for polder's water infrastructures. To tackle the issue landscape architects were involved to visualise locally and at landscape level the potential changes, difference between the alternatives. These alternatives were scored and valuated by local stakeholders and experts related to landscape management such as hunters, nature protection rangers, forest managers, tourist agencies. Different heights of dams of the polder and two different alternatives were assessed from the perspective of landscape to select the option, which has the least effects on the natural look of the Bereg.

The other important factor in the Bereg landscape is the high genetic variants of traditional fruit species and of course high biodiversity at protected sites. Since the habitats for these species are mainly water depended, 3 elevation categories were defined and the landscape classified based on the water- demand of habitats (including agricultural fields): from aquatic to drought resistant. These different plots were mapped, their potential water cover in each alternatives were analyzed, potential impacts assessed in the light of the water level and duration of water cover as well. These maps were turned into ranges of parameters and incorporated into the criteria system.

Water supply for marshes, wetlands, oxbows – contribution to nature protection and CC risk reduction and improvement of tourist potation of the landscape (fishing) has been based on previous surveys, and estimation of expected benefits preformed as expert judgement. Additionally survey on the ecosystem service assessment following the methodology proposed in the Millennium Assessment has started to quantify the services provided by the present system, the proposed system and a more intensified system, in the light of expected climate change as well [ADAM project report, 2008]. These results are

not complete yet, but it will have a link to the LU modelling and water balance modelling work performed in the catchments and will be supported with further public debate processes.

4. DECISION MAKING PROCESS IN BEREG INTERREG PROJECT

The alternatives developed for final decision making process are the followings:

a) Polder development on specified territory with embankments height of 2,5 meter minimum, up to 60-90 m3 water storage capacity. In case of no flood risk the areas water steering system is supported by a semi natural channel supporting water charge from the Tisza river to the further areas in the Bereg landscape.
b) Polder development on specified territory with embankments height of 2,5 meter minimum, up to 60-90 m3 water storage capacity with a permanent lake inside the polder for other tourist utilization possibility. . In case of no flood risk the areas water steering system is supported by a semi natural channel supporting water charge from the Tisza river to the further areas in the Bereg landscape.
c) Polder development with dynamic territory, following the natural elevation options, with embankments height of 1,2 meter minimum, up to 60 m3 water storage capacity with specific local dams (circular dams) around key build areas (villages, infrastructures, specific individual farms) including min 1,5 time larger territory, half of the water level (max . 80 cm in high flood case) compared to options a)-b). In this case the channel was not included.

4.1. Key procedural elements of decision making process

The Interreg project had 3 types of SH involved - those who are ***directly exposed*** to the decision (micro regional association of municipalities, representing the local SHs; nature directorate responsible for the 40% management of the territory); those who ***indirectly exposed*** (water authority responsible for water management in the Tisza river, including the flood risk reduction; civic organizations in the region) and those who are ***having responsibility*** to describe the situation, represent the national and EU directives (research institutes and higher policy bodies).

First of all the process of decision making had to be formulated. It had several rounds in the project team and out of the project team as well. Key considerations were:

- Does the micro regional association represents the local land owners, municipalities affected interest?
- Can the nature directorate harmonize the needs for livelihood and nature water demands?
- What level external institutes (such as the RISSAC) can take part in the decision making level?
- Which information level is satisfactory for directly exposed SHs - detailed plans including parcels or general plan which has the option to modify?
- Which extend modifications can be incorporated after making decisions?
- How the alternatives should be presented to SHs directly affected?

These questions were discussed and finally the team got to the following procedural conclusions (with consensus based decision):

- Different focus group meetings with **specific interest groups** to debate the benefits, the future costs and responsibilities of maintenance, the level of involvement in case of flooding, compensation measures to pay for requested land use change or potential loss of crops, income. These meetings were initiated by the civic organization to get a neutral attitude at the meetings, not having influence by higher-powered decision makers.
- **Public involved** covers directly and indirectly effected people in the micro region, including farmers, municipalities fishermen, foresters, hunters via public hearings where all concerns could be shared. The requirements mentioned at the hearing were incorporated into the MCA.
- **Municipalities were part of the pre-decision process,** where councils were discussing the alternatives, evaluating the local impacts, potential losses and

benefits with technical support provided by all planning partners, including research institutes and moderated by civic organization. Preliminary results of MCA presented and improved at the 4 most effected and concerned municipality. After having individual decisions a joint meeting for all municipality leaders were debating the alternatives, decisions were collected (alternatives reduced to a) and c) and finally joined pre-decision is set by the meeting to have alternative c) with main channel to support water steering.

- **Final decision making panel** consists of water authority, financial responsible ministry, agency for catastrophe management, micro region association, regional civic organization and nature directorate, they are responsible for decision making having considerations on all options, public hearings results, research institute assessments and recommendations and CBA, MCA valuation.
- **Final decision is made by the director of water authority ,** based on the discussion in the panel. The explanation of the process from the water authority side was the responsibility to apply for funding and cross harmonization with the Ukraine initiatives. (Embedded approach of decision making).

4.2. Selection of polder – specific issues of decision making

Following the process the local opinion was supporting the soft polder (option c).), while the planners and water managers were in favour of option a) /traditional polder/. The MCA had an important role in the final discussion driving the attention on the additional benefits, and long term concerns, including climate change impacts of water balance and livelihood improvement in general and for those participating in the polder implementation. (the integrated value was slightly higher for the "soft polder" than for the "traditional", while the Cost were increased by specific measures proposed by the water managers. Several measures were unnecessary but due to lack of time these discussions on final set of required measures and interventions are moved to the part of detailed planning.)

The soft /dynamic, more natural polder has been selected – but the work just starts now. Benefit transfer evaluation in the region and in the Upper Tisza catchments, the operational issues and other responsibilities for development of final plan is still under development, where further work to specify certain values of MCA parameters are turned to be important. The maps presenting the values and the GIS system, which can integrate criteria and allow statistical assessments, are also in development. Involved local people have not been satisfied with the transparency of the decision making process, and - which is more important - the reasoning on the measures proposed for final plans (eg. Higher roads which could be effected by flood level water (probability is 30 year /100 years) instead of allowing the water cover for 1-2 days not blocking the municipality access for larger cars; or protection or even replacement of high voltage pillars when in general baseline conditions with inland water stagnation similar water covers could occur).

The security of continuing the public debate and building local coalition among key SHs is also weak – after having finished the projects in the region it is up to the local SHs to get involved and keep receiving information from the water authority, which is a very difficult process since there is no tradition for open planning in Hungary and especially in less developed regions. Besides the river basin management planning can provide a frame of further actions, share of concerns and evaluate the use and non use values of the landscape functions, processes. This way the institutionalization of the actions can take place, set of relatively good measures can be selected.

5. CONCLUSIONS

The presented example of criteria-system is drawing the attention on the importance of combine the local and national, long term and short term perspectives both in landscape and water management investments as well as in the frame of the RBMP under the WFD. The criteria setting for the Bereg area had supported the development of a long-term infrastructure and land development program, which will define the future of the landscape and its society. It is promising that the society takes the responsibility of being part in the process, discussing the results of engineers, researchers and try to build consensus if resources (both financial and institutional) are secured to continue the process started. The

MCA application with combination of CBA provides new lessons for other floodplains as well, contribute a more sustainable development of the Tisza valley.

ACKNOWLEDGEMENTS

The author wish to thank László Koncsos (Budapest Technical University) and Zsuzsanna Nagy (Corvinus University for their support, and the Bereg Interreg project (HUSKUA/05/01/139), the Bereg FAO TCP project partners to support data input and criteria development. The research is linked to the SCENES project to develop long term water scenarios for the Tisza region.

REFERENCES

Carley, M., Christine, I.: Managing Sustainable development, Earthscan, London, 1997.

Delinger, K.: Land policies for Growth and Poverty reduction, World Bank and Oxford University Press, 2003.

Flachner, Zs., Németh, T.: The concept of the Methodology Development of the State of the Environement Assessment (SEA), MTA-KvVM, 2003.

Flachner, Zs.: Water retention based landuse changes at the Bodrogköz area – ecological processes and economic measures, Cereal Research Communications, Vol. 33., 2004.

Koncsos, L.: Modelling of water retention and flood risk in the Upper Tisza and in Bereg. Methods and results, FAO TCP project publication, MTVSZ publication, 2006.

Sendzimir, J. Magnuszewski, M., Flachner, Zs., Balogh, P., Molnar, G., Sarvari, A., Nagy, Zs.: Assessing the resilience of a river management regime: informal learning in a shadow network in the Tisza river basin. Ecology and Society, in press (2008)

Sijtsma, F.J.: Project evaluation, sustainability and accountability, REG-publication 27, Groningen, The Netherlands, 2006.

TRANSACT Consortium: Supporting decision making in transboundary catchment, VKI-CT2002-00124 Project Document, 2006.

Vásárhelyi törvény, 2004. LXVII., Magyar Közlöny (VTT law)

Werners, S. E., Z. Flachner, P. Matczak, and M. Falaleeva: Institutional Adaptation of Floodplain management to Climate Change in the Hungarian Tisza River basin. in Proceedings Earth System Governance, Amsterdam, 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
(Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

Sustainability Impact Assessment Tools (SIAT) for Regionalised European Impact Analysis: Focusing on the design process

S. Sieber[a], **P. Verweij**[b], **Jan-Erik Wien**[b], **D. Pohle**[a], **K. Fricke**[a], **K. Müller**[a], **C. Pacini**[c], **H. Haraldson**[d], **K. Tscherning**[a], **K. Helming**[a]

[a] *Leibniz-Centre for Agricultural Landscape Research e.V., Eberswalder Str 84, 15374 Müncheberg, Germany (stefan.sieber@zalf.de)*

[b] *Alterra, Postbus 47, 6700 AA Wageningen, The Netherlands (peter.verweij@wur.nl)*

[c] *Department of Agronomy and Land Management (DISAT), University of Florence, Piazzale delle Cascine 18, 50144 Firenze, Italia (gaiocesare.pacini@unifi.it)*

[d] *Dept of Chemical Engineering, Lund University, S-221 00 Lund, Sweden (hordurh@chemeng.lth.se)*

Keywords: Sustainability Impact Assessment, meta-model, SIAT, scenarios, simulation

1 Introduction

The SENSOR project aims at delivering ex-ante Sustainability Impact Assessment Tools (SIAT) to support decision making on policies related to multifunctional land use in European regions. Decision support tools are required that provide scientifically substantiated anticipations of the effects of future policy options on sustainability issues. SENSOR responds to these information needs at European land use policy level with the SIAT, which is interactively designed with end users (Helming et al. 2008). This paper illustrates the *design process* of the SIAT under the concurrent condition of a high usability for end users. To meet this goal, *end user involvements* by using techniques of *software prototyping* is applied.

Impact Assessment (IA) is an important instrument towards the fulfilment of the European Sustainable Development Strategy (EC 2001). This obligatory process has to be undertaken before decisions on policy proposals at European level are made (EC 2005). The European Commission provided in the European IA Guidelines detailed methodological steps on the procedure of IA (EC 2005). SIAT covers the two methodological steps of (a) analysing policy options against the divergence to defined objectives and (b) comparing policy options among each other.

IA procedures are prevalently supported by operational tools that are often restricted to precise, quantitative sector information. They focus mostly on single aspects of economic, social or environmental impacts and are mainly designed for ex-post analysis (Bartolomeo et al. 2004). Integrated and comprehensive questions are less answered (Tamborra 2002). SIAT intends to bridge this gap and thus focuses on impact assessment towards an integrated perspective. Region-explicit policy impacts are analysed across six sectors on economic, environmental and social indicators (Sieber et al. 2008, Verweij et al. 2006).

The major challenge is the transformation of interdisciplinary knowledge into an adequate *model-design* that meets the *end user requirements.* At the same time high standards of technical performance and evident functionality should be ensured: In particular, (a) short response time, (b) maximum flexibility regarding re-useability of applying scenarios, (c) fast integration of new intended policies for impact analysis, (d) high compatibility of different calculation methods, (e) effectual reliability and plausibility of simulation results.

The research question asks for a method to assure both, meeting the end user requirements as well as the above described standards. To meet these goals the method of software prototyping related to end user involvements is described in detail. Thus, the paper illustrates in chapter 2.1 the SIAT design and methodology, in chapter 2.2 the applied process of operational prototyping. Chapter 3 concludes in terms of so far achieved results.

2 Developing Model Design Features

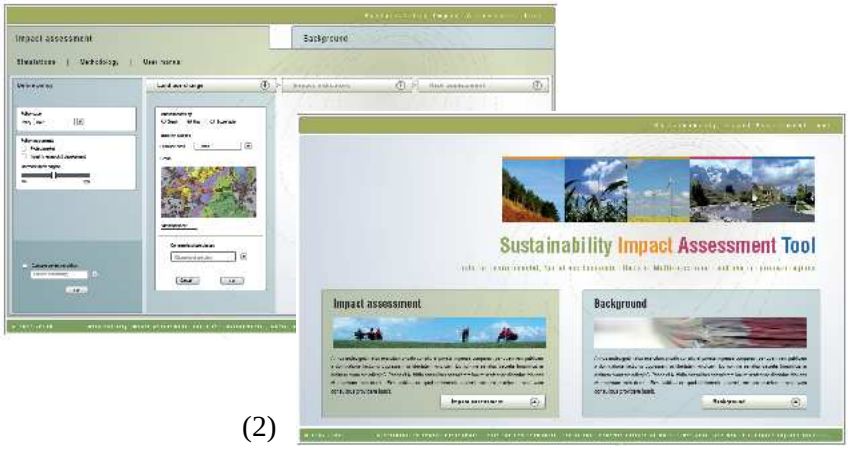
Software prototyping focuses on the design process towards the final model version, which should be tailored for the specific use of end user groups (Guida et al. 1999). Typically, subsets of features demonstrate functionalities of specific domains to be developed step-wise in collaborative discussions with interdisciplinary researcher (Davis 1992).

The SIAT prototype I was released in 2006 (Verweij et al. 2006). This version contains a small subset of the features to demonstrate the functionality of simulating land use-related policy impacts. The domain structure and model components have been defined to allow users to evaluate the proposals on design and functionality. Group discussions as well as interviews with potential end users surveyed expectations and requirements for the operational prototyping (Davis 1992). The feedback has supported accurate estimates on given capacities for software specification and applied modelling techniques. The given budget for software development adjusts the design of the software architecture towards possible solutions.

As a result of theses processes, SIAT prototype II achieved a complete system integration consisting of a new Graphical User Interface (GUI), server-data base and newly implemented calculation components with related content management (e.g. fact and information sheets). The SIAT prototype II will be released by end of May 2008. In the following chapter 2.1 the methodology and functionality of SIAT is explained. Subsequent, chapter 2.2 emphasises the development process of user involvements and prototyping.

2.1 The SIAT design

The integrated analysis by means of SIAT allows policy scenario solving across the sectors agriculture, forestry, energy, transport, nature conservation and tourism (Sieber et al. 2007, Verweij et al. 2006). SIAT conducts ex-ante impact assessment for the target year 2025 and covers 570 European regions at the level of the EU 27. End users are able to simulate policies by changing the intensities, which are expressed in sets of instruments (e.g. agricultural support subsidies in million €). Each simulation computes 35 impact indicators that illustrate the policy impact. Trade-offs are shown, when sets of policy options with varied intensities can be compared among each other. The impact indicators contain critical limits to valuate sustainability. Land Use Functions indicate the level of goods and services at regional level (see paragraph *methodology*).



(1)

(2)

Fig. 1. Design of prototype II (Verweij 2008)

Functionality

On the opening page of the GUI the user decide to choose the pathway (a) conducting Impact Assessment or (b) reading background information. Point (a) is the model application that forms the model core and defines the procedure to solve policy scenarios. A complete scenario comprises five steps:

Step (1) computes the macroeconomic reference scenario values of the impact indicators for the target year. Variations of the reference scenario are expressed in 'business as usual', 'high-growth' and 'low-growth'. They vary in terms of no change, positive and negative anticipated trends of the incorporated land use drivers: Oil Price, Expenditures for Research and Development, Labour Force, Demographic Changes and World Economic Demand.

Step (2) identifies the policy case (thematic area) to be simulated. Each policy case contains sets of instruments. Within each case the user can select and combine different policy instruments as well as the intensity of each instrument can be changed.

Step (3) illustrates the computed scenario results of impact indicators as consequence of the policy settings. Results are presented in interactive maps, tables and graphs. Photorealistic visualisations support impressions on changes within landscape views. Map layer (Google data) superimpose additional geographical information for specific analysis.

Step (4) evaluates impacts according to sustainability criteria that are expressed in critical limits (thresholds, targets). This valuation defines an allowable sustainability choice space, which is based on region-specific tolerance limits per indicator.

Step (5) aggregates groups of indicators to Land Use Functions (LUF) that indicate the level of goods and services at regional level. Nine LUFs have been defined. 'Provision of work', 'Human health and recreation', 'Cultural landscape identity', 'Residential and non-land based industries and services', 'Land based production and Infrastructure', 'Provision of abiotic resources', 'Support and provision of habitat' and 'Maintenance of ecosystem processes' (Perez-Soba et al. 2008).

Policy options can be compared and valuated at the level of single indicators or of LUFs, but the valuation of policy options by sustainability criteria depends on end users` opinion.

Methodology

In order to be able to compute a high number of simulation runs with short model response time, SIAT needs a specific meta-model concept. The algorithm for calculations in SIAT is composed of mathematical response functions, which are derived by quantitative modelling techniques (see fig. 2) using a model framework. The model framework is not implemented into SIAT, but the response functions.

In order to assess the response functions, the modelling framework interacts among linked components as follows: First, macro economical modelling is carried out by NEMISIS (Kouvaritakis 2004) in sectoral division for respective administrative regions. NEMESIS safeguards the statistic accounting frame and allocates needed sectoral demand-driven land claims. Based on the macro-economic agricultural and timbered land claim, the sectoral models CAPRI (Britz et al. 2003) and EFISCEN (Lindner et al. 2002) determine intra-sectoral interrelations and feedbacks land prices and physical land claims to the macroeconomic sector level. Thus, feedback loops between the macro- and sectoral models assure consistency between macro-economic and sector input-output relations. The consolidation of the model framework is reflected in equilibrium prices and physical supply-demand equilibrium for goods and services related to land allocation. The modelling results are disaggregated to grid level (1x1 km) using the CLUE model (Kok et al. 2000) in order to consistently superimpose model projections with the standard result regionalisation of computed impacts.

By applying this procedure for one policy intensity (fig. 2. figure above), one point estimation is assessed. Iteratively varied policy intensities provide a higher number of point

estimations: The dots on the presented function in the upper figure 2 express the mathematical correlation between either land use changes or other indicator values. The function shape as best fitting line among the point estimations is estimated through econometric techniques (Jansson 2006). One function represents the described correlations in one region. Assuming that only three subsets of policies across six sectors are defined and 30 indicators are implemented, approx. 600.000 functions for one policy case have to be integrated in SIAT.

This concept of implementing pre-calculated simulations results (1) reduces the model response time, (2) allow covering a high number of indicators with differently applied methods, (3) covers cross-sectoral integrated impact issues as quick-scan analysis for immediate decision support.

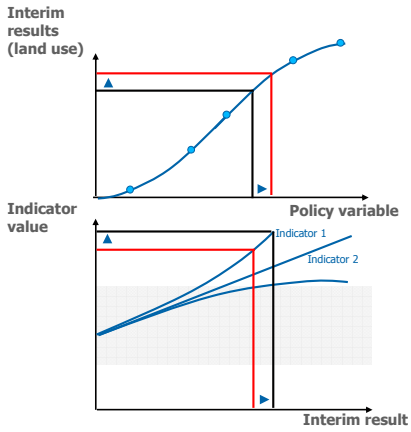


Fig. 2. Policy and indicator functions in SIAT

The SIAT prototype II has achieved a complete system integration of all implemented and interacting components. The three major groups of components were defined as follows:

(1) The SIAT-End user tool ensures full functionality to define policies for analysing their impacts and sustainability valuation. The effects of different policy options can be compared among each other. The (a) interaction model, (b) the software architecture and (c) the graphical design have been redesigned. (d) Standard visualization components as maps, tables and spider diagrams have been implemented.

(2) The SIAT-knowledge base contains all functional relations in a quantitative or qualitative form. Both concern the translation of policies into impacts and sustainability valuation as well as by applying (spatial) up- and down scaling methods. This knowledge is based on numeric data bases and fact- and information sheets to trace calculations and ensure transparency on assumptions.

(3) The visualisation of the complex models and scenario results beyond the SIAT-standard graphical solutions contains the following advanced methods: 2D and 3D-visualisation, namely (a) a 3D-Landscape Generation Model that serves as input for a photo-realistic 3D Visualisation (visualisation system L-VIS), a schematised Visualisation (Biosphere3D) that applied None-photorealistic Rendering Techniques. Complementary thematic map dialogues have been developed as additional Google map-layer to be superimposed with SIAT-result maps.

2.2 User involvements and prototyping

Essential and predominant prerequisite for a successful model design is meeting the end user requirements of end users. The functionality and methodology of SIAT has been developed by participatory processes with potential end users, who are involved in either decision making or impact assessment processes. By applying the method of operational prototyping stakeholders and key representatives of impact assessment have been involved at various levels (Guida et al. 1999). For this, the planning of structured processes is key requisite of success in terms of using given capacities preferably cost efficient.

User involvements

This chapter reveals major considerations from a theoretical perspective, which is then compared in the following chapter with the SIAT development.

The process of stakeholder involvements by potential need specific considerations in terms of the EU Commission as principal organisation as governmental representative. The selection of potential end users either as individuals or groups affects highly decisions on the model design. They influence both the functionality and the design (Hemmati 2002). Thus the organizational structure has immense influence, which has been taken into account as follows:

Institutional analyses have been performed both, from literature and as operating experience to take into account main requirements and organisational aspects into the current prototype design. Different roles, interactions and applied methods between participants have been analysed towards achieving a common SIAT design that ideally meet exactly the EC end users' requirements of a preferably broad audience (Checkland and Holwell 1999).

Supporting decision making limits the scope of the SIAT design process to a specific focus on an end-users' information needs. For any existing process of decision making the institutional structure plays an important role for the design. SIAT aims at providing relevant information in a manner, which improves the way in which the employees of the European Commission (EC) work together across the different organisational structures of Direction Generals (DGs). In order to meet the goal of an accepted SIAT design the organisation should be analysed with regard to organisational structure, internal processes and roles of actors.

Specific hierarchies and the degree of cross-organisational use cause different requirements on the design (Vetschera 1997). Generally, wider user groups and increasing cross-departmental decision spaces lead to an increase of support required for user-friendly handling. Due to abundant cross-sectoral thematic views, the analytical level is broader and focuses rather on comprehensive quick-scan analysis than on high performance of accuracy. The decision level of the potential SIAT user group aims primarily at a hierarchical system that supports decision making within the EU-Commission at the same organisational level. Hence, SIAT provides information which directly guides to the decision solutions (Fredman et al. 1999) at the same organisational level of the EC for cross-cutting analysis.

Different operational aspects of common objectives should be considered, as they affect the design of SIAT. Ideally SIAT will be used by the scientific consortium designing the tool and at the same time by externals at the EC level. The SIAT designer have to understand the demand on design in orientation and should use 'socio-technical' methods like Soft Systems Methodology (Winter et al. 1995) during the development processes to better reflect organisational needs in tool design. Often a good narrative is more engaging and useful than the best science (Checkland and Holwell 1999). Therefore, the SIAT interface and the entire model development itself should try to conform to the preferred communication systems of targeted end users.

In summary supporting organisational decision making at the EC level should minimise the risks by (1) establishing linkages with an adequate number of potential end users as catalysers in case of staff rotation and displacements respectively; (2) involving potential end users in the development process earliest possible, but with respect to different development phases of stakeholder involvements. (3) As key for creating awareness collaborative development should further be strengthened in terms of increasing the use of SIAT. (4) Continuity of the iterative process development towards a reliable and confidential relation between respective sharers is an essential success factor (Mcintosh et al. NN).

SIAT development

Taking the above described considerations on user involvements into account, the SIAT considered the following steps and actions during the development. As the previous chapter described an ideal situation, this consideration describes actual undertaken measures.

(1) Reviewing and benchmarking gathered knowledge of similar projects

Based on the existing in-house knowledge a critical review and benchmark of existing model approaches is essential. With regard to similar projects (e.g. EURURALIS, SEAMLESS) existing knowledge could be used with regard to the conceptual design. Reusing of existing model components is most effective to avoid redundancies. In this regard simulation procedures, calculation techniques, visualisation components as well as already established relations to sub-contractors could be efficiently used or shared among projects. The overall effect of these measures to reduce costs can be considerable.

(2) Adjustment of basic requirements

Based on the Documentation of Work (DOW) basic requirements with reference to analytical objectives and technical specifications have been surveyed. The DOW description was unspecific enough to have maximal freedom for 'own' specifications in terms of designing SIAT, but the analytical objective was clearly defined. Further surveys at the level of contracting body (EU Commission) created a common project understanding and priority setting of objectives. Thus, indispensable is a close contact to the responsible commissioner's view. Once, basic requirements and a priority setting were available, both have to be translated into a first prototype.

(3) Develop a simplistic Prototype

Based on both previous processes, the modelling group met for a one-week hands-on exercise on prototyping. The group was composed of researcher with different background from software engineers, landscape planners to agricultural engineers. Subcontracted graphical designer delivered on-demand design elements. The result was a fully functional SIAT prototype I that contained (a) an user interface, (b) a topic-structure for content management (c) a simplistic model application for one exemplary policy simulation and (d) visualisation. This SIAT version was evident key for an improved communication on the model design and functionality[1]. The gathered feedback by structured user involvements in group discussions (see number (4)) served as input for prototype II.

(4) Group discussions with end users

Group discussions based on SIAT prototype I aimed at gathering end user requirements in a structured way. Preferably a mixed group of software engineers, researcher and policy experts as potential end users have been involved. Apart from numerous SIAT presentations at scientific conferences, cross-institutional workshops have been organised: (a) DG Research / EU Commission 2005 [kick-off meeting on expectations], (b) ISPRA / Joint Research Institute 2006 [stakeholder meeting on conceptual design], (c) Directorates-General & EU Parliament 2006 [Stakeholder meeting with SEAMLESS FP6-project], (d) SENSOR / Research institutes 2006-2008 [various internal discussion on conceptual design], (e) SENSOR Peer-group meeting 2008 [feedback and advice on SIAT prototype], (f) IPTS / Joint Research Institute 2008 [scientific discussion and feedback], (g) DG Research (EU Commission) planned.

The workshops showed the importance of establishing key contacts with potential end users. The modelling group faced the major problem to establish a continuous group over time that allows iterative feedbacks. Main reasons for this were strategic behaviour, expressed shortage of time and a high fluctuation of positions among key stakeholders within institutions. Deliberate discussion guidelines intended to focus on general expectations on the functionality ("What should the model perform to make useful?") and

[1] Previous communications of the model design by means of a descriptive model definition were less successful in terms of an effective apprehension.

specific requirements to cover a preferable broad thematic area of application among all potential end users ("What do you need additionally for your specific scope of work?").

(5) Targeted input of experts

Along with group discussions single interviews with end users and specific experts (e.g. interaction model design) steered the fine tuning of the SIAT design. This additional technique is a valuable technique, since sincere options and expert judgments are only expressed in bilateral discussions. In contrast group discussion may lead to strategic behaviour on expressed end user requirements. A specific group within SENSOR conducts stakeholder and end user analysis and surveyed expectations in bilateral interviews. Apart from numerous internal expert interviews external surveys are being conducted and results will be expected by midyear 2008.

(6) Final negotiation with regard to given capacities

Given specifications and definite further requirements based on demonstrated prototypes need final decisions on ultimate changes. This last step includes negotiation with regard to given capacities, cost estimations on realistic possibilities. This process is not finished. The SENSOR project will last until June 2009.

3 Conclusions

The important aspects discussed in this article concern the model design and functionality of the meta-model SIAT, which have been developed on the base of end user involvements. This process was accompanied by applying the method of operational prototyping.

SIAT consists of response protocols, which are generated from a modelling framework consisting of a range of macro-, sector and land use models. The protocols allow simulating land use-related policies for Sustainability Impact Assessment at regional level of the EU with the predominant advantage of a minimised response time (quick-scan analysis). The meta-model concept causes specific needs for knowledge integration by means of non-standard technical solution finding. The combination of qualitative and quantitative integration techniques enables covering a high number of methodologically diverse indicators.

The research question emphasised the transfer of end user requirements into methodological and functional model advancements, which have been integrated iteratively into the SIAT meta-model. Concluding findings regarding the development process are:

- A first review and benchmark of existing tools among involved institutes is indispensable process for using cost efficient synergies. Considerable cost reduction effects for software development can be achieved.
- Based on experiences of analysing end user requirements, a simplistic first prototype is evident key to communicate the functionality and intended model design.
- Group discussions with end users is a valuable instrument, whereas key contacts to potential end users and permanently alternate meetings are most crucial obstacle.
- Targeted inputs of experts in bilateral interview form are essential, because 'strategic opinions' in group discussion are by-passed and thus sincere options can be focused.
- Final negotiation with regard to given capacities should be taken into account when the final version is developed. Mutual expectations on contractors and clients side are to be adjusted.
- Understanding the model development process helps to steer the model design in order to assure success in terms of acceptance, utility and high degree of utilisation.
- Knowing the institution regarding its organisational structure is an empiric key for efficient result-oriented end user collaboration on specific requirements of integrated impact assessment models (Sieber et al. 2008).

ACKNOWLEDGEMENTS

The authors wish to thank all members of SENSOR for numerous necessary inputs in order to design the SIAT prototype II.

REFERENCES

Bartolomeo M, Giugni M, Hertin J, Jacob K, Rennings K, Volkery A, Wilkinson D, Zanoni D. Approaches to Impact Assessment in six OECD countries and at the European Commission. IQ-Tools Project report, WP1A, 2004.

Britz W, Perez I, Wieck C. Mid-Term Review Proposal Impact Analysis with the CAPRI Modelling System, Mid-Term Review of the Common Agri-cultural Policy. July 2002 Proposals Impact Analyses, European Commission, Directorate-General for Agriculture ed, 2003.

Checkland P, Holwell S, Information, Systems and Information Systems. John Wiley & Sons, Chichester, 1999.

Davis Alan M. Operational Prototyping: A new Development Approach. IEEE Software, Page 71, 1992.

European Commission, A Sustainable Europe for a better world: A European Union Strategy for Sustainable Development. COM(2001) 264final, 2001.

European Commission, Impact assessment guidelines. SEC791, 2005.

Fredman J, Horndahl M, Strömberg LT. A Framework for the Design of Organizational Decision Support Systems, Department of Informatics, Göte-borg School of Economics, Goteborg University, 1999.

Guida, G., Lamperti, G., Zanella, M. Software Prototyping in Data and Knowledge Engineering, Springer, ISBN-10: 0792360168, 1999.

Helming, K., Bach, H., Dilly, O., Hüttl, R., König, B., Kuhlman, T., Perez-Soba, M., Sieber, S., Smeets, P., Tabbush, P., Tscherning, K., Wascher, D., Wiggering, H. Ex ante impact assessment of land use change in European regions – the SENSOR approach. In: Helming K, Tabbush P, Perez-Soba M (Eds). Sustainability impact assessment of land use changes. Springer, 77-105, 2008.

Hemmati, M. Multi-stakeholder Processes for Governance and Sustainability, Earthscan, London, 2002.

Hemmati, M. Multi-stakeholder Processes for Governance and Sustainability, Earthscan, London, 2002.

Jansson, T. Response functions and indicator functions in SIAT, SENSOR working paper, 2006.

Kok, K., Veldkamp, A. Using the CLUE framework to model changes in land use on multiple scales, Tools for land use analysis on different scales, Kluwer Academic Publishers, Dordrecht, pp. 35-63, 2000.

Kouvaritakis N, Zachariadis T. New Econometric Model for Environment and Strategies Implementation for Sustainable Development (NEMESIS), The Environmental Module. Final Report, National Technical University of Athens, Greece, 2004.

Lindner M, Sohngen B, Joyce LA, Price DT, Bernier PY., Karjalainen T. Integrated forestry assessments for climate change impacts. Forest Ecology Management 162: 117-136, 2002.

Mcintosh, B S, Giupponi, C, Smith, C, Voinov, A, Assaf, H, Crossman, N, Gaber, N, Groot, J, Haase, D, Hepting, D. Kolkman, R, Matthews, K, Monticino, M, Mysiak, J, Quinn, N, Scholten, H, Sieber, S., Bridging the gaps between design and use: developing tools to support management and policy: In: Jakeman, A.J., Voinov, A., Rizzoli, A.E. and Chen, S.H. (eds) (NN) Environmental Modelling and Software. Developments in Integrated Environmental Assessment Series. Elsevier Ltd.

Perez-Soba, M., Petit, S., Jones, L., Bertrand, N., Briquel, V., Omodei-Zorini, L., Contini, C., Helming, K., Farrington, J., Tinacci Mossello, M., Wascher, D., Kienast, F., De Groot, R. (2008): Land Use Functions – a multifunctionality approach to assess the impact of land use change on land use sustainability. In: Helming K, Tabbush P, Perez-Soba M (Eds). Sustainability Impact Assessment of Land Use Changes. Springer, 375-404, 2008.

Sieber, S., Müller, M., Verweij, P., Haraldsson, H., Fricke, K., Pacini, C., Tscherning, K., Helming, K., Jansson, T. (peer-reviewed) (2008): Transfer into Decision Support: The Sustainability Impact Assessment Tool (SIAT), In: Helming K, Tabbush P, Perez-Soba M (Eds). Sustainability Impact Assessment of Land Use Changes. Springer, 107-128, 2008.

Sieber, S., Müller, K., Helming, K., Fricke, K., Pohle, D., Tscherning, K., Verweij, P., Wascher, D., Perez-Soba, M., Kuhlmann, T., Van Meijl, H., Tabbush, P., Dilly, O., Bach, H., Haraldsson, H., Sverdrup, H., Pacini, C., Sustainability Impact Assessment Tools (SIAT): Modelling impacts of land use-related policies, IAJ – Integrated Assessment Journal, Bridging Science & Policy, Vol 7, No 1, Integrated Assessment, 2007.

Tamborra M, Socio-economic tools for sustainability impact assessment – the contribution of EU research to sustainable development. Energy, environment and sustainable development. EUR 20437.1 Office for official publications of the European Communities, Luxembourg, 2002.

Vetschera R. Decision Support Systems in Networked Organizations. Uni-versity of Vienna, Management Center Vienna, 1997.

Verweij PJ, Sieber S, Wien JJF, Müller K, SIAT - a Sustainable Impact Assessment Tool for Understanding the drivers in integrated impact assessment. International Conference IEMSS, 6 p, Vermont USA, 2006.

Winter MC, Brown DH, Checkland PB, The role of soft systems methodology in information systems development. European Journal of Information Systems 4: 130 – 142, 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

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

ABSTRACT

Due to its fuzziness the model of sustainable development has to be particularised and contextualised before any positioning with respect to the yardstick of sustainability. 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. Initially, problems of urban development were identified in a bottom-up approach by local stakeholders. These problem areas were contrasted with a set of sustainability rules which had been systematically derived from the basic sustainability norm, stating minimum requirements for sustainable development. At the interface of local problems and sustainability rules indicators had been identified which provide information on whether the city is over time becoming closer to or farther removed from the respective sustainability goals in its problem areas. This paper shows how such sustainability monitoring system can indicate municipal problems and contextualise the sustainability norm into the administrative structures.

Keywords: *sustainability, monitoring; operationalisation; indicators; Leipzig.*

Introduction

Over the last 20 years, the sustainability concept has become a guiding star for political activity all over the world, especially in cities, towns and smaller communities. 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.

However, if a local authority has to state its position regarding sustainability or if the basic concept of sustainability is used as a yardstick for political and administrative action, the stakeholders involved will find themselves forced to particularise the global concept of sustainability and place it into their local context. The reference models usually employed could only be used to a limited extent for such a systematic definition taking into account local conditions.

The definition put forward by the Brundtland Commission is too abstract, the popular three-pillar model too vague, and the Agenda 21 adopted at the Rio Summit far too loose to serve as a blueprint for local sustainability. The numerous sustainability indicator systems developed and propagated in Germany mainly designed for the local level reflect a considerable scope for interpretation regarding the operationalisation of the sustainability concept.

Apart from the desideratum that the local situation should be taken into account when a system of sustainability indicators is being developed, the need for more local contextualisation also means that ways need to be found of building on existing local information systems from the angle of sustainability. This approach is being taken at the Helmholtz Centre for Environmental Research – UFZ, which is collaborating with the city councils of Halle and Leipzig to develop and test an integrated sustainability information system (IGNIS). The system is primarily designed with the needs of local government in mind and is to be made accessible to all council departments over their intranet. The aim is to provide sustainability-related support for local political and administrative decisions.

The main questions of the project were: How the sustainability concept could be contextualised to the local level? In which form sustainability monitoring at best contributes to steer municipal development? How sustainability monitoring can be used to link different existing municipal monitoring activities (social, environmental, health, planning)?

The Integrated Sustainability Concept ("HGF Concept")

The development of the new information system's structure and methodology is geared to the Integrated Sustainability Concept recently put forward by the HGF, Helmholtz Association of National Research Centres (Kopfmüller et al. 2001). Starting from an understanding of sustainability whose central ethical postulate is equity, the HGF Concept systematically sets out the sustainability norm as defined by the Brundtland Commission. It has three core elements regarded as constitutive

for sustainability: *equity* (within and between generations), *globality* and *anthropocentricity*. This sets it apart from the two- and three-pillar concepts which have so far dominated both scholarly and political debate on sustainability. The diverse problems of these concepts (normative vagueness, the unsolved integration problem, a tendency towards 'sectoralisation', etc) prompted HGF researchers to derive (initially) three general interrelated aims for sustainable development from the above-mentioned constitutive elements: *securing human existence, preserving the productive potential of society, and maintaining development and action possibilities*.

These general aims were then used to develop an extensive set of 'sustainability rules' along the lines of the ecological management rules that have already been debated for some time. The sustainability rules particularise the targets for various themes and thus make up the normative core of the HGF Concept. Conceived as minimum requirements for sustainable development, they reflect universal principles to which policy directed towards sustainability (be it at the global, national or local level) needs to be geared.

Problem orientation: local problems

This normative approach to the deductive model definition (top-down) is augmented in the HGF Concept by an inductive, problem-based approach to the topic of sustainability (bottom-up). Starting from the current scholarly and social discourse, the main sustainability problems are identified and compared to the sustainability rules. The bottom-up approach, which ultimately regards sustainability as a social construct to be modified in discourse, reduces complexity by acting as a filter for the broad range of themes covered by the sustainability rules. In addition, however, it enables the universal set of rules to be adapted to the conditions of different spatial, temporal and social contexts – since a certain social group will at a given time in connection with a given level of spatial analysis develop its own specific understanding of sustainability and an equally specific definition of problem areas.

Consequently, the local contextualisation of the sustainability model takes place at the interface of norm-orientated top-down and problem-orientated bottom-up approaches. The model becomes further operationalised as indicators are identified which reflect the changes to the sustainability problems with reference to the corresponding rules (Fig. 1).

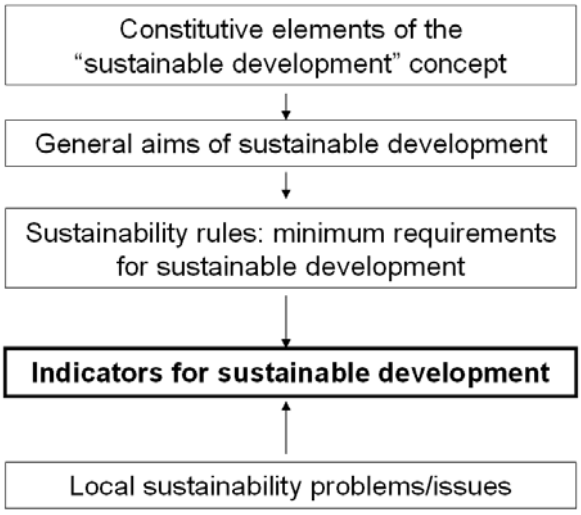


Fig. 1: Diagram illustrating how the top-down and bottom-up approaches are linked in the Integrated Sustainability Concept.

Figure 2 illustrates how the indicator process was carried out: the identification of indicators took place at the interface between rules and problems. The indicators allow to contextualise the concept of sustainability on the local level and to link it with local problems. In most of the cases there were chosen a set of indicators for every so-called rule-problem-complex. Using respective data sets it allows a quantitative measurement of sustainability (Fig.2).

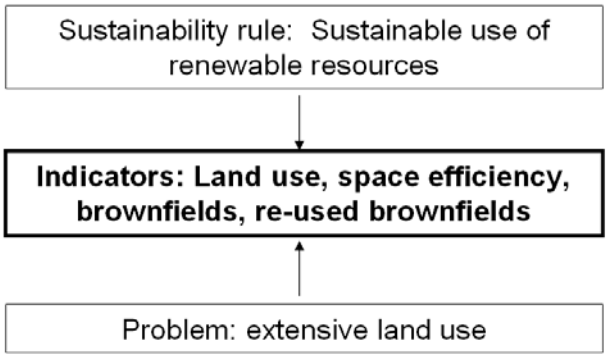


Fig. 2: Example illustrating the choice of indicators.

Although the HGF Concept was originally developed for the national level, all three aspects – the normative definition of the model, the contextualisation strategy implicitly contained, and the identification of indicators as the main element –

make the HGF Concept appear suitable for the operationalisation of sustainability at other structural levels, too. Therefore, the set of sustainability rules was adopted almost unchanged as a conceptual basis for the development of the local indicator system.

The indicators were identified in four stages, we started with assessments by local stakeholders of the current and future problems of the cities of Halle and Leipzig (‘problem orientation’). After introducing the sustainability rules of the HGF Concept as the target or yardstick (‘norm orientation’), rules and problem areas were interconnected (‘contextualisation’). Indicators were then determined for the resulting thematic complexes, which simultaneously led to the further particularisation of the issues covered by the rules and problem areas (‘operationalisation’).

Contextualisation: local problems in the light of sustainable development

Linking up the bottom-up and top-down approaches enabled the local problems of Halle and Leipzig to be compared with the rules of the Integrated Sustainability Concept, i.e. with aims derived from the sustainability model. In order to build the conceptual bridge between the local problems and aims, indicators now had to be identified which over the course of time could provide information about whether the local authority is becoming closer to or farther removed from the sustainability goals in its problem areas. Such indicators are required for both the continuous observation of the problem areas and the systematic controlling of political measures in this sphere. However, the relatively high degree of abstraction of the thematic complexes from sustainability rules and local problems necessitated further content circumscription. Hence the identification of indicators involved not just determining characteristic values for the measurement of sustainability in the area concerned but also the further operationalisation and particularisation of the model under the given local conditions. Therefore, this step can ultimately also be regarded as a discursive process moulded by both the participating scholars and the local authority stakeholders.

The selection process resulted in two extensive systems of indicators, each of which is designed to reflect locally specific, problem-orientated operationalisation of sustainability for the cities of Halle and Leipzig. Containing a total of 111 indicators for Halle and 105 for Leipzig, the two systems showed high congruence

(about 90%). As some indicators were selected for a number of different rule-problem complexes, the total numbers of indicators rose to 161 (Halle) and 155 (Leipzig).

Technical implementation

The indicator systems compiled are an important milestone on the road to integrated sustainability reporting for the cities of Halle and Leipzig. Once the indicators have been underpinned with data and integrated into an interactive web application, the two cities are to be equipped with a user-friendly, web-based georeferenced sustainability information system (IGNIS) allowing sustainability data to be accessed by all departments (Figure 3).

Such an information system can in principle fulfil a whole string of functions. It can for example 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).

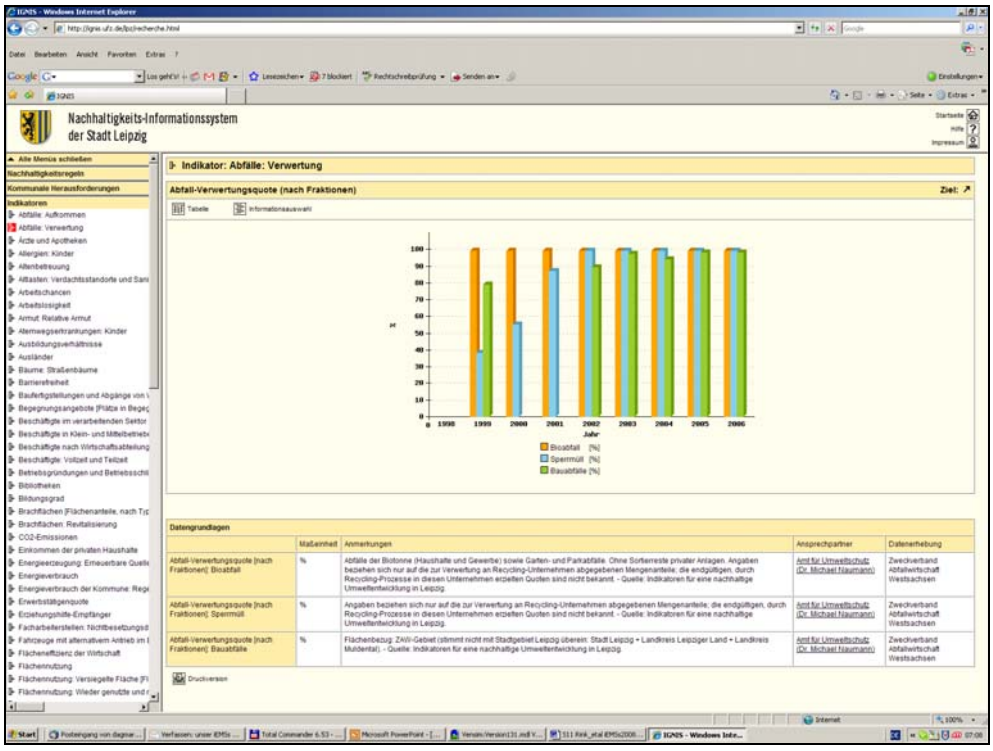


Fig. 3: IGNIS – local sustainability information system website.

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. And if the information system is used to check the success and effectiveness of certain development measures, it could also have an appraisal and verification function. In this particular case, the main consideration is probably to keep local government and authorities (and possibly also the general public) abreast of the implementation of local sustainability (i.e. sustainability monitoring). Whether this will lead to additional usage within local decision and control processes as described by means of the guidance and verification function (sustainability controlling) will mainly depend on how the stakeholders involved take to the information system.

The procedure chosen for the development of the indicator system has already proved successful. In particular, the systematic definition of the global sustainability norm and its incorporation into the local context have turned out to be especially helpful tools both for the selection of indicators for sustainable development at the local level and for the acceptance of the sustainability model among the stakeholders involved. Both monitoring systems were integrated into the municipal bargaining. By means of annual sustainability reports via the councillors the urban development shall steered towards sustainable direction.

Outlook

NAMOS is a specification of IGNIS and elaborates the sustainability concept to the specific problems of so-called “shrinking cities". Using the example of the city of Leipzig, we develop a monitoring system that aims to measure the sustainability of the shrinkage. By shrinkage we mean a multidimensional process in which economic and demographic decline join to physical decay.

Up to now, we identified a pool of indicators, formulated problems relevant to the shrinking process and related these to the rules of the sustainability concept of the Helmholtz society. The methodological concept is based on the IGNIS system and applies the latter to the phenomenon of shrinkage and will use an enhanced Geo-Information System for visualization.

The monitoring system NAMOS will provide a data base for urban modelling projects. Furthermore we plan an independent analysis that extensively evaluates the shrinking process following the rules of the HGF Concept. Besides we will also analyse processes of social change and relate them to urban social-science concepts such as segregation, reurbanisation and gentrification.

References

Hartmuth, G.; Rink, D.; Huber, K (2007): Operationalisation and contextualisation of sustainability at the local level: Stages in the development of a sustainability indicator system, in: Sustainable Development (accepted).

Kopfmüller, J.; Brandl, V.; Jörissen, J.; Paetau, M.; Banse, G.; Coenen, R.; Grunwald, A. (2001): Nachhaltige Entwicklung integrativ betrachtet. Konstitutive Elemente, Regeln, Indikatoren, Berlin.

Rink, D.; Hartmuth, G.; Huber, K. (Eds.) (2005): Raum für Nachhaltigkeit. Zur Kontextualisierung des Leitbilds, Berlin.

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 Methodology for Building Credible Models for Policy Evaluation

C. A. Aumann [a]

[a]*Land-Use Management Systems, Alberta Research Council, 250 Karl Clark Road, Edmonton, Alberta, Canada (craig.aumann@arc.ab.ca)*

Abstract:

The need for tools capable of evaluating the potential impacts of alternative policies has been expressed by many. This paper focuses on a methodology that describes how credibility can be constructed for models used to evaluate alternative policies. Relative to modeling conducted in scientific contexts, however, modeling for policy evaluation has notable differences that would lead some to say that achieving credibility in such results is not possible. We first introduce a model assessment framework that enables us to describe how models for policy evaluation can still be more or less credible despite the differences from scientific modeling contexts. The argument presented depends primarily on i) the scalar hierarchical structure used to represent the complex policy system, ii) the ability of experimental frames to include a variety of constraints and/or weak data patterns across the scalar levels used to represent the system, and iii) the way in which the framework facilitates critique by stakeholders.

Keywords: Assessment; Critique; Complexity; Validation; Verification; Hierarchy theory; Simulation

1 INTRODUCTION

The need for tools capable of evaluating the potential impacts of alternative policies has been expressed around the world [e.g., Owens et al., 2004; Gov. Canada, 2005]. Despite this need, research over the last 30 years has demonstrated that the impact of policy assessment on policy decisions does not occur via a linear process - meaning that the outputs from policy assessments are generally not carefully considered or used directly by decision makers [Owens, 2005]. Instead, the impact of these tools occurs in more subtle and nuanced ways such as by facilitating group learning among stakeholders and providing ammunition that can be used to persuade opponents [Owens, 2005]. In considering how credibility can be established in the context of policy evaluation, the main conclusion is that while possessing credible models does not guarantee that a policy change will occur, possessing models that are credible to stakeholders and domain experts is indeed necessary for policy change to occur in contentious situations.

This paper thus focuses on the question of how to build credibility in models used to evaluate alternative policies. Others have described the issues associated with building such models [e.g., Couclelis, 2005; Jakeman et al., 2006]. After introducing a number of concepts and definitions to provide the necessary context associated with simulation modeling generally, a general assessment framework is described that builds on previous work [e.g, Zeigler et al., 2000; Aumann, 2007]. Some of the unique challenges associated with models used to evaluate policies are then discussed along with how the assessment framework presented can be applied to build credibility in policy evaluation contexts. Given that credibility must be established with stakeholders, the way in which results from this assessment framework could be delivered using web-based tools are also described.

To clarify the concepts and definitions introduced, a single running example is used throughout. This example focuses on how a model developed to evaluate a policy of tradable disturbance credits (TDCs) can be assessed. Similar to SO_2 emission trading, under TDCs a cap is placed on the cumulative amount of landscape disturbance that can occur within a region and industrial players within the region can purchase and trade permits for the right to disturb land [e.g., Weber and Adamowicz, 2002]. This is one policy under evaluation for efficiently achieving ecological and economic objectives in the oil-sands area of NE Alberta, Canada. An agent-based model is currently being used for this evaluation, but only a small number of the model's components are presented here to illustrate the concepts in this paper. The components used in this paper include industry agents (which attempt to maximize their economic return by exploring for oil across the landscape and use the exploration information to prioritize the drilling of wells and building of pipelines, etc), a government agent (which sets disturbance cap, monitors cumulative disturbance across the landbase, and auctions disturbance permits), and the natural environment (which includes things like natural disturbance and vegetative succession).

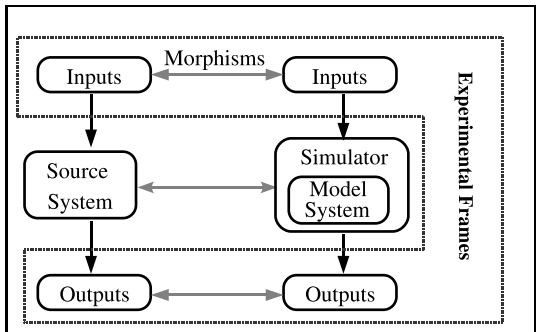


Figure 1: A general simulation framework consists of both source and model systems; the morphisms between the two systems in terms of input, outputs and model structure; and the experimental frames.

2 Assessment of Simulation Models

2.1 Simulation Framework

At a high level, a *model* is just a state transition function or mechanisms that instructs the *simulator* (e.g., a computer or algorithm capable of executing the model instructions) on how to generate outputs from inputs (Figure 1). The *source system* is the real or virtual system that is being modeled and the goal of modeling is to achieve a representation of the source system so that under "similar" inputs, both systems produce "similar" outputs. This idea is captured by the *experimental frames* which include a specification of the conditions under which both the source and model systems are to be observed or experimented with, along with mechanisms for comparing the two systems. The aim of model assessment is to build credibility in the model and is accomplished by ensuring that both systems "agree" over a sufficiently wide range of conditions encompassing the objectives motivating the modeling project.

One approach for representing complex systems is to decompose them according to a *scalar hierarchy* so that the objects at a given level contain, volumetrically and structurally, the objects of lower hierarchical levels [e.g., Kline, 1995; Giampietro, 2004]. As illustrated in Figure 2, one can think of the levels below a focal level as providing the "mechanism" while higher levels provide a "purpose" or context for the lower levels. Emergent properties are taken to define each level [Aumann, 2007].

At each level in this scalar hierarchy, model components can be specified with varying degrees of detail according to a *specification hierarchy*. An *I/O Behavior* specification is like a black-box in which inputs are mapped directly onto outputs. For example, in Figure 2 the way in which the Exploration component of the Oil Company is represented might follow an I/O behavior specification. In a *I/O System* specification, the model maintains an internal state that can be changed by model inputs. Thus, whether the Oil Company engages in exploration at a given time depends on it possessing sufficient TDCs. A *Coupled Component* specification is a model composed of other I/O Behavior, I/O System, and possibly other coupled component models and is illustrated by the entire Oil Company sub-model.

Scalar and specification hierarchies enable definition of what is meant by "similarity" between two systems [Zeigler et al., 2000]. Two systems are said to be *morphic* at a given specification level if it is possible to establish a direct correspondence between the defining elements of each system at the same specification level within some experimental frame. For an I/O Behavior specification, the morphism is simply the comparison of the inputs and outputs of both systems. For an I/O System specification, the introduction of a state space requires the introduction of a mapping between the state spaces of both systems. This mapping is said to be *homomorphic* if there is a defined correspondence (but not necessarily an identity) between the states in both systems so that both systems progress through similar pathways to achieve similar model outputs. Finally, since component model specifications can involve all three model specification types, it must be ascertained not only that the output of the overall model is correct (similar to I/O Behavior), but also that the outputs are produced for the right reasons (i.e., that the homomorphisms between the components of the two systems hold).

2.2 Model Assessment Framework

In the context of the simulation framework presented above, the goal of model assessment is to establish the strengths of the morphisms between the source and model systems. Building such credibility is accomplished via a processes of model verification, model validation, and critique. It should be noted that the meanings of these terms is not entirely consistent across fields with some eschewing the use of the term "validation" [e.g., Anderson and Bates, 2001], others noting the problems implied by the term while acknowledging it's widespread use [e.g., Oreskes and Belitz, 2001], and other distinguishing numerous different kinds of validation including operational, conceptual, data, and even processes [e.g., Rykiel Jr., 1996].

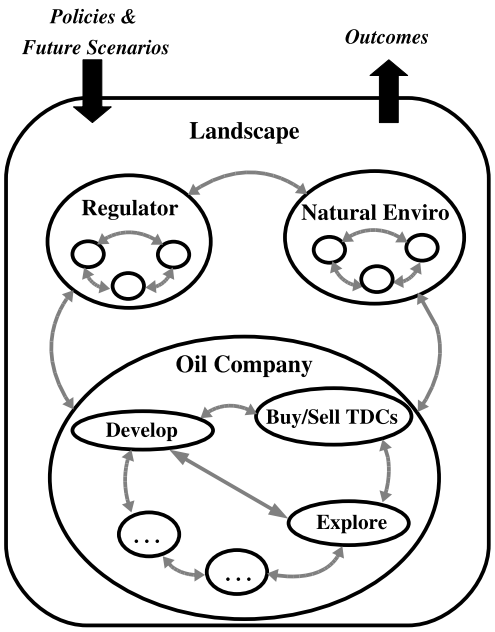


Figure 2: Simplistically, the entities relevant to evaluating the TDC policy include the oil companies, a regulator, and the natural environment. These entities all operate over a defined landscape area, and in turn are comprised of other lower-level components. In the case of an oil company, such components include units in charge of exploring for oil, developing oil (i.e., building roads and pipelines to get the oil to market), and also buying and selling TDCs with the Regulator or other companies.

In this paper, *model verification* means verifying that the actual model implementation is consistent with the model design specifications. Relative to the other assessment processes, verification is relatively easy to accomplish and thus this paper focuses on the processes of model validation and critique and builds on previous work [e.g, Balci, 1997; Rykiel Jr., 1996; Zeigler et al., 2000; Aumann, 2007].

Model validation is about substantiating that the behavior of the model "mimics" the behavior of the system with sufficient accuracy so that it is impossible to distinguish the behaviors of both systems in the experimental frames. *Experimental frames* are an operational formulation of the objectives motivating a modeling project and practically function as a type of measurement or observer system consisting of a *generator* that generates the input to the systems, an *acceptor* that monitors the "experiment" to ensure the desired experimental conditions are met by both systems, and a *transducer* that observes, analyzes and stores the output.

Saying it is impossible to distinguish the behaviors of two systems requires the concept of *replica-*

tive validity, namely that for all experiments possible within the experimental frame, the behavior of the model and the source system agree within the specified tolerance at the I/O Behavior level. For models specified at the I/O System level, the introduction of a state-space necessitates a stronger notion of *structural validity* meaning that the model mimics in a step-by-step, component-by-component fashion the way in which the source system performs its state transitions. Structural validity ensures that the model is generating the correct I/O behavior for the right reasons, and not because incorrect behaviors in one model component are compensated by behaviors in other model components. Methodologically, structural validity can be achieved by applying experimental frames to each of the model components within a given scalar level and across hierarchical levels to assess the larger scale consequence of these behaviors (see Figure 3 and discussion below).

One final, important step for achieving model credibility is a *critique* of the processes of model design, verification, and validation. Such critique is essential because model validity is only established relative to the study objectives as implemented in the experimental frames. If these objectives are incorrectly specified and/or the model is incorrectly defined, the model can still be valid with respect to this incorrect specification even though the simulation results will not be credible when viewed from a broader perspective. However, for such a critique to occur the model verification and model validation steps must be accessible, transparent, and understandable to non-modelers. How this can be achieved in the context of the current simulation and assessment framework will be discussed below.

We will say that a model is credible in a particular problem context if it has been verified, validated, and critiqued. No model can be absolutely credible, but rather models should be thought of in terms of degrees of credibility. Ultimately, the greater the need for high levels of credibility, the higher will be model development and assessment costs. Thus, the level of credibility required for the project needs to be bounded before the model is constructed and model assessment performed.

3 BUILDING CREDIBILITY IN MODELS FOR POLICY EVALUATION

While the above simulation and assessment frameworks appear quite natural for most systems studied by the physical sciences, a number of substantial differences arise in the context of policy evaluation that impact the way in which assessment can be carried out and ultimately the kind and level of credibility that is achievable. While models in scientific contexts are about constructing scientific explanations based on facts and well established theories, in a policy context the purpose is to decide on a course of action based on our values about the kind of unknown (and unknowable) future we desire based on knowledge that is always incomplete [Couclelis, 2005]. In addition, while scientific modeling will likely only ever be critiqued by a relatively small scientific community, a much larger community of diverse stakeholders will seek to comment on any policy assessment - especially if the financial stakes are high. These differences are explored more fully in this section.

3.1 Differences in the Simulation Framework

A major difference between typical scientific modeling and policy evaluation is that no source system exists for the later - since the point of such evaluation is to do it **before** a policy is ever enacted. Evaluation of a policy's performance must thus be judged relative to the future and not relative to observed behaviors - as is typical in scientific contexts. This unknown future is typically expressed using a number of alternative, coherent, plausible and relevant future scenarios capturing the exogenous conditions under which the policy might operate [e.g., Couclelis, 2005]. Thus, relative to the framework shown in Figure 1, the inputs to the model system also contain a set of future scenarios. The non-existence of the source system coupled with the alternative futures used means that limited data will exist for model validation, particularly at higher scalar levels that are critical to evaluating policy performance (e.g., experimental frame C in Figure 3). These differences help explain why the kind of credibility possible in scientific contexts cannot exist for policy evaluation.

3.2 Model Assessment - Validation

Despite these differences and their attendant challenges, we will argue that it is still possible for results of a policy evaluation to be more or less credible. The main tenets of the argument are based on the utility of constraints and "weak data patterns" when applied across the scalar levels of the model to validate model components, and the ability to apply experimental frames across multiple models.

In the context of the running example considered here, the impacts of a policy change would ideally be assessed regionally (in terms of economic and ecological indicators) and at lower scalar levels to capture the impacts on individual firms. The lack of empirical data may lead some to conclude that achieving any kind of credibility is impossible. However, such a position overstates the necessity of data in achieving model credibility. In general, the ability of any model to mimic data at the I/O Behavior level is only weak evidence that the important underlying processes have been correctly captured since models that inaccurately capture underlying processes can still predict quite well [Oreskes and Belitz, 2001]. For example, statistical models can predict quite well provided they are applied in a similar context to which they were built.

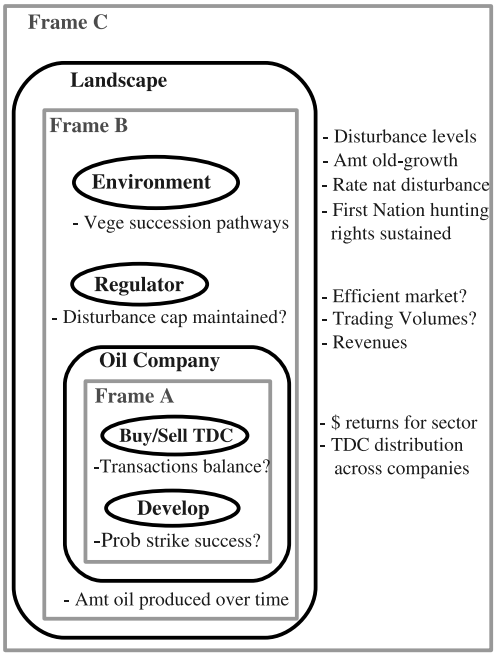


Figure 3: Experimental frames (in gray) encapsulate the lower-level components in Figure 2 where each frame contains criteria on the behavior of the sub-model that need to be respected. For example, the sub-model dealing with buying and selling TDCs needs to ensure that the transactions balance across the entire company, that the prices it is paying for the TDCs are not causing the company's financial ruin, etc. At higher hierarchical levels, the amount of oil produced by a company over time should not fluctuate widely, the market should be efficient, and overall trading volumes for TDCs should be large given a large number of players in the market.

In the context of a policy that alters any existing context for which data may exist, validity can be established using structural integrity criteria at the I/O System level. Structural integrity criteria ensure that the answers achieved at a given scalar level are also achieved for the "right reasons" - meaning that all of the experimental frames applied to lower levels are also satisfied. In the context of the scalar hierarchical model decomposition used, it must be decided whether the behaviors of sub-models are the same under the new policy context. For example, under TDCs the vegetation succession pathways will remain the same, as should the probability of a strike success as exploration proceeds across a basin (see Figure 3). In this case, existing data can be used.

For sub-model behaviors that change under alternative policies, the experimental frames can include a number of alternative options: i) constraints can be applied to these components (e.g., Do a company's number of TDCs and the disturbance it creates balance? Is the disturbance cap being maintained across the landscape?); ii) qualitative behavior assessment criteria (e.g., Are companies developing in areas of actual high petroleum potential? Are companies generally selling TDCs to each other when it makes sense to do so? Is the market efficient?); and iii) "vague" or "weak" data patterns (e.g., Do the natural disturbance patterns created by the model agree with what is observed empirically?). Thus, the lack of "hard data" for the system being modeled is no excuse for failing to perform model validation in policy contexts.

Applying such types of "data" across the scalar hierarchical levels using the experimental frames can provide as much information for model validation as a single strong empirical data pattern [Grimm et al., 2005, & Online Material] and also [Aumann et al., 2006, Appendix B]. The reason for this is that **all** experiments specified for the scalar hierarchical levels below the current level must be satisfied **simultaneously**. Provided the criteria used in the experimental frames are not trivial, satisfying all frames simultaneously rapidly becomes challenging. Further, the way in which experimental frames trade-off with each other can be used to guide refinement of model structure and also refine the experimental frames [Reynolds and Ford, 1999; Wiegand et al., 2003].

Another way in which model validity can be established is by comparing the behaviors of alternative models constructed using different modeling assumptions. In the context of such multiple models, the experimental frames are general enough to allow these models to be validated against each other. However, the challenge here is ensuring that the experimental frames are *applicable* - meaning that the conditions required by an experimental frame can be satisfied by all the models.

In summary, the non-existence of the source system and the lack of data does not mean model validation can be ignored nor that validation is impossible. Instead, the above validation process allows us to say why the model's behaviors are being produced and to demonstrate that these behaviors are being produced for the right reasons. The reason we are justified in believing the outcomes produced by the model results from the behavior of the lower level model components being deemed to be valid based on the experimental frames applied at these level(s) coupled with the experimental frames applied across higher hierarchical levels. As a result, we can have confidence that these lower levels are not spuriously influencing the behavior of higher levels, and that the behaviors of model components are all within reasonable bounds. Because the outputs are produced for reasons that we think are justifiable (if we didn't think this, then we would add additional experimental frames or include models built using alternative assumption), we are compelled to view the results as valid within the assumptions of the assessment.

This approach does, however, also have some notable limitations. First, even if alternative models are constructed and these models all agree across the applicable experimental frames, there is a very real possibility that all of such models are simply wrong due to a common lack of knowledge about the non-existent system being modelled. Thus, agreement across diverse models may simply be a result of common ignorance. While disagreement across models might improve understanding at the stage of critique, this is unlikely to occur under common ignorance and is particularly challenging since no source system exists to act as the ultimate authority. Another limitation in the above method is that evaluations can only be done under a small and finite set of alternative futures. While utilizing a greater variety of alternative futures will help to ensure that a "robust policy" is identified [e.g., Lempert et al., 2006], constructing a large number of alternative futures that are coherent, plausible and realistic presents its own challenges.

3.3 Model Assessment - Critique

A larger critique of model design and validation is essential because model validity is a necessary but insufficient condition for establishing the credibility of modeling results [e.g., Balci, 1997]. As is clear from the discussion above, model validity is only established relative to the experimental frames used. If these experimental frames are inadequate for the study objectives, the model can still be deemed to be valid even though the simulation results will not be credible. Credibility can be enhanced via a process of external review by stakeholders. However, the challenge is delivering the large quantities of information produced during validation in a manner that is accessible, transparent, and understandable to such reviewers who are likely not modelers. This section illustrates conceptually how such delivery could be accomplished under the simulation framework presented.

Since the system being modeled is conceptualized in terms of a scalar hierarchy, these entities are represented as separate sub-models within the overall model. Each of these sub-models is also "wrapped" (e.g., Figure 3) in one or more experimental frames that may encompass one or more

sub-models. Each experimental frame has associated with it criteria governing the inputs that are allowable for the sub-model (generator), criteria to ensure that the assumptions underlying the sub-model are maintained (acceptor), and the frame also monitors the output produced by the sub-model to ensure its acceptability (transducer). Problems could originate in any of these three areas, and the aim is to present such failures to the user in an easy to understand way and enable them to browse the criteria used in the experimental frames. One possibility is illustrated in Figure 4.

Current programming tools enable the hierarchical structure of models to be displayed. Thus, the challenge is to expand these tools to display the experimental frames and also add in an overarching monitoring system that monitors and reports on the status of all the frames as the model runs. Conceptually, users could browse such information as illustrated in Figure 4. The challenge is implementing such a tool in a general manner so that it can deal with models from diverse problem contexts.

4 CONCLUSION AND RECOMMENDATIONS

With the ever increasing power of computers, the amount of modeling done and the complexity of the models built in both scientific and policy fields will continue to increase. However, to date the field of model assessment lags far behind our current computational abilities. Indeed, much work is still needed to arrive at the concepts and a framework for model assessment that is general across disparate disciplines. This paper attempts to advance the field of model assessment, at least as it relates to policy assessment, by introducing a number of concepts that have proven useful in scientific modeling contexts and describing how these same concepts apply to policy evaluation.

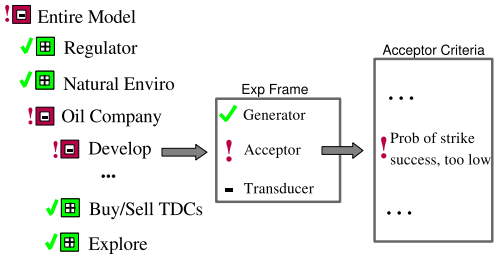


Figure 4: Using the hierarchical structure of the model and the fact that experimental frames cover different sub-models, it becomes possible for a user to quickly identify which sub-models are failing and why, and communciate these results to stakeholders.

The strengths of this framework are that it can be applied in both scientific and policy contexts, suggesting that it is at least somewhat general. Conceptually, the framework lends itself to the communication of model validation results along with the experimental frames in a manner that is accessible, transparent, and likely understandable to stakeholders (Figure 4). Drawbacks to the framework include that implementing it will require substantially more programming effort and running it along with the model will require considerable computational resources.

This framework would seem to be a natural starting place for enabling autonomic model assessment tools - meaning tools that are capable of assisting in the complex process of model validation and the identification of inadequate model components or experimental frames. While autonomic computing aims to create computer systems that are capable of self management - the processes of model verification, model critique, and the specification of the experimental frames all require levels of human involvement that go far beyond anything a computer can do for us. In a policy context where one of the aims of modeling is to foster dialog among stakeholders, we do not see the inability to fully automate this model assessment framework as a drawback. Instead, a more urgent research need is figuring out how to best incorporate model assessment tools into stakeholder consultation to facilitate public consensus on which policy is prefered.

REFERENCES

Anderson, M. G. and P. D. Bates. Hydrological science: model credibility and scientific understanding. In Anderson, M. G. and lastname Bates, P. D., editors, *Model Validation: perspectives*

in hydrological science, chapter 1, pages 1–10. John Wiley & Sons, New York, 2001.

Aumann, C. A. A methodology for developing simulation models of complex systems. *Ecological Modelling*, 202:385–396, 2007.

Aumann, C. A., L. A. Eby, and W. F. Fagan. How transient patches affect population dynamics: the case of hypoxia and blue crabs. *Ecological Monographs*, 76(3):415–438, 2006.

Balci, O. Principles of simulation model validation, verification, and testing. *Transactions of the Society for Computer Simulation International*, 14(1):3–12, 1997.

Couclelis, H. "where has the future gone?" rethinking the role of integrated land-use models in spatial planning. *Environment and Planning A*, 37:1353–1371, 2005.

Giampietro, M. *Multi-scale Integrated Analysis of Agroecosystems*. CRC Press, New York, 2004.

Gov. Canada. Integrated landscape management models for sustainable development policy making. `http://www.policyresearch.gc.ca/doclib/ILMM_Workshop_Report_e.pdf`, 2005.

Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science*, 310(11):987–991, November 2005.

Jakeman, A. J., R. A. Letcher, and J. Norton. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software*, 21:602–614, 2006.

Kline, S. J. *Conceptual Foundations for Multidisciplinary Thinking*. Stanford University Press, 1995.

Lempert, R., D. g. Groves, S. W. Popper, and S. C. Bankes. A general, analytic method for generating robust strategies and narrative scenarios. *Management Science*, 52(4):514–528, 2006.

Oreskes, N. and K. Belitz. Philosophical issues in model assessment. In Anderson, M. G. and lastname Bates, P. D., editors, *Model Validation: perspectives in hydrological science*, chapter 3, pages 23–41. John Wiley & Sons, New York, 2001.

Owens, S. Making a difference? some perspectives on environmental research and policy. *Transactions of the Institute for British Geographers*, 30:287–292, 2005.

Owens, S., T. Rayner, and O. Bina. New agendas for appraisal: reflections on theory, practice, and research. *Environment and Planning A*, 36:1943–1959, 2004.

Reynolds, J. H. and E. D. Ford. Multi-criteria assessment of ecological process models. *Ecology*, 80(2):538–553, 1999.

Rykiel Jr., E. J. Testing ecological models: the meaning of validation. *Ecological Modelling*, 90: 229–244, 1996.

Weber, M. and V. Adamowicz. Tradable land-use rights for cumulative environmental effects management. *Canadian Public Policy*, 28(4):581–595, 2002.

Wiegand, T., F. Jeltsch, I. Hanski, and V. Grimm. Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. *Oikos*, 100:209–222, 2003.

Zeigler, B. P., H. Praehoffer, and T. G. Kim. *Theory of Modeling and Simulation*. Academic Press, New York, 2 edition, 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

Evaluating consistency of biosphere models: software tools for a web-based service

G. A. Alexandrov and T. Matsunaga

Office for Global Environmental Database, Center for Global Environmental Research, National Institute for Environmental Studies, Onogawa 16-2, Tsukuba, Ibaraki 305-8506, Japan (g.alexandrov@nies.go.jp)

Abstract: A biosphere model is a geographical extension of an ecosystem model, and so modelling ecosystems at global scale we are facing the same problem as at the local scale -- structural uncertainty. The structural uncertainty includes competing conceptual frameworks, lack of agreement on model structure, ambiguous definitions of system boundaries, inadequate description of significant processes. The typical approach to this problem is assessing the maturity of the underlying science through retrospection of modelling efforts. This displays either consensus building or paradigm shift. In this paper we present a general scheme and software tools for performing such analysis, discuss how this scheme can be used for benchmarking a newly developed model, and specify a web-based service relevant to this purpose.

Keywords: Biosphere; Modelling; Model Consistency; Benchmarking; Web-based service

1. INTRODUCTION

A number of biosphere models has been developed in connection to the International Geosphere-Biosphere Program that backs the work of the Intergovernmental Panel on Climate Change. They are essential for projecting climate change scenarios and assessing mitigation and adaption options, however no criterion exists to pick one model over another [Cramer and Field, 1999]. Instead, the models form an ensemble, which is assumed to be a consistent estimator -- that is, to be converging to the quantities being estimated as the number of models grows. Is this assumption close to reality? Are confidence intervals shrinking? To answer this question we need regular checks of model consistency based on well-agreed methodology.

A biosphere model is a geographical extension of an ecosystem model, and so modelling ecosystems at global scale we are facing the same problem as at the local scale -- structural uncertainty. The structural uncertainty includes competing conceptual frameworks, lack of agreement on model structure, ambiguous definitions of system boundaries, inadequate description of significant processes [Manning et al, 2004]. The typical approach to this problem is assessing the maturity of the underlying science through retrospection of modelling efforts [Oikawa, 2007]. This displays either consensus building or paradigm shift.

In this paper we present a general scheme and software tools for performing such analysis, discuss how this scheme can be used for benchmarking a newly developed model, and specify a web-based service suitable to this purpose.

2. GENERAL SCHEME

2. 1 Methodological background

The evolution of scientific theories is often considered as a Darwinian process of natural selection that determines which theory survives and drifts them toward consensus [Bradie, 1994]. The concept of natural selection explains well

1. why well-established beliefs normally remain stable
2. and why independent researches addressing the same question ultimately come up with the same answer.

This does not imply, however, that an ultimate consensus is a legitimate goal. Stability of well-established beliefs stems from stability of research methods, and therefore it could be temporal.

2.2 Normative data

The well-established beliefs are based on normative data -- that is, the data supporting a judgement about what ought to be. The normative data represent the bulk of estimates obtained in previous researches, and in an ideal case, they characterize the maturity of knowledge in statistical terms -- by the confidence intervals of the mean values for the measured quantities. The estimates of independent studies are expected (Figure 1)

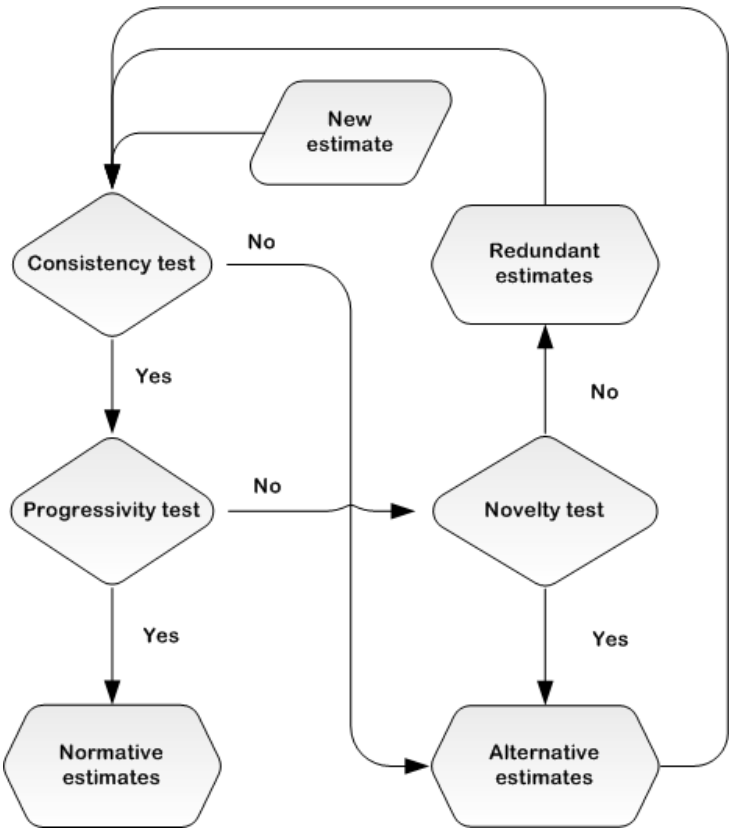


Figure 1. Flowchart for the process of building normative data

1. to fall within the bounds defined by normative data (consistency test),
2. to shift the mean value (novelty test),
3. and to narrow its confidence interval (progressivity test).

2.3 Alternative data

The process of "artificial selection" formalized above works against estimates suggesting too large shifts in mean values. These estimates form the pool of alternative data. With the passage of time, the bulk and consistency of alternative data may comprise to that of normative data, reflecting a paradigm shift.

Since the normative mean value is changing with every new estimate added to normative data, the pool of alternative data is re-processed and some estimates are moved to the pool of normative data. Therefore, the probability of sudden paradigm shift is quite low. This, however, does not completely remove the effect of order in which estimates appear, implying certain stability of beliefs originated from the pioneer studies.

2.4 Redundant data

The estimates that neither shift the normative mean value nor reduce its confidence interval form the pool of redundant data. This pool is also re-processed periodically and some estimates are moved to the pool of normative data.

3. SOFTWARE TOOLS

Program language

The software tools are written in *Mathematica* language [Wolfram, 1999] and arranged into a *Mathematica* package. The package contains the functions needed to perform tests (Figure 1) and to visualize the results (Figure 3).

Consistency test

Consistency test shows how far is the estimate (y) from the average ($\bar{x}$) in comparison to the lowest ($x_{\min}$) , or highest ($x_{\max}$) estimate:

$$z = \begin{cases} 100\dfrac{y-\bar{x}}{x_{\min}-\bar{x}}, & y < \bar{x} \\ 100\dfrac{y-\bar{x}}{x_{\max}-\bar{x}}, & y \geq \bar{x} \end{cases}$$

The test (z) is positive when estimate falls within the bounds defined by normative data ($x_{\min} \leq y \leq x_{\max}$), and negative otherwise. The estimate is said to be 100% consistent with the normative data if it coincides with the average estimate.

Progressivity test

Progressivity test is positive if inclusion of the estimate into normative data narrows the confidence interval of the average estimate. It returns the relative decrease in the width of the confidence interval:

$$z = 100\frac{x-y}{x}$$

where x is the original width of the confidence interval, y is the width of confidence interval changed due to inclusion of the estimate into the normative data.

Novelty test

Novelty test is positive if inclusion of the estimate into normative data shifts the average estimate. It returns the relative value of the shift with respect to the width of the confidence interval:

$$z = 100\frac{y-x}{u}$$

where x is the original mean value, y is the mean value changed due to inclusion of the estimate into the normative data, u is the half-width of the original confidence interval.

Case study

Net Primary Production (NPP) has been a focus of biosphere studies over the last three decades. First, the global pattern of NPP was characterized by the data collected during International Biological Program. Then, the data has been turned into empirical models that relate gradations in NPP to environmental factors of known geographic distribution such as mean annual temperature and precipitation (Miami model), actual evapotranspiration (Montreal model), and annually integrated NDVI. The ensemble (ToyBiMo) of three empirical models [Box et al., 1994] derived from the same NPP data shows high inconsistency of estimates in the case of larch forests and some other biomes (Figure 2). (Inconsistency is measured as half-width of the confidence interval of the normative mean value suggested by the model ensemble under concern.)

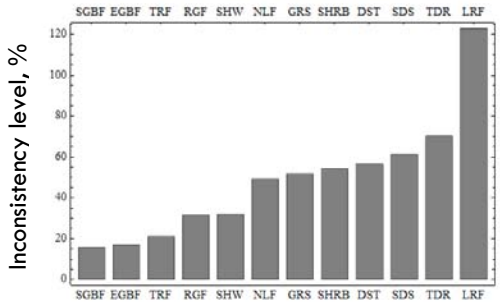


Figure 2. Inconsistency of the ToyBiMo ensemble. Legend: EGBF - evergreen broad-leaved forests, RGF - raingreen forests, TRF - tropical rainforests, SGBF - summer-green broad-leaved forests, SHW - subhumid woodlands, TDR - tundra, GRS - grasslands, NLF - needle-leaf forests, SDS - semi-desert scrubs, DST - deserts, SHRB - shrublands, LRF - larch forests.

A process-based model (TsuBiMo) calibrated with the same data [Alexandrov et al., 1992] reduces the ranges of uncertainty in the case of larch forest biome, but increase it in the case of the biome of evergreen broadleaf forest (Figure 3). Although the positive effects overweigh the negative effects, the case study illustrates well the fact that one can hardly expect an automatic reduction of uncertainties with every new model.

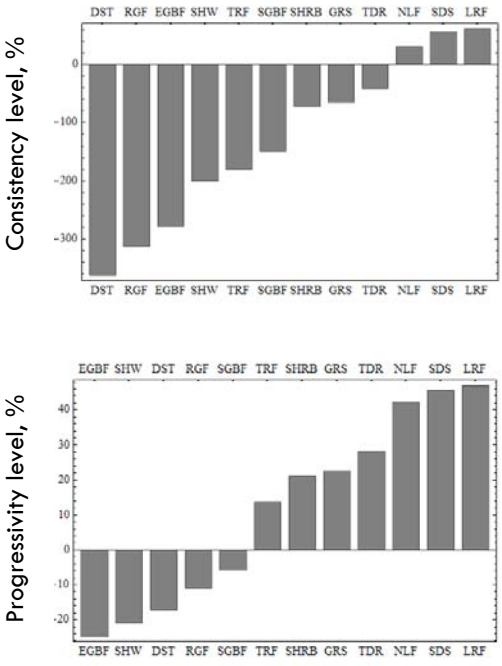


Figure 3. Consistency and progressivity charts: TsuBiMo vs ToyBiMo ensemble. Legend is the same as at the Fig 2.

The negative results of the consistency test stem from the fact that ToyBiMo suggests quite narrow range for NPP values, despite the wide confidence intervals. For example, in the case of evergreen broadleaf forests the highest and lowest values differ from average by 50 gC m^{-2} yr^{-1}. The width of confidence interval reflects the "sample size", which is rather small.

Such a narrow range for NPP values makes it difficult for a model to pass this consistency test, and so many of current models would be considered as alternative models with respect to the ToyBiMo ensemble. Hence, one may anticipate that retrospection of modelling efforts would detect a paradigm shift in judgement about what the productivity of some biomes ought to be. The results of the tests are visualized in ascending order to display the "problem biomes".

4. TOWARD A WEB-BASED SERVICE

4.1 Rationale

"In fact, the value of the software is proportional to the scale and dynamism of the data it helps to manage", [O'Reilly, 2005]. A retrospection of modelling efforts as well as benchmarking a new model requires a specialized database of model outputs. Without the data, the software tools are useless. The current version of the package includes some data. However, the database should be updated on regular basis, and the most efficient way of doing this is to employ the Web 2.0 "shiso" [Hayashi, 2007] that "aims to transform a society into an aggregated intelligence acting like a huge cyborg, by connecting people's individual intelligence (assumed as CPUs) through information and communications technology".

4.2 Technology

Since the software tools are written in *Mathematica* language, they can be used via the web interface [Wickham-Jones, 2006] to the *Mathematica* kernel. This interface, *webMathematica*, allows a web site to deliver JavaServer pages that call *Mathematica* commands. When the commands are evaluated, the computed result is placed in the page.

4.3 Specifications

In addition to the web interface mentioned above, the web-based service should include

1. a submission system for storing a model output in the specialized database
2. a reference system for getting information about the models available for retrospection analysis or benchmarking

4.4 Feasibility

The system for storing, retrieving and analysing 2-dimensional data as related to terrestrial carbon sink, recently developed by the Office for Global Environmental Database (OGED/CGER/NIES), includes all the components mentioned above. Therefore, feasibility of this web-based service depends mainly on the attitude of biosphere modellers, their inclination "to harness collective, net-enabled intelligence".

5. DISCUSSION

Modern computational and observational tools have caused an explosion of scientific data related to biosphere studies. This would improve eventually the consistency of biosphere models through creating multiple constraints for positioning 'true' values for model parameters. The limiting factor is thus the rate of building consensus on interpretation of the model outputs that will result from the new observations.

For example, terrestrial productivity is currently estimated at 60 GtC/yr. The consensus about this value was built in 1970s, although the estimates varied from 40 to 80 GtC/yr at that time, and re-analysis of the data (Alexandrov et al., 1999) revealed that estimates depend on how the data were classified with respect to the major regions of the world. In other words, this estimate is partly a social construction. Terrestrial productivity could be estimated at 50 or at 70 GtC/yr from the same observations.

The web-based service described above is to facilitate internalization of the new model results that may deviate from the existing consensus on biosphere characteristics. The scheme of building normative data simulates the natural process, but sets transparent criteria for distinguishing between normative and alternative data. It also suggests that an existing consensus should be re-considered when the bulk and consistency of alternative data comprise to that of normative data.

6. CONCLUSION

The biosphere models are essential for projecting climate change scenarios and assessing mitigation and adaption options. Many of them have been developed in connection to the International Geosphere-Biosphere Program (IGBP) that backs the work of the

Intergovernmental Panel on Climate Change (IPCC). Hence, there is certain demand for normative data on Net Primary Production (and its sensitivity to climate) as well as for normative data on other biogeochemical and eco-hydrological components of the climate system [Denman, et al., 2007].

This demand is partly satisfied by IPCC reports. Although one can hardly find the words "normative data" in IPCC reports, IPCC sets up certain norms by summarizing research results to produce policy-relevant recommendations. The recommendations are based on selected data, but the procedure of data selection is the Darwinian one: there are no explicit criteria of fitness.

There is nothing wrong in employing the method of "natural selection", with except to the risk of coming to an evolutionary deadlock. This risk can be significantly reduced through retrospection of modelling efforts, setting explicit criteria for distinguishing between normative and alternative data, and detecting paradigm shifts in timely fashion. OGED would spur these activities by providing a modern platform for processing the flows of data related to collaborative works of this sort.

ACKNOWLEDGEMENTS

The original manuscript was carefully revised in response to the valuable comments of Craig Aumann, Andrei Kirilenko, and Miquel Sanchèz-Marrè.

REFERENCES

Alexandrov, G.A., T. Oikawa and G. Esser, Estimating terrestrial NPP: what the data say and how they may be interpreted? *Ecological Modelling*, 117:361-369, 1999.

Alexandrov, G.A., T. Oikawa and Y. Yamagata, The scheme for globalization of a process-based model explaining gradations in terrestrial NPP and its application, *Ecological Modelling*, 148, 293-306, 2002.

Box, E.O., D.G. Dye, K. Fujiwara, R. Tateishi and X. Bai, Global environmental data sets from the Toyota Crown laboratory global engineering research project. Univ. of Tokyo, Tokyo, 1994.

Bradie, M., Epistemiology from an Evolutionary Point of View, In: Conceptual Issues in Evolutionary Biology (Ed.: E. Sober), MIT Press, 1994, 453-476.

Cramer, W. and C.B. Field, Comparing global models of terrestrial net primary productivity (NPP): introduction, *Global Change Biology*, 5, III-VII, 1999.

Denman, K.L., Brasseur, A., Chidthaisong, A., Ciais, P., Cox, R.E., Dickinson, D., Hauglustaine, D., Henze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., da Silva Dias, P.L., Wofsy, S.C., Zhang, X., Couplings Between Changes in the Climate System and Biogeochemistry. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, M., Tignor, M., and Miller, H.L. (eds.)], Cambridge University Press, Cambridge, 2007.

Hayashi, S., "Information and Communications Technology and Shiso", *Science & Technology Trends-Quarterly Review*, 23, 11-22, 2007 [http://www.nistep.go.jp/achiev/ftx/eng/stfc/stt023e/qr23pdf/STTqr2301.pdf]

Manning, M.R., M. Petit, D. Easterling, J. Murphy, A. Patwardhan, H-H. Rogner, R. Swart, and G. Yohe (Eds), IPCC Workshop on Describing Scientific Uncertainties in Climate Change to Support Analysis of Risk and of Options: Workshop report. Intergovernmental Panel on Climate Change (IPCC), Geneva, 2004.

Oikawa, T., Private retrospection on ecosystem model development. Invited lecture presented at the OGED seminar, NIES, Tsukuba, 24 August, 2007.

O'Reilly, T., "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software", 2005. [http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/ what-is-web-20.html]

Wickham-Jones, T., webMathematica: A User Guide, Wolfram Research, 2006.

Wolfram, S., Mathematica book, Cambridge University Press, Cambridge, 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

The Carbon-Land Model Intercomparison Project (C-LAMP): A Protocol and Evaluation Metrics for Global Terrestrial Biogeochemistry Models

Forrest M. Hoffman [a], **James T. Randerson**[b], **Inez Y. Fung**[c], **Peter E. Thornton**[d]
Yen-Huei "Jeff" Lee[d], **Curtis C. Covey**[e], **Gordon B. Bonan**[d], **and Steven W. Running**[f]

[a]*Oak Ridge National Laboratory (ORNL), Computer Science & Mathematics Division, Oak Ridge, Tennessee, USA (forrest@climatemodeling.org)*

[b]*University of California - Irvine, Department of Earth System Science, Irvine, California, USA (jranders@uci.edu)*

[c]*University of California - Berkeley, Department of Earth & Planetary Science, Berkeley, California, USA (ifung@berkeley.edu)*

[d]*National Center for Atmospheric Research (NCAR), P.O. Box 3000, Boulder, Colorado, USA (thornton@ucar.edu, jeff@ucar.edu)*

[e]*Lawrence Livermore National Laboratory (LLNL), Program for Climate Model Diagnosis and Intercomparison (PCMDI), Livermore, California, USA (covey1@llnl.gov)*

[f]*University of Montana, Numerical Terradynamic Simulation Group (NTSG), Missoula, Montana, USA (swr@ntsg.umt.edu)*

Abstract: Described here is a protocol and accompanying metrics for evaluation of scientific model performance of global terrestrial biogeochemistry models. Developed under the guise of the NCAR Community Climate System Model (CCSM) Biogeochemistry Working Group, the Carbon-Land Model Intercomparison Project (C-LAMP) experimental protocol improves and expands upon the Coupled Carbon Cycle-Climate Model Intercomparison Project (C^4MIP) Phase 1 protocol. However, unlike traditional model intercomparisons, C-LAMP has established scientific model performance metrics based upon comparison against best-available satellite- and ground-based measurements. Moreover, C-LAMP has partnered with the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison (PCMDI) to collect, archive, and distribute—via the Earth System Grid (ESG)—model results from C-LAMP experiments performed by international modeling groups in the same fashion as was done for the model results used in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). In addition, because future IPCC Assessment Reports are expected to be based on results from integrated Earth System Models (ESMs), C-LAMP is helping to establish the metadata standards for model output from terrestrial biogeochemistry components of ESMs. Proposed as an extension to the netCDF Climate and Forecast (CF) 1.1 Convention, these metadata standards will facilitate future model-model and model-measurement intercomparisons. A prototype diagnostics tool has been developed for C-LAMP that summarizes model results, produce graphical representations of these results as compared with observational data sets, and score models on their scientific performance.

Keywords: C-LAMP; model verification; model intercomparison; carbon cycle; terrestrial biogeochemistry

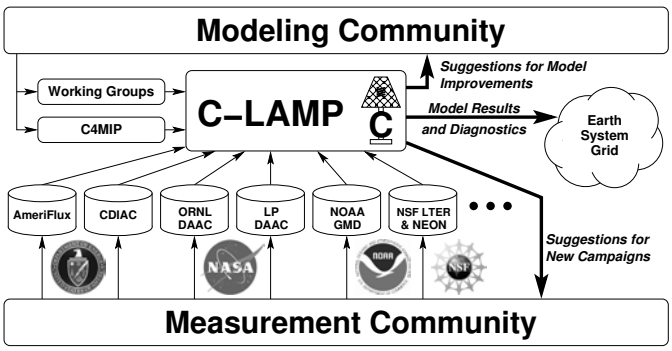


Figure 1: The Carbon-Land Model Intercomparison Project (C-LAMP) helps to bridge the gap between the measurement and modeling communities by comparing models against best-available observational data sets. C-LAMP provides feedback to both communities by offering suggestions for model improvements and by suggesting new measurement campaigns. All C-LAMP model results and diagnostics are distributed via the Earth System Grid (ESG).

1 INTRODUCTION

For the continued advance of climate change research it is particularly important for general circulation models (GCMs) to be extended to capture the global effects and feedbacks of carbon and other biogeochemical cycles. This need has resulted in new efforts to include atmospheric chemistry and land and ocean biogeochemistry into the next generation of climate models, now often referred to as Earth System Models (ESMs). While a number of terrestrial and ocean carbon models have been coupled to GCMs, recent work has shown that such models can yield a wide range of results [Friedlingstein et al., 2006]. This study suggests that a more rigorous set of offline and partially coupled experiments along with detailed analyses, including comparisons with measurements, are warranted.

The Carbon-Land Model Intercomparison Project (C-LAMP) provides a protocol and metrics for the intercomparison of terrestrial biogeochemistry models through a set of carefully crafted simulation experiments. Originally developed under the guise of the Community Climate System Model (CCSM) Biogeochemistry Working Group to test a number of such models within the CCSM3 framework [Hoffman et al., 2007], C-LAMP has been extended to include the larger international research community. Unlike traditional model intercomparisons, C-LAMP has established scientific model performance metrics based upon comparison against best-available satellite- and ground-based measurements. C-LAMP provides feedback to the modeling community by offering suggestions for model improvements and to the measurement community by suggesting new measurement campaigns. In addition, all model results will be made available through the Earth System Grid (ESG), the same system that distributed results used in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4).

2 THE C-LAMP PROTOCOL

2.1 Experiment 1: Uncoupled Simulations

Experiment 1 consists of uncoupled simulations of the terrestrial carbon model specifically designed to examine the ability of the models to reproduce surface carbon and energy fluxes at multiple sites and to examine the influence of climate variability, prescribed atmospheric carbon dioxide (CO_2), nitrogen (N) deposition, and land cover change on projections of terrestrial carbon fluxes during the 20^{th} century. These simulations are forced using an improved NCEP/NCAR reanalysis meteorology data set from Qian et al. [2005] that covers the years 1948–2004. The prescribed global atmospheric CO_2 is from the C^4MIP reconstruction of Friedlingstein et al. [2006], extended out to the year 2005. A nitrogen deposition climatology is used for the pre-industrial

simulations, while a time series is used starting in year 1890. Both of these data sets were developed as part of the SANTE FE project [Lamarque et al., 2005]. Historical global land cover data sets are those developed by Feddema et al. [2005] for climate change studies. The static land cover is that from the year 1798. Recently added were experiments designed to test the response of the models to a sudden increase in atmospheric CO_2 against the results from field measurements of the Free Air CO_2 Experiments (FACE) reported by Norby et al. [2005].

Initially, the protocol provided two equilibrium criteria for model spin up. These were 1) the absolute value of global land net ecosystem exchange (NEE) must be less than 0.05 PgC/y when taken as an average over a full 25-year repeat cycle of the meteorology drivers, and 2) the absolute value of NEE in every model grid cell must be less than 1.0 $gC/m^2/y$ when taken as an average over a full 25-year repeat cycle of the meteorology drivers. These criteria have proven to be too stringent; however, determining an adequate degree of equilibrium attainment is an open research question that is likely to be model dependent. Acceleration techniques are typically employed to reduce the simulation time required to reach an adequately spun up model state.

The Experiment 1 simulations are listed in Table 1. Experiments 1.1 and 1.2 are the spin up and control runs, respectively, and both cycle the first 25 years of the meteorology drivers with a fixed pre-industrial CO_2 concentration, climatological N deposition, and static land cover. Experiment 1.3 is initialized from year 1948 of the control run, and it uses the full meteorology time series to isolate the effect of varying only the climate. Experiment 1.4 begins in 1798 and includes the effects of varying climate, CO_2 concentration, and N deposition. Experiment 1.5 also begins in 1798 and adds the effects of historical land use change. Experiments 1.6 and 1.7 are the FACE control and transient simulations, respectively. They branch off Experiment 1.4, with static land cover, at year 1997 and extend out to year 2100, cycling the last 25 years of the meteorology drivers. Experiment 1.6 holds atmospheric CO_2 concentration and N deposition constant at year 2005 values, while Experiment 1.7 holds atmospheric CO_2 at 550 ppm, the nominal value from FACE.

2.2 Experiment 2: Partially Coupled Simulations

Experiment 2 consists of partially coupled simulations of the terrestrial carbon model with an active atmosphere model exchanging energy and moisture. In these experiments, atmospheric CO_2 is radiatively active and follows the prescribed historical trajectory used in Experiment 1. As in C^4MIP, the climate system is forced using sea surface temperatures (SSTs) and corresponding sea ice concentrations from the Hadley Centre for years 1875–2003, and extended to 2005 by Keith Lindsay. However, because of problems encountered in the Ocean Carbon Model Intercomparison Project (OCMIP) data set used in C^4MIP, prescribed ocean CO_2 fluxes come instead from an ocean simulation performed by Doney, *et al.* Fossil fuel emissions are annual estimates from the SRES A2 scenario, except in Experiment 2.6 where these emissions have been seasonalized to monthly values following the technique described by Erickson et al. [2008]. Because radiative CO_2 is prescribed, the CO_2 from land, ocean, and fossil fuel emissions are advected individually as inert tracers in the atmosphere.

The Experiment 2 simulations are listed in Table 2. Experiment 2.1 is the spin up run. It is initialized from the spun up state of Experiment 1.1, it cycles SSTs from years 1875–1899, and it uses a pre-industrial CO_2 forcing. Experiment 2.2 is the control simulation, run from 1800–2004, using the same forcing as Experiment 2.1. Experiment 2.3 also runs from 1800–2004, it adds the full time series of SSTs for years 1875–2004, and it employs the reconstructed time series for radiatively active CO_2 forcing. This experiment isolates the effects of changing only the climate. Experiment 2.4 is similar to Experiment 2.3, but it adds tracers for A2 fossil fuel emissions, ocean fluxes, and land NEE. In addition, a time series of atmospheric CO_2 concentration and N deposition forcing are applied to the land, so it includes the effects of varying climate, CO_2 concentrations, and N deposition. Experiment 2.5 adds the effects of historical land use change. Finally, Experiment 2.6 is like Experiment 2.4, but it uses seasonalized A2 fossil fuel emissions, instead of annual values, to demonstrate their effect on the seasonal cycle of atmospheric CO_2.

Table 1: The specifications for simulations contained in Experiment 1.

Exp	Time Period	NCEP/NCAR Forcing			Land [CO_2]				N Deposition			Land Cover	
		Cycling 1948–1972	Time Series	Cycling 1980–2004	Pre-Ind Constant	Time Series	2005 Time Series	FACE Constant	Clima-tology	Time Series	2005 Time Series	Constant	Time Series
1.1	As needed	✓			✓				✓			✓	
1.2	1798–2004	✓			✓				✓			✓	
1.3	1948–2004		✓		✓				✓			✓	
1.4	1798–1889	✓				✓			✓			✓	
	1890–1947	✓				✓				✓		✓	
	1948–2004		✓			✓				✓		✓	
1.5	1798–1889	✓				✓			✓				✓
	1890–1947	✓				✓				✓			✓
	1948–2004		✓			✓				✓			✓
1.6	1997–2100			✓			✓				✓	✓	
1.7	1997–2100			✓				✓			✓	✓	

Table 2: The specifications for simulations contained in Experiment 2.

Exp	Time Period	SST and Sea Ice		Radiative [CO_2]		Fossil Fuel		Ocean CO_2 Fluxes	Land [CO_2]		N Deposition		Land Cover	
		Cycling 1875–1899	Time Series	Pre-Ind Constant	Time Series	A2	A2S		Pre-Ind Constant	Time Series	Clima-tology	Time Series	Constant	Time Series
2.1	As needed	✓		✓					✓		✓		✓	
2.2	1800–2004	✓		✓					✓		✓		✓	
2.3	1800–1874	✓			✓				✓		✓		✓	
	1875–2004		✓		✓				✓		✓		✓	
2.4	1800–1874	✓			✓	✓		✓		✓	✓		✓	
	1875–1889		✓		✓	✓		✓		✓	✓		✓	
	1890–2004		✓		✓	✓		✓		✓		✓	✓	
2.5	1800–1874	✓			✓	✓		✓		✓	✓			✓
	1875–1889		✓		✓	✓		✓		✓	✓			✓
	1890–2004		✓		✓	✓		✓		✓		✓		✓
2.6	1800–1874	✓			✓		✓	✓		✓	✓		✓	
	1875–1889		✓		✓		✓	✓		✓	✓		✓	
	1890–2004		✓		✓		✓	✓		✓		✓	✓	

3 PERFORMANCE METRICS

C-LAMP has established model performance metrics that employ comparison against best-available satellite-, aerial-, and ground-based measurements. These metrics will continue to evolve as new and improved observational data sets become available. The metrics are targeted at examining the sensitivity of biogeochemical fluxes and pools to changes in driver variables, rather than the absolute value of those fluxes and pools, recognizing that pools are poorly constrained by observations and that fluxes critically depend on these pools and the physical climate within the model. For example, measurements of net primary production (NPP) normalized by precipitation are compared to model results normalized in the same way, to adjust for possible atmospheric model biases in precipitation. Similarly, Pearson's r correlation coefficient is computed between satellite observations and model distributions of NPP and leaf area index (LAI), which provides a measure of the correspondence of the variability of these quantities instead of the correspondence of actual values.

A number of NPP metrics have been defined. First, NPP from control runs is compared with NPP from the Ecosystem Model-Data Intercomparison (EMDI) Class A observations at 81 sites. Here actual values from entire model grid cells are compared against site observations, and the scale mismatch can contribute to biases and errors in determining model performance. Scatter plots of observed vs. modeled NPP are produced as diagnostics. Second, NPP normalized by precipitation is compared between the same EMDI observations and model results. This reduces effects of biases in the atmospheric model's hydrological cycle, but in some areas, NPP is limited by temperature (or the length of the growing season) and not by precipitation. Diagnostic plots of NPP vs. mean annual precipitation are produced to support this evaluation. Third, Pearson's r correlation coefficient is computed between MODIS (Moderate Resolution Imaging Spectroradiometer) MOD17 annual net primary productivity from Zhao et al. [2005] and model results, both globally and zonally by latitude. The former tests the models' ability to capture observed spatial variability, while the latter is designed to identify possible extra-tropical or tropical biases in model performance. Maps of global annual NPP and latitudinal zonal mean plots are generated for this evaluation.

Three LAI metrics have been established: correspondence with the annual mean, maximum, and phase (*i.e.*, month of maximum LAI) from MODIS MOD17 by land/biome class. Again, Pearson's r correlation coefficients are computed between satellite and model distributions. It is recognized that satellite-derived estimates of LAI are strongly dependent on atmospheric and canopy radiative transfer models that require validation, and biases in observations are likely to impact model performance scores. Phase should be less sensitive to these types of potential biases. Maps of LAI annual mean, maximum, and phase from both the observations and the model results, as well as difference maps, are produced to support LAI evaluations.

Metrics for the seasonal cycle of atmospheric CO_2 test the combined effects of the seasonal timing and magnitude of NPP and heterotrophic respiration in northern hemisphere biomes. Good performance provides some confidence in the temperature sensitivity of respiration for those biomes in the model and suggests that prognostic leaf area, and thus gross primary production (GPP) and NPP, is being simulated correctly. The observations are obtained from the NOAA Globalview data set of measurements from surface stations. Correspondence in latitudinal zones as well as correlation and amplitude ratios at individual sites are tested as part of these metrics. Since ocean and fossil fuel fluxes contribute only weakly to the CO_2 seasonal cycle, it serves as a good diagnostic of biosphere-atmosphere exchange in northern hemisphere ecosystems. However, biosphere fluxes from the model require a model of atmospheric transport, and biases in the horizontal or vertical mixing in the transport model can influence performance for these metrics. Plots of monthly mean observations vs. model results for latitudinal zone and for individual stations are generated.

Measurements of carbon stocks are very limited, but a recent estimate of above-ground live biomass in the Amazon basin by Saatchi et al. [2007] provides one set of observations useful for comparison with model results. For this metric, both the total above-ground live biomass and its spatial pattern are compared between observations and model results. Maps of biomass from

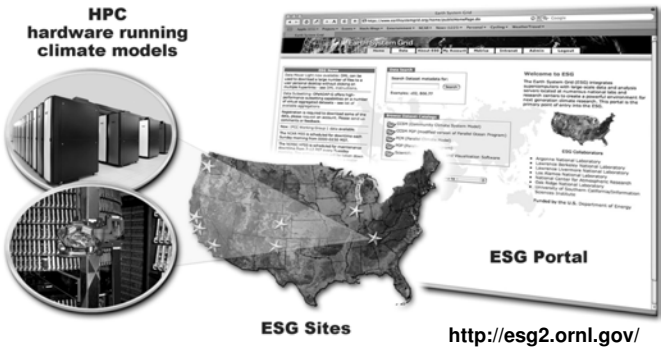


Figure 2: The Earth System Grid (ESG).

observations and model results and difference maps are produced for this evaluation.

The wide deployment of eddy covariance flux towers offers the opportunity to constrain modeled fluxes of latent and sensible heat, net radiation, and CO_2 across a diversity of world biomes. Measurements from the Ameriflux sites are readily available, and these are used to evaluate model performance for latent heat, sensible heat, NEE, GPP, and ecosystem respiration (ER). Models are scored for correspondence with each of these factors across 74 sites. Plots containing measured and modeled results for each of these quantities for every site are generated.

Additional diagnostics of transient dynamics are produced to further characterize model behavior. These include calculations of turnover times for leaf, wood, fine root, litter, coarse woody debris, and soil carbon pools by biome type and tables of carbon stocks and fluxes meant to elucidate responses to El Niño phenomena. As additional observational data sets become available, they can be added to the C-LAMP diagnostics to improve its utility in evaluating performance of biosphere models.

4 THE EARTH SYSTEM GRID (ESG)

The model output from the C-LAMP experiments will be made available to the wider international research community through the Earth System Grid Center for Enabling Technologies (ESG-CET) [Ananthakrishnan et al., 2007]. The Earth System Grid (ESG) (`http://www.earthsystemgrid.org/`) is a large, production, distributed system that allows registered users to download model output, code, and ancillary data over the Internet [Bernholdt et al., 2005] Two ESG Portals have been established, and a new one has been deployed at ORNL to support C-LAMP (see Figure 2). PCMDI is assisting in the deployment of this server at ORNL, which will archive and distribute the standard model output fields resulting from C-LAMP. With over 6,000 registered users and more than 250 TB of data, ESG was the primary means for distribution of IPCC data that resulted in over 300 scientific publications supporting AR4 [IPCC, 2007].

C-LAMP is leading an effort to develop metadata standards for terrestrial biosphere model output. These standards will be needed to support IPCC AR5 model results since biogeochemistry and atmospheric chemistry are likely to be included in the new Earth System Models (ESMs) participating in the main simulations. In particular, proposals are being developed to extend the netCDF Climate and Forecast (CF) metadata conventions [Eaton et al., 2008] to include better representation of common biosphere model output fields.

5 CONCLUSIONS

The C-LAMP experiments provide a means for rigorous testing and intercomparison of terrestrial biogeochemistry models. A growing number of C-LAMP metrics can be used to suggest where, when, and why such models exhibit deficiencies. Tracking changing scores offers a quantitative means for measuring model improvements. A diagnostics tool has been developed that summarizes model results, produces graphical representations of the model results as compared with observational data sets, and automatically scores models based on their scientific performance. This tool may be extended in the future to provide a user-friendly method for modelers to test and score their own model results prior to contributing them to ESG. C-LAMP is an open, community project that benefits from the suggestions and input of community researchers. More information about C-LAMP, the experimental protocol, model performance metrics, and early results from participating models is available at http://www.climatemodeling.org/c-lamp.

ACKNOWLEDGMENTS

This research was partially sponsored by the Climate Change Research Division (CCRD) of the Office of Biological and Environmental Research (OBER) and the Computational Science Research and Partnerships (SciDAC) Division of the Office of Advanced Scientific Computing Research (OASCR) within the U.S. Department of Energy's Office of Science (SC). 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. The Lawrence Livermore National Laboratory is managed by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy under Contract No. DE-AC52-07NA27344. The National Center for Atmospheric Research (NCAR) is operated by the University Corporation for Atmospheric Research (UCAR) and receives research funding primarily from the U.S. National Science Foundation (NSF).

REFERENCES

Ananthakrishnan, R., D. E. Bernholdt, S. Bharathi, D. Brown, M. Chen, A. L. Chervenak, L. Cinquini, R. Drach, I. T. Foster, P. Fox, D. Fraser, K. Halliday, S. Hankin, P. Jones, C. Kesselman, D. E. Middleton, J. Schwidder, R. Schweitzer, R. Schuler, A. Shoshani, F. Siebenlist, A. Sim, W. G. Strand, N. Wilhelmi, M. Su, and D. N. Williams. Building a global federation system for climate change research: The Earth System Grid Center for Enabling Technologies (ESG-CET). *Journal of Physics: Conference Series*, 78:012050 (7pp), December 2007. doi:10.1088/1742-6596/78/1/012050.

Bernholdt, D., S. Bharathi, D. Brown, K. Chanchio, M. Chen, A. Chervenak, L. Cinquini, B. Drach, I. Foster, P. Fox, J. Garcia, C. Kesselman, R. Markel, D. Middleton, V. Nefedova, L. Pouchard, A. Shoshani, A. Sim, G. Strand, and D. Williams. The Earth System Grid: Supporting the next generation of climate modeling research. *Proc. IEEE*, 90(3):485–495, March 2005. doi:10.1109/JPROC.2004.842745.

Eaton, B., J. Gregory, B. Drach, K. Taylor, S. Hankin, J. Caron, and R. Signell. NetCDF Climate and Forecast (CF) metadata conventions. Technical Specifications Version 1.1, January 2008. http://cf-pcmdi.llnl.gov/documents/cf-conventions.

Erickson, D. J., R. T. Mills, J. Gregg, T. J. Blasing, F. M. Hoffman, R. J. Andres, M. Devries, Z. Zhu, and S. R. Kawa. An estimate of monthly global emissions of anthropogenic CO_2: Impact on the seasonal cycle of atmosheric CO_2. *J. Geophys. Res.*, 113(G01023), March 2008. doi:10.1029/2007JG000435.

Feddema, J. J., K. W. Oleson, G. B. Bonan, L. O. Mearns, L. E. Buja, G. A. Meehl, and W. M. Washington. The importance of land-cover change in simulating future climates. *Science*, 310 (5754):1674–1678, December 2005. doi:10.1126/science.1118160.

Friedlingstein, P., P. M. Cox, R. A. Betts, L. Bopp, W. von Bloh, V. Brovkin, S. C. Doney, M. Eby, I. Fung, B. Govindasamy, J. John, C. D. Jones, F. Joos, T. Kato, M. Kawamiya, W. Knorr, K. Lindsay, H. D. Matthes, T. Raddatz, P. Rayner, C. Reick, E. Roeckner, K.-G. Schnitzler, R. Schnur, K. Strassmann, S. Thompson, A. J. Weaver, C. Yoshikawa, and N. Zeng. Climate-carbon cycle feedback analysis, results from the C^4MIP model intercomparison. *J. Climate*, 19 (14):3373–3353, July 2006. doi:10.1175/JCLI3800.1.

Hoffman, F. M., C. C. Covey, I. Y. Fung, J. T. Randerson, P. E. Thornton, Y.-H. Lee, N. A. Rosenbloom, R. C. Stöckli, S. W. Running, D. E. Bernholdt, and D. N. Williams. Results from the Carbon-Land Model Intercomparison Project (C-LAMP) and availability of the data on the Earth System Grid (ESG). *Journal of Physics: Conference Series*, 78:012026 (8pp), December 2007. doi:10.1088/1742-6596/78/1/012026.

IPCC. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Intergovernmental Panel on Climate Change, Cambridge, United Kingdom and New York, NY, USA, 2007.

Lamarque, J.-F., J. T. Kiehl, G. P. Brasseur, T. Butler, P. Cameron-Smith, W. D. Collins, W. J. Collins, C. Granier, D. Hauglustaine, P. G. Hess, E. A. Holland, L. Horowitz, M. G. Lawrence, D. McKenna, P. Merilees, M. J. Prather, P. J. Rasch, D. Rotman, D. Shindell, and P. E. Thornton. Assessing future nitrogen deposition and carbon cycle feedback using a multimodel approach: Analysis of nitrogen deposition. *J. Geophys. Res.*, 110(D19303), December 2005. doi:10.1029/2005JD005825.

Norby, R. J., E. H. DeLucia, B. Gielen, C. Calfapietra, C. P. Giardina, J. S. King, J. Ledford, H. R. McCarthy, D. J. P. Moore, R. Ceulemans, P. D. Angelis, A. C. Finzi, W. H. Schlesing, and R. Oren. Forest response to elevated CO_2 is conserved across a broad range of productivity. *Proc. Nat. Acad. Sci.*, 102(50):18052–18056, December 2005. doi:10.1073/pnas.0509478102.

Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson. Simulation of global land surface conditions from 1948–2002. Part I: Forcing data and evaluation. *J. Hydrometeorology*, 7(5), October 2005. doi:10.1175/JHM540.1.

Saatchi, S. S., R. A. Houghton, R. C. D. S. Alvala, J. V. Soares, and Y. Yu. Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology*, 13(4):816–837, April 2007. doi:10.1111/j.1365-2486.2007.01323.x.

Zhao, M., F. A. Heinsch, R. R. Nemani, and S. W. Running. Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, 95(2):164–176, March 2005. doi:10.1016/j.rse.2004.12.011.

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

Reducing Uncertainty in Ecological Niche Models with ANN Ensembles

A. P. Kirilenko[a], R. S. Hanley[b]

[a] *Department of Earth Systems Science and Policy, University of North Dakota, Grand Forks, ND 58202-9011, USA. 1-701-777-6761, 1-701-777-2940 (fax), andrei.kirilenko@und.edu*
[b] *Department of Earth Systems Science and Policy, University of North Dakota, Grand Forks, ND 58202-9011, USA. 1-701-777-3909, 1-701-777-2940 (fax), rshanley@aero.und.edu*

Abstract: Transformation of spatial distributions of species is a key element for understanding global and regional impact of climate change on environment in the future. Yet little is known about current distributions of species, especially in relation with physical parameters of environment. This is primarily due to limited availability of data on species occurrences, which prevents reconstruction of their geographical distribution with standard methods of spatial statistics. To address this problem, specialized applications called distributional models, based on the hypothesis of fundamental niche, are being developed. We have built a new tool for predicting species' geographical distributions based on presence-only data. The tool is developed around an Artificial Neural Network (ANN) simulation engine, which employs the Stuttgart Neural Network Simulator (SNNS). Our tool is capable of using multiple environmental layers as predictors to generate the patterns at the species' presence and pseudo-absence localities, selecting and training an ANN using these patterns, selecting the optimal ANN, testing the selected ANNs on independent sets of data, applying the selected model to project species distribution at current or modified climate conditions, and porting the resultant presence probability maps to ESRI ArcGis. One of the frequent criticisms with ANN-based computer applications is connected with their generalization ability, demonstrated through "model overtraining". When the pre-selected number of training steps for the model is too high, the model-generated predictive surface looses its smoothness, becoming too tightly locked at the training dataset. To reduce the generalization abilities of ANN predictions of insect species' occurrences, we employed model ensembles. The redundant ensemble of ANN models was used to simulate spatially distributed probabilities of species presence from the train presence data set; the results of modeling with each individual network were then pooled together using their linear combination. Such modeling based on employing redundant model ensembles yielded a significant improvement in overall model performance and reduced model-related uncertainty.

***Keywords**:* Artificial Neural Networks, NeuroNiche, Model Ensembles, Ecological Niche, Model Uncertainty.

1. INTRODUCTION

Modeling the impact of environmental change on a species' geographical location is frequently an attempt to estimate the transformations of the species' ecological niche. The concept of the ecological niche, introduced by Hutchinson (1958) suggests that a species needs a certain combination of the environmental factors, such as the temperature, precipitation, soil, etc., for its successful survival and reproduction. Mathematically, the concept of the fundamental species' niche can be formalized by considering a mapping μ

from 2-dimensional geographical space X into a multidimensional space of physical factors important for the species' existence P, i.e. μ: X → P. The projection of the species' known localities then constitutes the species' fundamental niche. Once the fundamental niche is found, the impact of the changed environmental conditions on geographical distribution of the species can be found from the inverse mapping μ^{-1}: P → X, assuming it exists. Practically, however, the number of known species' localities (i.e., the species' presence points) is never large enough to make an inference about the fundamental niche without of making additional assumptions. To address this problem, specialized applications called distributional models, based on the hypothesis of fundamental niche, are being developed. First, a set of spatially distributed data layers believed to determine species geographical distribution is defined based on biological considerations. Then, model parameters are being computed to provide the best match between the predicted and observed species presence data. Finally, the model is applied to the entire set of original or modified environmental layers to project species distribution in the areas with missing presence data and/or under the impact of environmental change.

A number of statistical and Artificial Intelligence methods, such as Maximum Entropy (MaxEnt, by Phillips et al., 2006), Bioclimate Prediction System (BIOCLIM, by Nix, 1986), and Genetic Algorithm for Rule-set Prediction (GARP, by Stockwell & Peters, 1999) have been used for estimating the species' fundamental niche and prediction of the the geographical transformations of the species' localities under the environmental change. Each of these methods has its own limitations. The recently developed Bayesian probabilistic method MaxEnt has probably become the most frequently used approach for modeling species spatial predictions. The rationale for this popularity is MaxEnt's ability to produce compact predictive distribution for probabilities of species' occurrence with performance tests presumably superior to other modeling techniques. However, with one method clearly dominating the predictions, modeling uncertainty of predicted species distribution is becoming harder to estimate.

We introduce a new tool for modeling predicting the species' geographical distributions, which is based on the Artificial Neural Networks (ANNs). The ANNs, originally designed to mimic signal processing in a brain, nowadays are increasingly used in a multitude of different applications. Essentially, an ANN presents a system for parallel information processing that consists of a large number of simple adaptive elements. The large degree of interconnection between the elements in the ANN leads to its complex, non-linear behaviors, and the ability of the elements (neurons) to adapt permits the "learning by example" type of behavior, as the system re-builds itself so that the processed input information would closely match the provided training data.

The ANNs are capable of approximating a measurable function with any arbitrary set degree of accuracy (Hornik et al., 1989). As a consequence, an ANN-based model has an extremely good predictive ability, yet the downside is loosing the model generalization ability, the process, which is commonly referred to as "overtraining". A widely used method of finding an acceptable balance between underfitting and overtraining of the ANN-based model (Poulton 2001) requires a small separate set of the testing data. Then, after a pre-set number of training epoch, the ANN is validated on this testing data set, and the training is allowed to continue as long as the error of validation is decreasing. However, for the insect presence data, the limited number of recorded locations for rare species make this method of overtraining control less useful.

Additionally to overfitting, the stochastic nature of the ANN training process leads to the substantially different results returned by the ANNs, even when they are train on identical data. To reduce the structural uncertainty of the model and to increase its generalization ability, we take an advantage of using an ensemble of the ANN models, post-processing the model results. We further illustrate our approach on modeling the geographical distribution of the bumblebee *Bombus pennsylvanicus* (Degeer).

2. NEURONICHE: A NEW TOOL BASED ON THE ARTIFICIAL NEURAL NETWORKS

2. 1 Neuroniche

We have developed a new tool for predicting species' geographical distributions based on the species presence data we call NeuroNiche. The tool is built around an Artificial Neural Network (ANN) simulation engine Stuttgart Neural Network Simulator (SNNS, 2007). NeuroNiche is capable of using multiple environmental layers as predictors to generate the patterns at the species' presence and pseudo-absence localities, selecting and training an ANN using these patterns, selecting the optimal ANN, testing the selected ANNs on independent sets of data, applying the selected model to project species distribution at current or modified climate conditions, and porting the resultant presence probability maps to ESRI ArcGis. The NeuroNiche uses a feedforward neural network with one hidden layer, which is the most popular type of ANN in the majority of applications. Further, we set the number of the elements (neurons) in the hidden layer N equal to the number of inputs to the ANN. For the learning function, we used the Std_Backpropagation function of SNNS with the weights set to random. The number of training cycles can be arbitrary, however after experimenting with multiple sets of the species location data we found the optimal number of training cycles to be between 2000 and 10000.

The data on insect occurrences tend to contain only the presence records. The so-called "profile methods" (Pearce and Boyce, 2006) of computing the species geographical distribution that require only the species presence data are common. The most popular in this category is the "climatic envelope" ANUCLIM method (McMahon et al., 1996) as implemented in the widely used BIOCLIM (Busby, 1991) software. The presence-only data distributional models successfully employ a variety of methods to envelope the projections of the known presence locations to the multidimensional space of environmental parameters. Yet the common criticism of the profile-based models is that the resultant distribution is "too wide" as the models tend to over-predict the species occurrence and may be highly sensitive to outliers (e.g., BIOCLIM - Pearce, Boyce, 2006).

Alternative to the profile methods are the methods the contrast the environmental conditions in the known species presence locations with the environmental conditions in a random sample of geographical locations, called the "pseudo-absence" locations. Neuroniche is able of generating a sample of uniformly distributed pseudo-absences or it can generate the sample using a custom probability density function. In our study, we generated the pseudo-absence samples using the inverse of the normalized density of the presence locations, generated in the ESRI ArcMap software.

2. 2 Case study: geographical distribution of bumblebee *Bombus pennsylvanicus*

To illustrate the application of NeuroNiche to finding the fundamental niche of a species, we apply the model to mapping the geographical distribution of *Bombus pennsylvanicus*. Bumblebees in general represent an ideal group for the analysis because their distributional data is more-or-less abundant, their taxonomy in North America is well understood, they play a critical ecological role across the continent as pollinators, and they are highly sensitive to the well-mapped environmental factors, e.g. to climate. The distributional data of the bumblebees examined were obtained from the Global Biodiversity Information Facility (http://www.gbif.org/), and included 107 unique occurrence points collected throughout the continent for the three species (figure 1).

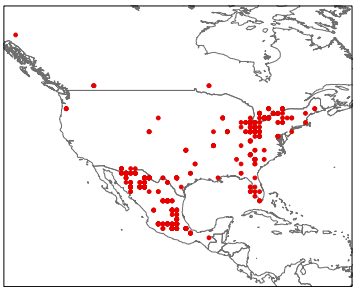


Figure 1. The data on the bumblebee *Bombus pennsylvanicus* presence.

We divided the location data into the training and testing sets of equal size. Further, we followed the methodology described above to generate ten sets of pseudo-absence data,

1000 points each, with distribution determined by the normalized density of presence points (figure 2). These redundant sets of the presence-absence data were used for model training and testing.

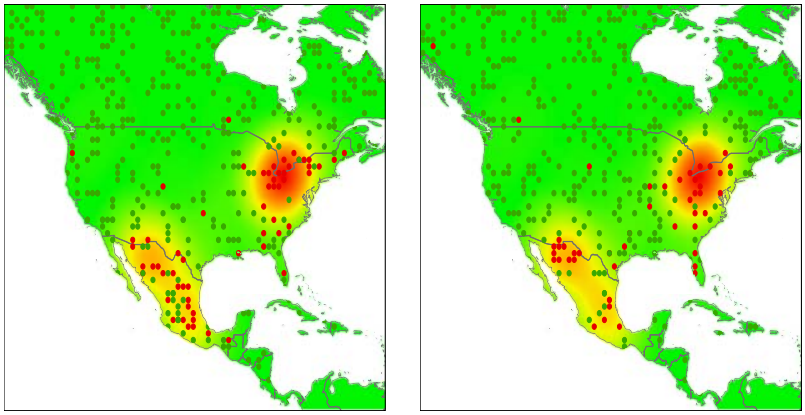


Figure 2. The species location database, divided into the training (left) and testing (right) sets, with random pseudo-absence points added. Presence points are shown in red; pseudo-absence – with green.

2. 3 Environmental data

As a base climate, we used the 1961-1990 monthly values for temperature and precipitation (Mitchell et al., 2004). These values were further processed using stochastic weather generator (Friend, 1988) to obtain the following seven climate layers for the North America: mean annual temperature and precipitation for the reference period, summer temperature, winter temperature, mean values of the highest and lowest temperatures through a year, and the length of frost period. All modeling in this study require that distributional points be placed within a continuous, gridded ecological context. To do this, we used topographical data (elevation, slope, aspect, flow accumulation, flow direction, and compound topographical index) from the US. Geological Survey's (2008) HYDRO-1k data set. All data were gridded at 0.1 degree spatial resolution for analysis. In total, the following 12 environmental layers were used:

- Annual mean temperature
- Total annual precipitation
- DJF temperature
- JJA temperature
- Length of thaw free period
- Highest temperature (expectation)
- Lowest temperature (expectation)
- Elevation
- Slope
- Aspect
- Flow accumulation
- Compound topographical index

2. 4 Model ensembles

One of the frequent criticisms with ANN-based computer applications is connected with their generalization ability, demonstrated through "model overtraining" (Sharkey, 1999). When the pre-selected number of training steps for the model is too high, the model-generated predictive surface looses its smoothness, becoming too tightly locked at the training dataset. The apparent result of model overtraining is a "grainy" structure of simulated spatial distribution of the probability of species occurrence. Paradoxically, the more time is spent for training the model, the poorer results will it demonstrate on testing data. The traditional method to increase model generalization ability is early termination of model training, which however increases model-related uncertainty of prediction. To

reduce the generalization abilities of ANN predictions of insect species' occurrences, we employed model ensembles (Sharkey, 1999). We used randomly generated seed values to create a set of ten ANN models, which we then trained on the same insect presence data. This redundant ensemble of ANN models was then used to simulate spatially distributed probabilities of species presence from the test presence data set. Finally, the results of modeling with each individual network were pooled together using their linear combination.

3. RESULTS

NeuroNiche performance has demonstrated a satisfactory model performance as estimated from receiver operating characteristic (ROC) sensitivity vs. 1-specifisity plots (table 1). As a single estimator for model quality, we used the area under the ROC curve (AUC), a measure, which is traditionally used in signal detecting (Swets, Pickett, 1982), and shown to be equivalent to Mann-Whitney U test. We found ROC to vary between 0.95 and 0.98 (mean value 0.97) for training data, and between 0.72 and 0.87 (mean value 0.83) for testing data set when the models with randomly selected input parameters were trained using backpropagation algorithm for an arbitrary pre-selected 5000 steps each. When the number of model training steps was selected based on the best performance demonstrated by the model during the test, AUC would increase to 0.85-0.87. These results are similar to the ones we received when using MaxEnt method with same dataset.

Table 1. The area under the ROC curve (AUC), computed to estimate the overall performance of the model trained for 5000 and 10000 cycles. Notice that using the redundant model ensembles considerable improves model performance on testing data.

Experiment	Train 5000	Test 5000	Test 10000
1	0.95	0.74	0.77
2	0.97	0.75	0.71
3	0.96	0.81	0.69
4	0.97	0.87	0.7
5	0.96	0.79	0.69
6	0.96	0.72	0.71
7	0.94	0.87	0.75
8	0.96	0.75	0.78
9	0.96	0.74	0.64
10	0.96	0.8	0.72
Mean	**0.959**	**0.784**	**0.716**
Ensemble	**0.98**	**0.83**	**0.82**

Despite the satisfactory performance of each of the generated ANN – based models, the projections of the individual models demonstrated similar, yet different in details, results (figure 3). For demonstration purposes, model training was interrupted early, after 2000 cycles. The projected probability of species occurrence exhibit two peaks with the highest possibility of locating the species. The maps also demonstrate a considerable noise, which is the evidence of the average generalization ability of each of the ANNs.

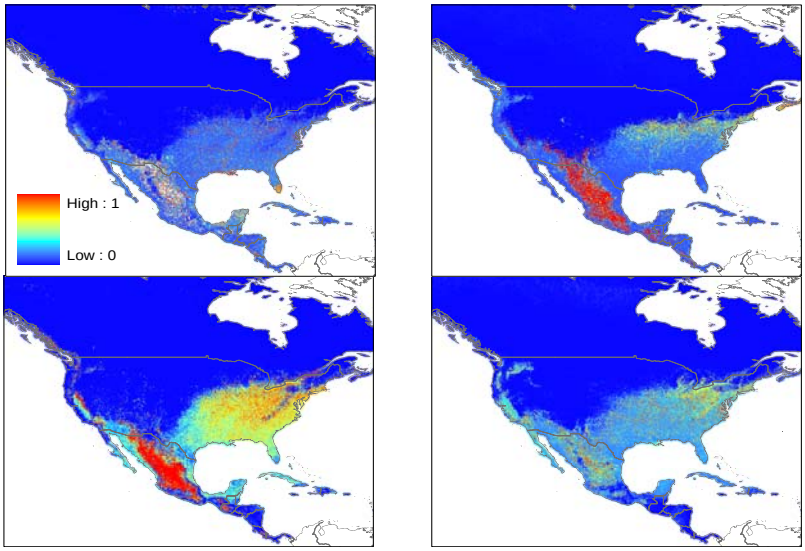


Figure 3. An example of variability in projections of ANN models, trained for 2000 cycles with difference seed values and using different sets of pseudo-absence data.

The results of each of the individual models was be significantly improved by employing redundant model ensembles. The AUC value for a model trained for 5000 steps has increase from 0.78 on average up to 0.83. The performance for slightly overtrained models increased even further: e.g., the AUC of the models trained for 10000 steps vary from 0.64 to 0.77 (mean 0.72) when model results were compared to test data. Same set of models, taken as a model ensemble, has increased AUC up to 0.82.

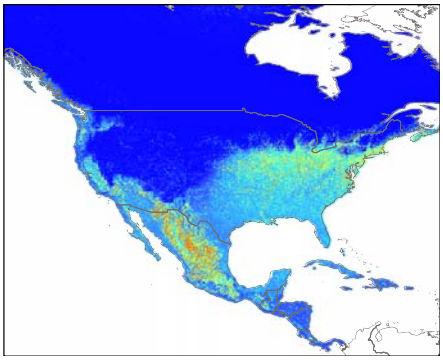


Figure 4. A projection of *Bombus pennsylvanicus* geographical distribution, obtained as a linear combination of the projections of ten ANN models (see figure 3 to compare with the projections of individual models).

4. CONCLUSIONS

We introduced a new approach to model insect species' spatial distributions based on the hypothesis of the fundamental ecological niche. The approach employs species presence-only data and pseudo-absence data to train an ANN using the backpropagation algorithm. We used the model to predict the continental-scale distribution of *Bombus pennsylvanicus* using a medium-size training dataset of 53 locality records. Model performance was measured using ROC estimator with a testing dataset containing 54 locality points and was found to be satisfactory. We also tested the model with a reduced volume of training dataset to 20 points with equally satisfactory results.

To improve model generalization ability and further reduce the prediction error, we utilized a redundant ensemble of ANN models. The projection of each individual ensemble member

was different from the others due to the variations in both the input data for model training and in the model seed values. Between multiple methods for combining the ensemble members (Ahmad, Zhang, 2002; Shu and Burn, 2004) we selected the most frequently used linear combination approach. The previous studies (Hansen, Salamon, 1990; Opitz, Maclin, 1999) has demonstrated that a model ensemble with as few as ten ANNs yields a significant improvement in the overall model performance. Following these studies, we employed an ensemble of ten ANNs, and found both model generalization and model error significantly improved.

Several complications to our analysis should be mentioned. First, we used simplified approaches for both creating and combining the results of the ensemble members; more elaborate methods such as ANN boosting and model stacking should further improve NeuroNiche performance. Second, relatively small volume of the presence locality dataset was one of the reasons for the significant variability between the projections of the individual models. Finally, the selection of bumblebees for the analysis sets a considerable restriction on the predictive ability of a species distribution model. The bumblebees are closely associated with particular flowering plants as a source of nectar. This analysis did not examine the distributions of those plant species, which almost certainly are limiting factors in determining the distributions of bumblebees. These considerations suggest that substantial work remains to be done in improving the projections of the ecological niche models using ANN ensembles.

ACKNOWLEDGEMENTS

This research was supported through a NASA Applied Sciences Program Grant (#NNX06AE16G) awarded to the authors.

REFERENCES

Ahmad, Z., and J. Zhang, A comparison of different methods for combining multiple neural networks models, Proceedings of the 2002 World Congress on Computational Intelligence, IEEE Press, Piscataway, N.J., 12–117, 2002.

Busby, J.R., BIOCLIM - A bioclimatic analysis and prediction system. In: Margules, C.R., and M.P. Austin (eds.), Nature conservation: cost effective biological surveys and data analysis, Canberra, CSIRO, 64–68, 1991.

Friend, A.D., Parameterization of a global daily weather generator for terrestrial ecosystem modeling, *Ecological Modelling*, 109, 121–140, 1998.

Hansen, L., and P. Salamon, 1990, Neural network ensembles, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12, 993–1001.

Hutchinson, G.E., Concluding remarks, Cold Spring Harbor Symposium, *Quantitative Biology*, 22, 415–427, 1958.

Hornik, K., M. Stinchcombe, and H. White, Multilayer feed forward networks are universal approximators, *Neural Networks*, 2, 359–366, 1989.

McMahon, J.P., M.F. Hutchinson, H.A. Nix, and K.D. Ord, ANUCLIM version 1 user's guide, Canberra, ANU, CRES, 1996.

Mitchell, T.D., T.R. Carter, P.D. Jones, M. Hulme, and M. New, A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100), Tyndall Centre for Climate Change Research, Working Paper 55, 2004.

Nix, H.A., A biogeographic analysis of Australian elapid snakes, Atlas of elapid snakes of Australia. Australian Flora and Fauna Series 7, Australian Government Publishing Service, Canberra, 1986.

Opitz, D., and R. Maclin, Popular ensemble methods: an empirical study, *Journal of Artificial Intelligence Research*, 11, 169–198, 1999.

Pearce, J.L., and M.S. Boyce, Modelling distribution and abundance with presence-only data, *Journal of Applied Ecology*, 43, 405–412, 2005.

Phillips, S.J., R.P. Anderson, and R.E. Schapire, Maximum entropy modeling for species geographic distributions, *Ecological Modelling*, 190, 231–259, 2006.

Poulton M.M. Computational neural networks for geophysical data processing, Pergamon, 2001.

Sharkey, A.J.C. (Ed.), Combining artificial neural nets: ensemble and modular multi-net systems, Springer-Verlag, New York, 1999.

Shu, C., and D.H. Burn, Artificial neural network ensembles and their application in pooled flood frequency analysis, *Water Resources Research*, 40, W09301, doi:10.1029/2003WR002816, 2004.

SNNS - Stuttgart Neural Network Simulator V4, Manual. Internet. http://www-ra.informatik.uni-tuebingen.de/SNNS (last accessed January 14, 2007).

Stockwell, D.R.B., and Peters, D.P., The GARP modelling system: Problems and solutions to automated spatial prediction, *International Journal of Geographic Information Systems*, 13, 143–158, 1999.

Swets, J.A., and R.M. Pickett, Evaluation of Diagnostic Systems—Methods from Signal Detection Theory, Academic Press, 1982.

U.S. Geological Survey, HYDRO-1k. Internet. http://edc.usgs.gov/products/elevation /gtopo30/hydro/index.htm (last accessed January 14, 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

Towards model evaluation using Self-Organizing Maps

M. Herbst and M. C. Casper

Department of Physical Geography, University of Trier. FBVI, Campus II, Behringstrasse, D-54286 Trier, Germany (herbstm/casper@uni-trier.de)

Abstract: Basically, any statement on hydrological model behaviour depends on our possibilities to differentiate between model time series. Applied within a model identification context, aggregating statistical performance measures are inadequate to capture details on time series characteristics as essentially different model results can be produced with close to identical performance measure values. It has been readily shown that the loss of information on the residuals imposes important limitations on model identification and -diagnostics and thus constitutes an element of the overall model uncertainty. In this contribution we present an approach using a Self-Organizing Map (SOM) to circumvent the identifiability problem induced by the low discriminatory power of aggregating performance measures. Instead, a Self-Organizing Map is used to differentiate the spectrum of model realizations, obtained from Monte-Carlo simulations with a distributed conceptual watershed model, based on the recognition of different patterns in time series. Further, the SOM is tentatively used as an alternative to a classical optimization algorithm to identify the model realizations among the Monte-Carlo simulations that most closely approximate the pattern of the measured discharge time series. The results are analyzed and compared with the manually calibrated model as well as with the results of the Shuffled Complex Evolution algorithm (SCE-UA).

Keywords: SOM; Self-Organizing Map; model evaluation; optimization.

1. INTRODUCTION

Model evaluation and model identification usually resort to aggregating statistical measures to compare observed and simulated time series [Legates and McCabe Jr., 1999]. In this context however these measures involve considerable problems [Yapo et al., 1998; Lane, 2007]. Aggregating measures of performance have in common that the information contained in the errors is aggregated into a single numerical value, regardless of the characteristic and the actual pattern of the error. In consequence, essentially different model results can be obtained with close to identical performance measure values although the parameter sets used to generate them are widely scattered throughout the parameter space. Because of their low discriminatory power traditional performance measures are not well suitable to give evidence of the difference or equivalence (i.e. equifinality) between alternative model realizations [Gupta et al., 2003; Beven and Binley, 1992]. This, in turn, implies serious limitations to model calibration and identification; in a sense it constitutes another source of model uncertainty [Wagener et al., 2003].
As a step toward improved extraction of information from existing data we introduce an approach that circumvents the ambiguity induced by standard objective functions: A Self-Organizing Map (SOM) [Kohonen, 2001] is used to represent the spectrum of model realizations obtained from Monte-Carlo simulations with a distributed conceptual watershed model based on the recognition of different patterns of model residual time series.

Self-Organizing maps have found successful practical applications in speech recognition, image analysis, categorization of electric brain signals [Kohonen, 2001] as well as data mining and process monitoring [Alhoniemi et al., 1999; Simula et al., 1999; Vesanto, 2000]. In the context of the hydrological sciences, however, applications of SOM are still rather uncommon.

2. METHODS

2. 1 Self-Organizing Map

A Self-Organizing Map is a type of artificial neural network (ANN) and unsupervised learning algorithm that is used for clustering, visualization and abstraction of multi-dimensional data: It maps vectorial input data items with similar patterns onto contiguous locations of a discrete low-dimensional grid of neurons, i.e. it has no output function like other types of ANN. Nearby locations on this map are attributed similar data patterns. Thus, in the course of the training, each of the map's neurons is 'tuned' to a different domain of the patterns contained in the vectorial training data items. The map units act as decoder for different types of patterns contained in the input data [Kohonen, 2001]. An input data item $\boldsymbol{x} \in X$ is considered as a vector

$$\mathbf{x} = \left[x_1, x_2, \ldots, x_n\right]^T \in \Re^n \qquad (1)$$

with n being the dimension of the input data space. In our case n is the length of the time series. A fixed number of neurons is arranged on a regular grid whose dimensions can be determined by means of heuristic algorithms if no other preferences are made. Each neuron is being associated to a weight vector

$$\mathbf{m}_i = \left[\mu_{i1}, \mu_{i2}, \ldots, \mu_{in}\right]^T \in \Re^n \qquad (2)$$

also called reference vector, which has the same dimensionality as the input vectors $\boldsymbol{x} \in X$. These weights connect each input vector $\boldsymbol{x}$ in parallel to all neurons (indexed i) of the map. Moreover the neurons are connected to each other. In our case this interconnection is defined on a hexagonal grid topology. Fig. 1 explains the functioning of a SOM: In each iteration step the Euclidean distance between a randomly chosen input data item $\boldsymbol{x}$ and the reference vectors $\boldsymbol{m}_i$ is calculated. The neuron with minimal Euclidean distance to this data item is called the best-matching unit (BMU). Subsequently the reference vectors in the neighbourhood of this BMU are updated. However, the rate of change of the reference vectors decrease proportionally to the difference between $\boldsymbol{x}$ and $\boldsymbol{m}_i$ and the number of iteration steps. Moreover, also the radius of the neighbourhood decreases proportionally to the number iteration steps (commonly a Gaussian function is used to define the neighbourhood). The mapping "self-organizes" upon repeated cycling through the input data sets. For more detailed information on SOM and its properties please refer to Kohonen [2001] or Haykin [1999]. A concise description of the algorithm is given in Herbst and Casper [2008]. In this contribution we especially make use of the fact that SOM can also be applied to project an input data vector $\boldsymbol{y}$ onto the map which has not been part of the training data manifold. This means that the node with the reference vector $\boldsymbol{m}_c$ is selected for which the Euclidean distance between $\boldsymbol{y}$ and $\boldsymbol{m}_c$ is minimal. This "image" of the projected data item then represents the domain of input data patterns from X that is most similar to $\boldsymbol{y}$. Moreover, as the number of available neurons is much smaller than the number of vectors used for the training, the selected neuron will be associated to a range of input data patterns from X which will represent the domain of input data patterns that is closest to $\boldsymbol{y}$.

2.2 Data, preparation and experimental setup

In the present example 4000 model time series constituted the input data vectors of the training data set. These were obtained from Monte Carlo simulations with the distributed conceptual watershed model NASIM [Hydrotec, 2005] running at hourly time steps over a period of two years (1 November 1994 to 28 October 1996), i.e. each input data vector

consisted of 17472 elements. The details of the model are beyond the scope of this contribution. Instead, we adopt the decision-maker's point of view and treat the model as a black-box.

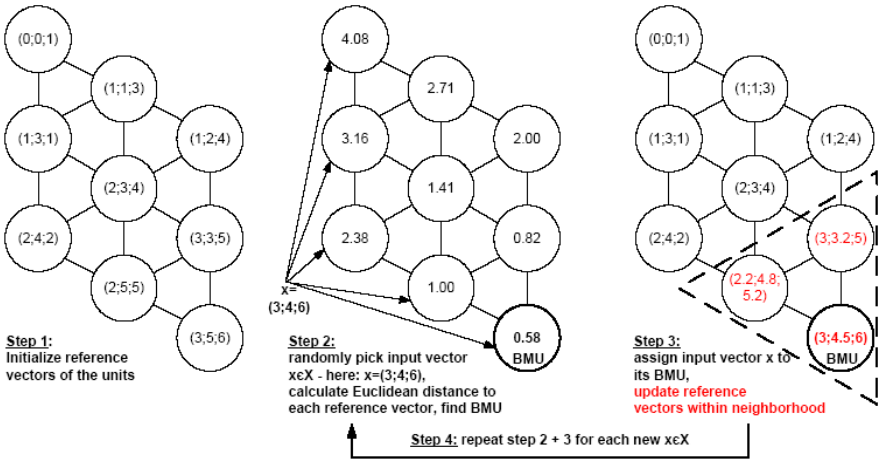


Figure 1. The basic steps of the SOM algorithm.

Seven model parameters were selected for Monte-Carlo random sampling (Tab. 2). The ranges of these parameters (Tab. 2) were chosen to be identical to those that participated in the course of a manual expert calibration for the test watershed. The bounds therefore reproduce the plausible parameter space for this catchment. The input data for the model was taken from the 129 km² low-mountain range test watershed "Schwarze Pockau" Saxony (Germany), situated near the border to Czechoslovakia and tributary of the Freiberger Mulde, a sub-basin of the Elbe River. Before the training, normalization of the data after Eq. (3) was carried out to avoid that high data values (vector elements) dominate the training because of their higher impact on the Euclidean distance measure Eq. (3) [Vesanto et al., 2000].

$$x' = \left(x - \bar{x}\right) / \sigma_x \qquad (3)$$

In order to compare the results of the aforementioned Monte-Carlo simulation and the properties of the SOM, seven measures of performance, listed in Tab. 1, were calculated for each model run. Consecutively, a reference data set, which has not been part of the training data, consisting of the time series of observed data was projected onto the SOM according to 2.1. As to the model optimization using the SCE-UA algorithm [Duan et al., 1993] the criterion for successful termination was set to a change of less than 0.05 percent of the performance criterion in three consecutive loops.

2. RESULTS

3.1 Properties of the SOM

After the training each neuron of the 22x15 SOM is expected to be activated by a narrow domain of residual patterns from the input data manifold. The neurons and their respective location on the map are identifiable by index numbers. To examine the map the means of different performance measures as well as the mean values of the model parameters on each map element are calculated. This allows to assess the properties of the map's ordering principle with respect to well known attributes such as a) the distribution of performance measures and b) the distribution of different model parameter values over the map lattice.

Table 1. Statistical goodness-of-fit measures calculated for the model output (Qobs: observed discharge, Qsim: simulated discharge).

Name	**Description**	**Formula**
BIAS	Mean error	$\frac{1}{N}\sum_{k=1}^{N}\left(Qobs - Qsim_k\right)$
RMSE	Root of mean squared error	$\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(Qobs - Qsim_k\right)^2}$
CEFFlog	Logarithmized Nash-Sutcliffe coefficient of efficiency	$\frac{\sum_{k=1}^{N}\left(\ln(Qobs) - \ln(Qsim_k)\right)^2}{\sum_{k=1}^{N}\left(\ln(Qobs) - \ln(\overline{Q}obs)\right)^2}$
IAg	Willmott's index of agreement [Willmott, 1981]; $0 \le IAg \le 1$	$1 - \frac{\sum_{k=1}^{N}\left(Qobs - Qsim_k\right)^2}{\sum_{k=1}^{N}\left(\left\|Qsim_k - \overline{Q}obs\right\| + \left\|Qobs - \overline{Q}obs\right\|\right)^2}$
MAPE	Mean average percentual error	$\frac{100}{N}\sum_{k=1}^{N}\frac{1}{Qobs}\left\|Qsim_k - Qobs\right\|$
VarMSE	Variance part of the mean squared error	$\frac{\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(Qobs - \overline{Q}obs\right)^2} - \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(Qsim - \overline{Q}sim\right)^2}}{\frac{1}{N}\sum_{k=1}^{N}\left(Qobs - Qsim\right)^2}$
Rlin	Coefficient of determination	$\frac{\sum_{k=1}^{N}\left[\left(Qsim - \overline{Q}sim\right)\left(Qobs - \overline{Q}obs\right)\right]}{\sqrt{\sum_{k=1}^{N}\left(Qsim - \overline{Q}sim\right)^2\sum_{k=1}^{N}\left(Qobs - \overline{Q}obs\right)^2}}$

Table 2. Free NASIM model parameters of the Monte-Carlo simulation with their respective parameter ranges.

Name	Description	Range
RetBasis	Storage coefficient for baseflow component [h]	0.5 – 3.5
RetInf	Storage coefficient for interflow component [h]	2.0 – 6.0
RetOf	Storage coefficient for surface runoff from unsealed surfaces [h]	2.0 – 6.0
StFFRet	Storage coefficient for surface runoff from urban areas [h]	2.0 – 6.0
hL	Horizontal hydraulic conductivity factor	2.0 – 8.0
maxInf	Maximum infiltration factor	0.025 – 1.025
vL	Vertical hydraulic conductivity factor	0.005 – 0.105

Referring to a) seven performance measures have been calculated for each model realization (Tab. 1). For each of them individual SOM lattices were colour-coded according to the mean of the performance measure of the model runs associated with each map unit. Fig. 2 shows the distribution of the performance measures from Tab. 1 on the SOM lattice. The same procedure was repeated for the values of the free parameters such that the distribution of mean parameter values can be shown for each parameter individually (Fig. 3). In each lattice of Fig. 2 and Fig. 3 the positions of the neurons remain identical such that each map element refers to identical model realizations in both figures.

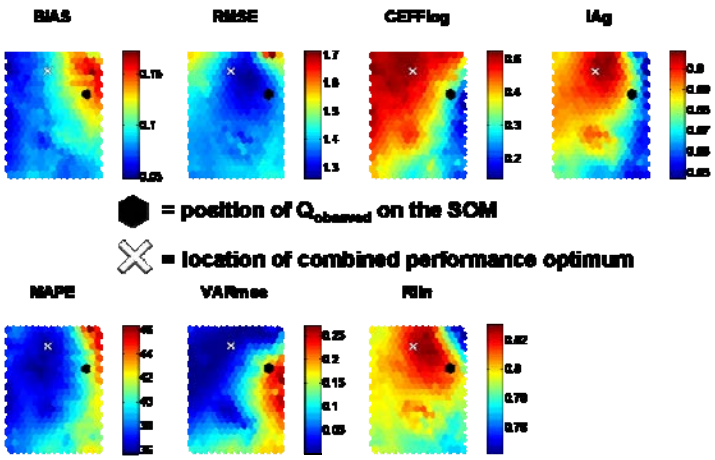


Figure 2. Distribution of the mean values of each performance measure from Table 2 over the SOM lattice, location of the best-matching unit (black dot) for measured discharge time series and common optimum (balance point) of the seven performance measures on the SOM grid (white cross).

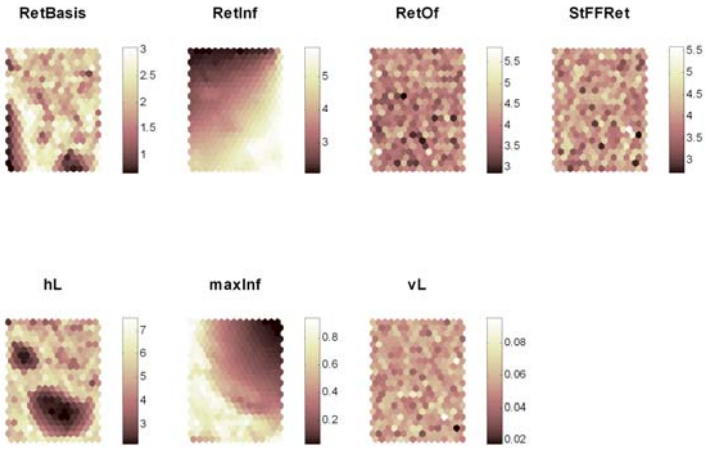


Figure 3. Distribution of the mean values of each model parameter from Table 1 over the SOM lattice.

In Fig. 2 it can be seen that, without providing information on the performance measures with the training data, the different performance values are not distributed randomly across the map but significantly relate to different regions of the lattices. As to Fig. 3, a visibly ordered relation of the map regions to different parameter values can only be stated for two parameters (RetInf and maxInf), whereas the values of RetOf, StFFRet and vL do not appear to relate to any ordering principle. A similar random pattern can be observed for the two remaining parameters (RetBasis and hL) throughout wide areas of the map. As can be seen from the locally ordered colour distribution, some intercalated areas in these lattices markedly display again a relationship between the parameter values and map locations (which stand for a certain domain of simulated time series pattern). To facilitate the interpretation of these findings we compared Fig. 3 with scatterplots of performance measures. These corroborate the assumption that only the parameters RetInf and maxInf are sensitive. The findings indicate that the parameters RetBasis and hL are subject to interaction with other free parameters, i.e. changes to a parameter influence the operation of another parameter.

3.2 Projection of a reference vector on the SOM

To locate the best-matching unit (BMU) of the measured discharge time series on the map the transformation Eq. (3) is carried out, using the normalization parameters obtained from the input data set, to ascertain that this reference data set can be projected. In Fig. 2 the location of the resulting vector is displayed on top of the performance measure distributions. Additionally, the location of the combined optimum for the seven performance measures is marked, which has been determined as the geometric center of mass of the individual performance measures optima on the map. It can be seen that the position of the BMU neither coincides with any of the expected objective function optima (which are indicated by the colour coding) nor with the common optimum location of the seven performance measures. Tab. 3 summarizes the parameter values of the 11 model realizations that are associated to the BMU for representing the model time series that are most "similar" to the measured time series. By comparing these parameters from the corresponding model runs to the ranges in Tab. 2 it becomes obvious that, with the exception of RetInf and maxInf, all parameter values span the full range of the Monte-Carlo sampling bounds. The resulting model outputs for these 11 realizations is shown in Figure 4c along with the total envelope range of all 4000 simulation outputs in the background and the observed discharge (only the period from 14 January 1995 to 21 October 1995 is reproduced here). It can thus be seen that, compared to the whole set of Monte-Carlo outputs, these realizations obviously comprise a compact subset of "similar" time series.

Table 3. Summary of the parameter values of the 11 model realizations associated to the Best-Matching map Unit when the time series vector of observed discharges is projected onto the SOM.

	RetBasis	RetInf	RetOf	StFFRet	hL	maxInf	vL
min	0.699	4.336	2.379	2.202	2.191	0.107	0.008
max	3.143	4.787	5.731	5.581	6.540	0.134	0.105
mean	1.756	4.555	4.278	3.548	4.674	0.122	0.065

Additionally, the model results obtained from an expert manual calibration and the single-objective automatic calibration using the SCE-UA algorithm [Duan et al., 1993] with the RMSE as objective function (Tab. 1) are shown in Fig. 4b. Although the SOM procedure, unlike the manual calibration, emphasizes all features of the hydrograph equally, the time series associated to the BMU of the measured discharge appear to outperform the result of the expert calibration (Fig. 4a). Not surprisingly, the SCE algorithm minimizes the RMSE and the same can be stated for most of the remaining performance measures (Tab. 4). The hydrograph corresponding to the optimized model provides a reasonable representation of the measured time series. The recession limbs, however, are more accurately reproduced by the BMU-realizations. This might be attributed to the fact that the SOM training does not tend to put emphasis on particular hydrograph features, which however can be expected when using RMSE as optimization criterion.

Table 4. Comparison of model performances for results obtained from manual calibration, optimization with SCE-UA and the SOM application. In case of the SOM mean values of 11 results are given.

	BIAS	RMSE	CEFFlog	IAg	MAPE	VARmse	Rlin
calibration	0.32	1.58	0.50	0.86	42.36	0.01	0.75
SCE-UA	0.10	1.25	0.49	0.91	36.37	0.06	0.83
SOM	0.13	1.34	0.30	0.88	40.71	0.19	0.81

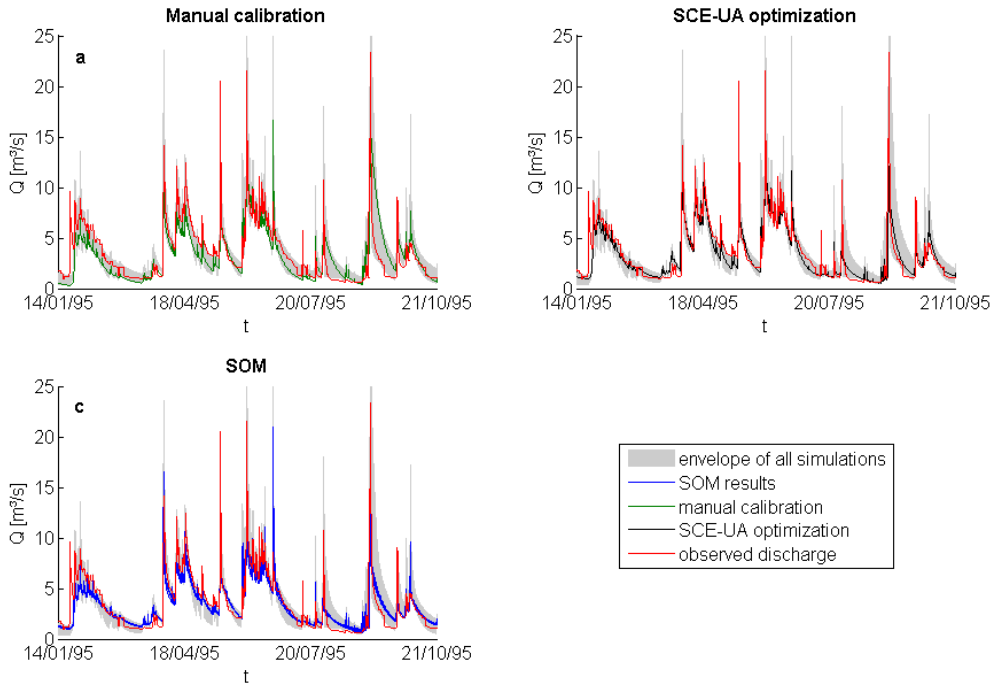


Figure 4. The model realizations as resulting from a) manual calibration, b) optimization with the SCE-UA algorithm and c) the BMU of the SOM for the measured discharge time series. The time-series are compared to the measured discharge and the envelope of the Monte-Carlo simulation.

4. DISCUSSION AND CONCLUSIONS

From the patterns of the performance measures on Fig. 2 it can be seen that certain correlation structures inherent to these statistical measures appear to be reflected by the map. Thus we deduce that the information that can be extracted by these aggregating statistical measures is assimilated and preserved by the SOM. The findings reproduced in Fig. 3, Fig. 4 and Tab. 3 demonstrate that the SOM application is capable of revealing information about parameter sensitivities and, to a certain degree, parameter interactions. We consider these results an indication of the high discriminative power of the SOM application with respect to the characteristics of different simulated discharge time series. Moreover, we were not able to obtain similar findings with traditional methods that are based on the evaluation of performance measures, e.g. parameter response surfaces for different objective functions.

The second experiment demonstrates that the information which is processed by the SOM allows differentiating the spectrum of model realizations, given with the Monte-Carlo data, such that a rather narrowly confined set of model time series which are similar to the observed time series can be identified. Of course, the resolution of the method is dependent upon the number of model time series that participated in the training. The results (Fig. 4c) were achieved with a rather small number of model data items. Nevertheless the model realizations that have been attributed to the BMU already exhibit qualities similar to the result which was based on optimization with the SCE algorithm. It is important to notice that the results were achieved without resorting to aggregating statistical measures and therefore, the “similarity” represented in the SOM is not directly quantifiable in traditional terms. Instead, it rather accounts for the complexity that is inherent to time series data and which cannot be reduced to a rank number.

The discriminatory power of the SOM that has been demonstrated in this article also highlights that uncertainty induced by the properties of the performance measure should be included in the discussion of model uncertainties and equifinality, because any statement on model behaviour depends on our possibilities to differentiate between model time series.

ACKNOWLEDGEMENTS

The authors wish to thank Oliver Buchholz (Hydrotec GmbH) and René Wengel for their invaluable support. This work was carried out using the SOM-Toolbox for Matlab by the "SOM Toolbox Team", Helsinki University of Technology (http://www.cis.hut.fi/projects/somtoolbox)

REFERENCES

Alhoniemi, E., Hollmén, J., Simula, O., and Vesanto, J.: Process Monitoring and Modeling using the Self-Organizing Map, Integrated Computer Aided Engineering, 6, 3-14, 1999.

Beven, K. J., and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Processes, 6, 279-298, 10.1002/hyp.3360060305, 1992.

Duan, Q., Gupta, V. K., and Sorooshian, S.: Shuffled complex evolution approach for effective and efficient global minimization, Journal of Optimization Theory and Applications, 76, 501-521, doi: 10.1007/BF00939380, 1993.

Gupta, H. V., Sorooshian, S., Hogue, T. S., and Boyle, D. P.: Advances in Automatic Calibration of Watershed Models, in: Calibration of Watershed Models, edited by: Duan, Q., Gupta, H. V., Sorooshian, S., Rousseau, A. N., and Turcotte, R., Water Science and Application, AGU, Washington D.C., 9-28, 2003.

Haykin, S.: Neural networks - a comprehensive foundation, 2nd ed., New Jersey, 842 pp., 1999.

Herbst, M., and Casper, M. C.: Towards model evaluation and identification using Self-Organizing Maps, Hydrology and Earth System Sciences, 12, 657-667, 2008.

Hydrotec: Rainfall-Runoff-Model NASIM - program documentation (in German), Hydrotec Ltd., Aachen, 579, 2005.

Kohonen, T.: Self-Organizing Maps, 3rd ed., Information Sciences, Berlin, Heidelberg, New York, 501 pp., 2001.

Lane, S. N.: Assessment of rainfall-runoff models based upon wavelet analysis, Hydrol. Processes, 21, 586-607, 2007.

Legates, D. R., and McCabe Jr., G. J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35, 233-241, 1998WR900018, 1999.

Simula, O., Vesanto, J., Alhoniemi, E., and Hollmén, J.: Analysis and Modeling of Complex Systems Using the Self-Organizing Map, in: Neuro-Fuzzy Techniques for Intelligent Information Systems, edited by: Kasabov, N., and Kozma, R., Physica Verlag (Springer Verlag), 3-22, 1999.

Vesanto, J.: Using the SOM and Local Models in Time-Series Prediction, Workshop on Self-Organizing Maps (WSOM'97), Espoo, Finland, 1997, 209-214, 1997.

Vesanto, J.: Using SOM in Data Mining, Licentiate's thesis, Helsinki University of Technology, 57 pp., 2000.

Vesanto, J., Himberg, J., Alhoniemi, E., and Parhankangas, J.: SOM Toolbox for Matlab 5, Helsinki University of Technology, Espoo, 60, 2000.

Wagener, T., Wheater, H. S., and Gupta, H. V.: Identification and Evaluation of Watershed Models, in: Calibration of Watershed Models, edited by: Duan, Q., Gupta, H. V., Sorooshian, S., Rousseau, A. N., and Turcotte, R., Water Science and Application, AGU, Washington D.C., 29-47, 2003.

Willmott, C. J.: On the validation of models, Phys. Geogr., 2, 184-194, 1981.

Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Multi-objective global optimization for hydrologic models, Journal of Hydrology, 204, 83-97, doi:10.1016/S0022-1694(97)00107-8, 1998.

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 Software Component for Model Output Evaluation

G. Bellocchi[a,*], M. Donatelli[a,,†], M. Acutis[b], E. Habyarimana[a]**

[a] Agriculture Research Council, 40128 Bologna, Italy;
[*] currently at European Commission Joint Research Centre, Institute for Health and Consumer Protection, Biotechnology and GMOs Unit, Ispra, Italy;
[**] currently seconded to European Commission Joint Research Centre, Institute for the Protection and Security of the Citizen, Agriculture Unit, Agri4cast Action, Ispra, Italy
[b] Department of Crop Science, University of Milan, Italy
[†] Corresponding author, marcello.donatelli@jrc.it

Abstract: As the role of biophysical models in ecological, biological and agronomic areas grows in importance, there is an associated increase in the need for suitable approaches to evaluate the adequacy of model outputs (model testing, often referred as to "validation"). Effective testing techniques are required to assess complex models under a variety of conditions, including a wide range of validation measures, possibly integrated into composite metrics. Both simple and composite metrics are being proposed by the scientific community, continuously broadening the pool of options for model evaluation. However, such new metrics are not available in commonly used statistical packages. At the same time, the large amount of data generally involved in model testing makes the operational use of new metrics a labour-consuming process, even more when composite metrics are meant to be used. An extensible and easily reusable library encapsulating such metrics would be an operational way to share the knowledge developed on model testing. The emergence of the component-oriented programming in model-based simulation has fostered debate on the reuse of models. There is a substantial consensus that component-based development is indeed an effective and affordable way of creating model applications, if components meet via their architecture a set of requirements which make them scalable, transparent, robust, easily reusable, and extensible. This paper illustrates the Windows .NET2 component IRENE (Integrated Resources for Evaluating Numerical Estimates) and a first prototype application using it, SOE (Simulation Output Evaluator), to present a concrete application matching the above requirements in the area of model testing.

Keywords: IRENE; Modelling; Model testing; SOE; Software component.

1. INTRODUCTION

The evaluation of model adequacy is an essential step of the modelling process [Jakeman et al., 2006] because it indicates the level of accuracy (how closely model-estimated values are to the actual values) of the model estimations. This is an important phase either to build confidence in a model or to allow selection of alternative models. The concept of validation is quite generally interpreted in terms of model suitability for a particular purpose, that means a model is valid and sound if it accomplishes what is expected from it [e.g. Sargent, 2001]. One of the principles of validating models dictates that complete testing is not possible [Balci, 1997], thus proving that a model is absolutely valid is an issue without solution. Exhaustive validation requires testing all possible model outputs under virtually all possible inputs (i.e., conditions). Combinations of feasible values of model inputs/outputs can generate a very large number of logical paths during execution. Due to

time and budgetary constraints, exhaustively testing the accuracy of so many logical paths is often unfeasible. Consequently, in model evaluation, the purpose is to increase confidence that the accuracy of the model meets the standards required for a particular application rather than establish that the model is absolutely correct in all circumstances. This suggests that the challenge for model evaluation is to ensure that, in addition to ensuring that minimal (application specific) standards are met, the testing should also increase the credibility of the model with the users while remaining cost effective. As a general rule, the more tests that are performed in which it cannot be proven that the model is incorrect, the more confidence in the model is increased. Yet the low priority typically given to validation in model project proposals and development plans indicates a tendency towards the minimum standards approach alone being adopted. Several validation methods are available, as reviewed in the international literature [e.g. Bellocchi, 2004; Tedeschi, 2006] but, typically, only a limited number of methods are used in modelling projects, often due to time and resource constraints. In general, limited testing may hinder modeller's ability to substantiate sufficient model accuracy.

The availability of appropriate software tools is necessary to assist model validation. The freeware, Microsoft COM-based tool IRENE_DLL (Integrated Resources for Evaluating Numerical Estimates_Dynamic Link Library, Fila et al., 2003), is a flexible tool providing extensive, integrated, statistical capabilities for use in model validation. IRENE_DLL-based spreadsheet applications were used into modelling projects to perform statistical evaluation of model outputs [Bechini et al., 2006; Incrocci et al., 2006; Diodato and Bellocchi, 2007a, b, c; Amaducci et al., 2008]. The DLL was also used to tailor a dedicated application for evaluation of pedotransfer functions [Fila et al., 2006], and coupled with the model for rice production WARM [Bellocchi et al., 2006]. Since IRENE_DLL was developed, the component-oriented paradigm has evolved, specifying new requirements in order to increase software quality, reusability, extensibility, and transparency for components providing solutions in the biophysical domain [Donatelli and Rizzoli, 2008]. Also, the COM paradigm has become *de facto* obsolete since the advent of the .NET platform of Windows which provides, among other features, much simpler provisions for installation and versioning.

In this paper, the role of computer-aided validation support is considered for its effectiveness to support modelling projects which are applicable across a broad range of fields. In particular, the features of the Windows .NET component IRENE (Integrated Resources for Evaluating Numerical Estimates) and one prototype application SOE (Simulation Output Evaluator) using it are presented, showing the advantages achievable via a component-oriented architecture in building a software component in the domain of model validation statistics.

2. STATISTICS FOR MODEL VALIDATION

The component implements a range of statistical measures and visual techniques that can be used to assess goodness-of-fit of a given model (when its outputs are assessed against actual data) and to compare the performance of a suite of models (Table 1). There are several papers that contain the equations of the statistics and methods used for model validation, hence the equations are not reported here (see online documentation).

Table 1. Model validation statistics implemented in IRENE.

Group of statistics	Type of statistics
Difference-based	Simple, absolute, squared indices; test statistics
Association-based	Regression parameters (from alternative fitting methods); test
Pattern	Pattern indices (range-based, F-based)
Aggregate	Multi-level composite indicators
Time mismatch	Time mismatch indices

2. 1 Difference-based Statistics

Difference-based (or residual-based) statistics quantify the departure of the model outputs from the measurements, consisting of simple, absolute and squared statistics [Fox, 1981]. Mean bias, the mean difference between observed and model-estimated values, is likely to be the oldest statistic to assess model accuracy. More common is the mean square error, or equivalently its square root, the root mean square error (or derived statistics such as the relative root mean square error). Mean absolute error measures the mean absolute difference between observed and estimated values, and is also used as the mean absolute percent error. Willmott [1981] index of agreement and the modelling efficiency statistic [Nash and Sutcliffe, 1970], interpreted as the proportion of variation explained by the model, are to be used for a better interpretation of squared errors.

2.2 Association-based Statistics

Measures of statistical association such as correlation and regression coefficients [Addiscott and Whitmore, 1987] provide quantitative estimates of the statistical covariation between estimated and measured values. Linear regression parameters (intercept and slope) are calculated according to two methods: ordinary least squares and reduced major axis.

2.3 Pattern Statistics

The presence of patterns in the residuals versus independent variables (e.g., a model input or a variable not considered in the model) is quantified by computing pattern indices of two types: range-based and F-based [Donatelli et al., 2004a].

2.4 Probability Distributions

To identify model adequacy for stochastic models, both probability density and cumulative distribution functions from observed and estimated values are compared [Reynolds and Deaton, 1982].

2.5 Aggregation of Statistics

Multiple validation metrics can be combined according to the methodology originally developed by Bellocchi et al. [2002], based on an expert weighting expression (fuzzy-based rules) of the balance of importance of the individual metrics and their aggregation into modules (first-stage aggregation). The metrics which can be aggregated are the ones either made available within the component, and/or resulting from an extension of it. The modules (or modules and simple metrics) may be further aggregated into an overall indicator (second-stage aggregation). Several aggregation levels can be composed hierarchically. Figure 1 shows the tree-view of a two-level aggregation index.

2.6 Time Mismatch

Either simple or composite statistics can be used to assess the mismatch in time series comparison [Donatelli et al., 2002].

3. SOFTWARE COMPONENT DESIGN

The software component IRENE (Integrated Resources for Evaluating Numerical Estimates) contains functions for the computation of simple and composite statistics for evaluation of model estimates. Each simple statistic is implemented as a class. Composite statistics are the result of an association of classes, obtained dynamically via an XML file. Each statistic computed via internal methods is fully described as description and units, as well as maximum, minimum and default values. The software architecture of this component further develops the one used in previously developed components [Carlini et

al., 2006; Donatelli et al., 2006a, b], and is fully described by Donatelli and Rizzoli [2008]. Transparency and ease of maintenance are granted, also providing functionalities such as the test of input data versus their definition prior to computing any simple or composite validation metric; same tests can be run on outputs. The component can be extended independently by third parties without requiring re-compilation. The component is freely available to scientists and institutions developing component-oriented models and applications in the agro-ecological field. The component is written in C# for .NET2, but extensions can be written in any .NET language. The component is deployable and reusable in any application developed using the Microsoft .NET framework. The IRENE software development kit includes sample projects which show how to use and extend the component. Code documentation is also provided and the online help file is available at: http://www.apesimulator.it/help/utilities/irene.

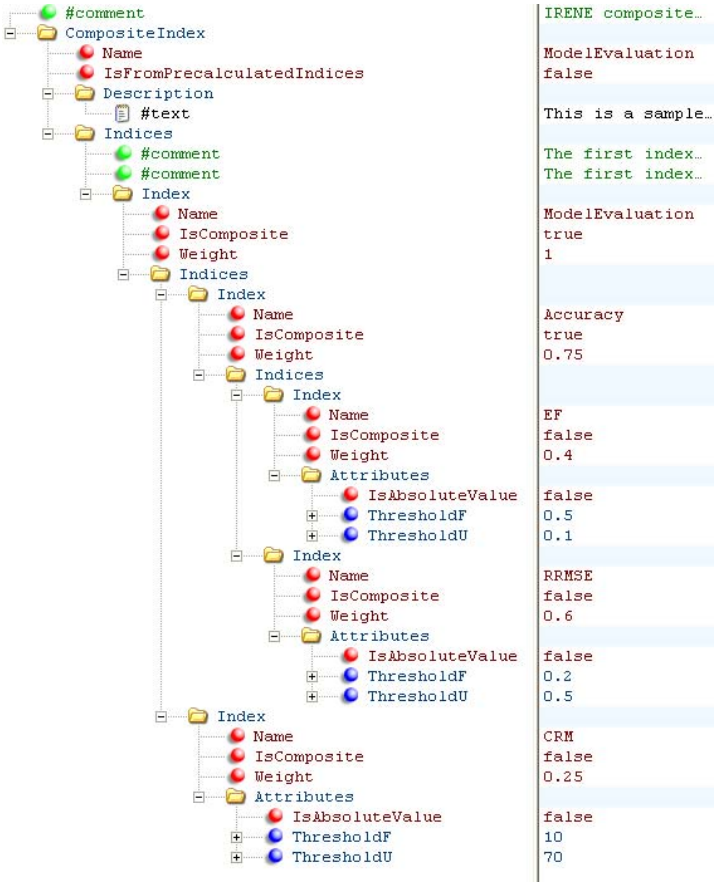


Figure 1. Tree view of a two-level composite indicator (see 3.1).

3.1 Composite Indicators

Composite indicators can be composed in hierarchical structure (including any hierarchical level), using statistics computed internally to IRENE, made available from extensions of IRENE, or provided as inputs. In the latter case, the component runs only as a fuzzy-based aggregator for construction of aggregated indicators. The structure of composite indicators can be saved as XML file for reuse. As simple example (XML code in Figure 2), RRMSE (relative root mean square error) and EF (modelling efficiency) are aggregated into an indicator (named “Accuracy”), and the latter is further combined to CRM (coefficient of residual mass) to give a second-level indicator (namely, ”ModelEvaluation”).

```
<!--IRENE composite indices CRA.IRENE.dll -->
<CompositeIndex Name="ModelEvaluation" IsFromPrecalculatedIndices="false">
   <Description>This is a sample two-levels composite
     index</Description>
  <Indices>
    <!--The first index node has always weight = 1 -->
    <!--The first index node always IsComposite = true -->
    <Index Name="ModelEvaluation" IsComposite="true" Weight="1">
      <Indices>
        <Index Name="Accuracy" IsComposite="true" Weight="0.75">
          <Indices>
            <Index Name="EF" IsComposite="false" Weight="0.4">
              <Attributes IsAbsoluteValue="false">
                 <ThresholdF>0.5</ThresholdF>
                 <ThresholdU>0.1</ThresholdU>
               </Attributes>
             </Index>
            <Index Name="RRMSE" IsComposite="false"
                Weight="0.6">
              <Attributes IsAbsoluteValue="false">
                 <ThresholdF>0.2</ThresholdF>
                 <ThresholdU>0.5</ThresholdU>
               </Attributes>
             </Index>
           </Indices>
         </Index>
        <Index Name="CRM" IsComposite="false" Weight="0.25">
          <Attributes IsAbsoluteValue="false">
             <ThresholdF>10</ThresholdF>
             <ThresholdU>70</ThresholdU>
           </Attributes>
         </Index>
       </Indices>
     </Index>
   </Indices>
 </CompositeIndex>
```

Figure 2. XML code of a two-level composite index (see tree view in Figure 1).

4. SIMULATION OUTPUT EVALUATOR

Simulation Output Evaluator (SOE) is a data analysis tool which makes use of the component IRENE, designed to provide easy access to model evaluation techniques and graphical views. It is illustrated here with the purpose of providing one proof-of-concept of a client-side application of IRENE. For each analysis, two datasets (which can be observed and model-estimated data points) must be loaded (in XML-format with schema, MS Excel worksheets, or any more compact binary format, the latter via a reusable component also made available, CRA.Core.IO.dll). When the second dataset is loaded, combo boxes of tables and variables are populated with the data from the supporting format that the user can select for analysis. Another variable, to be used as covariate in the pattern analysis, can also be loaded. A number of analyses can be performed and displayed in various screen tabs, providing graphical data views as well as the indices computed via IRENE. One of the tabs contains a summary of the analyses done which can be saved as XML file; the graphics can also be saved as .JPG-format files. A tree-structure graphic reflecting the hierarchical composition of statistics into aggregated indicators and their impact is also available. The current version of SOE is a first prototype, currently being refactored. Its help file is available online at http://www.apesimulator.it/help/tools/soe.

5. DISCUSSION

Software tools specifically created for model validation provide effective computer-aided support. In particular, they can significantly reduce the testing time and effort. A key step in this direction is the coupling between model components and validation techniques, the latter also implemented into component-based software [e.g. Bellocchi and Confalonieri, 2006]. To meet the substantial model quality challenges, it is necessary to improve the tools and technologies currently available, and their cost-benefit characterizations. The

emergence of new technologies in simulation modelling has, in fact, fostered debate on the reuse of models. A large number of existing agricultural and ecological models have been implemented as software that cannot be well maintained or reused, except by their authors, and therefore cannot be easily transported to applications developed by third parties. Therefore the focus needs to be on the use of design patterns that encourage usability, reusability and cross-language compatibility, thus facilitating model development, integration, documentation and maintenance [Donatelli et al., 2004b]. The component-oriented paradigm is the leading methodology in developing systems in a variety of domains, including agro-ecological modelling [e.g. Argent, 2005]. Component-oriented development has emerged steadily as a paradigm that focuses on explicit and semantically rich interfaces. The substantial consensus that emerged within the scientific community around component-based development as effective and affordable way of creating model applications relies upon the possibility offered by the component architecture to meet a set of requirements to make components scalable, transparent, robust, easily reusable, and extensible. The distribution of validated model components can substantially decrease model validation effort when reused, providing functionalities to improve the process of model verification. The component IRENE makes available a number of metrics in a discrete software unit, hence allowing operational access to the computation of a variety of model evaluation metrics. Its extensibility, also available independently for third parties, allows easy maintenance and further development.

6. CONCLUSIONS

The goals of IRENE development are to extend access to model validation statistics to multiple users, and provide architecture to ensure reuse and extension of coded statistics. IRENE attempts to overcome some of the technical challenges that to date have limited the development of reusable validation capabilities. Via its documentation and the metadata associated with each variable, IRENE is also a way to share data processing knowledge in an extensible way.

ACKNOWLEDGEMENTS

This publication has been funded by the SEAMLESS integrated project, EU 6th Framework Programme for Research, Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract no. 010036-2).
The authors gratefully acknowledge Ing. A. Sorce for the design and implementation of the composite indices modules within the IRENE component.

REFERENCES

Addiscott, T.M., and A.P. Whitmore, Computer simulation of changes in soil mineral nitrogen and crop nitrogen during autumn, winter and spring, *Journal of Agricultural Science (Cambridge)*, 109, 141-157, 1987.

Amaducci, S., M. Colauzzi, G. Bellocchi, and G. Venturi, Modelling post-emergent hemp phenology (*Cannabis sativa* L.): Theory and evaluation, *European Journal of Agronomy*, 28, 90-102, 2008.

Argent, R.M., A case study of environmental modelling and simulation using transplantable components. *Environmental Modelling & Software*, 20, 1514-1523, 2005.

Balci, O., Principles of simulation model validation, verification, and testing, *Transactions of S.C.S.I.*, 14, 3-12, 1997.

Bechini, L., S. Bocchi, T. Maggiore, and R. Confalonieri, Parameterization of a crop growth and development simulation model at sub-model components level. An example for winter wheat (*Triticum aesivum* L.), *Environmental Modelling & Software*, 21, 1042-1054, 2006.

Bellocchi, G., Appendix A: Numerical indices and test statistics for model evaluation. In: Pachepsky, Y., and W. Rawls (Eds.), *Development of pedotransfer functions in soil hydrology*, Elsevier, Amsterdam, 394-400, 2004.

Bellocchi, G., M. Acutis, G. Fila, and M. Donatelli, An indicator of solar radiation model performance based on a fuzzy expert system, *Agronomy Journal*, 94, 1222-1233, 2002.

Bellocchi, G., R. Confalonieri, and M. Donatelli, Crop modelling and validation: integration of IRENE_DLL in the WARM environment, *Italian Journal of Agrometeorology*, 11, 35-39, 2006.

Carlini, L., G. Bellocchi, and M. Donatelli, Rain, a software component to generate synthetic precipitation data, *Agronomy Journal*, 98, 1312-1317, 2006.

Diodato, N., and G. Bellocchi, Estimating monthly (R)USLE climate input in a Mediterranean region using limited data, *Journal of Hydrology*, 345, 224-236, 2007a.

Diodato, N., and G. Bellocchi, Modelling reference evapotranspiration over complex terrains from minimum climatological data, *Water Resources Research*, 43, doi:10.1029/2006WR005405, 2007b.

Diodato, N., and G. Bellocchi, Modelling solar radiation over complex terrains using monthly climatological data, *Agricultural and Forest Meteorology*, 144, 111-126, 2007c.

Donatelli, M., M. Acutis, G. Bellocchi, and G. Fila, New indices to quantify patterns of residuals produced by model estimates, *Agronomy Journal*, 96, 631-645, 2004a.

Donatelli, M., M. Acutis, G. Fila, and G. Bellocchi, A method to quantify time mismatch of model estimates. In: Villalobos, F.J., and L. Testi (Eds.), *Proceedings of the 7th European Society for Agronomy Congress*, July 15-18, Cordoba, Spain, 269-270, 2002.

Donatelli, M., G. Bellocchi, and L. Carlini, Sharing knowledge via software components: models on reference evapotranspiration, *European Journal of Agronomy*, 24, 186-192, 2006a.

Donatelli, M., L. Carlini, and G. Bellocchi, A software component for estimating solar radiation, Environmental Modelling & Software, 21, 411-416, 2006b.

Donatelli, M., A. Omicini, G. Fila, and C. Monti, Targeting reusability and replaceability of simulation models for agricultural systems. In: Jacobsen, S.E., C.R. Jensen, and J.R. Porter (Eds.), *Proceedings of the 8th European Society for Agronomy Congress*, July 11-15, Copenhagen, Denmark, 237-238, 2004b.

Donatelli, M., and A. Rizzoli, A design for framework-independent model components of biophysical systems. *These proceedings*, 2008.

Fila, G., G. Bellocchi, M. Donatelli, and M. Acutis, IRENE_DLL: A class library for evaluating numerical estimates, *Agronomy Journal*, 95, 1330-1333, 2003.

Fila, G., M. Donatelli, and G. Bellocchi, PTFIndicator: an IRENE_DLL-based application to evaluate estimates from pedotransfer functions by integrated indices, *Environmental Modelling & Software*, 21, 107-110, 2006.

Fox, D.G., Judging air quality model performance: a summary of the AMS workshop on dispersion models performance, *Bulletin of the American Meteorological Society*, 62, 599-609, 1981.

Incrocci, L., G. Fila, G. Bellocchi, A. Pardossi, C.A. Campiotti, and R. Balducchi, Soil-less indoor-grown lettuce (*Lactuca sativa* L.): approaching the modelling task, *Environmental Modelling & Software*, 21, 121-126, 2006.

Jakeman, A.J., R.A. Letcher, and J.P. Norton, Ten iterative steps in development and evaluation of environmental model, *Environmental Modelling & Software*, 21, 602-614, 2006.

Nash J.E., and J.V. Sutcliffe, River flow forecasting through conceptual models, Part I - A discussion of principles, *Journal of Hydrology*, 10, 282-290, 1970.

Reynolds, J.M.R., and M.L. Deaton, Comparisons of some tests for validation of stochastic simulation models, *Communications of Statistical Simulation and Computation*, 11, 769-799, 1982.

Sargent, R.G., Verification, validation and accreditation of simulation models. In: Peters, B.A., J.S. Smith, D.J. Medeiros, and W.W. Rohrer (Eds.), *Proceedings of the 2001 Winter Simulation Conference*, Arlington, VA, USA, December 10-13, 106-114, 2001.

Tedeschi, L.O., Assessment of the adequacy of mathematical models, *Agricultural Systems*, 89, 225-247, 2006.

Willmott, C.J., On the validation of models, *Physical Geography*, 2, 184-194, 1981.

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 SOAP to Develop Software for Air Trajectory Calculation

J. Zeng[a], T. Matsunaga[a] , H. Mukai

[a] National Institute for Environmental Studies, Onogawa 16-2, Tsukuba, Ibaraki, 305-8506 Japan
Corresponding author: zeng@nies.go.jp

Abstract: We developed a program with Simple Object Access Protocol and constructed a data server system to deliver meteorological data from the server to clients of the program for air trajectory calculation. Our practice shows that with proper design, the performance of running the program with client-server configuration is well acceptable comparing with a program of traditional design. The request-for-data-on-demand approach helps to reduce the burden of maintaining huge amount of data in local machine by end-user and provides a more transparent way to calculate air trajectories.

Keywords*: SOAP; Air Trajectory; Meteorology; Network Computing.*

1. INTRODUCTION

Software for air trajectory calculation is a popular tool for studying airflow pattern [Miller, 1981; Harris, 1992] and source-receptor relation [e.g., Ashbaugh, 1983; Stohl, 1996; Tsuang et al., 2002] in environmental research, especially in pollution or greenhouse gas monitoring [e.g., Moody and Galloway, 1988; Mukai and Suzuki, 1996; Lam et al., 2001]. However, users of such an application often face the difficulties of maintaining huge amount of meteorological data and feeding data to the model of the program. The first problem seems disappearing with the unprecedented increase of storage capacity of personal computers. But new computer power often leads to refining of meteorological models and thus the output volume from models. As a result, there seems never enough storage space for data. Even when storage space is not a problem, downloading GB to TB of data can be time consuming. Meteorological data are subject to changes when new methods for improving data quality become available; therefore updating data is necessary from time to time, which adds additional burden to data maintenance. The second problem comes from the fact that software for air trajectory calculation is usually designed to use data in a specific format. For example, FLEXTRA [Stohl, 1999] is designed for data in GRIB (http://dss.ucar.edu/docs/formats/grib/gribdoc/) format on sigma levels and HYSPLIT [Draxler and Hess, 2004] only accepts data on hybrid levels. Modifying such software for a differently formatted dataset or converting one format to another requires advanced programming skills that often are out of the expertise of researchers. Therefore, a better way to design such software is to make the input transparent by separating the input component from the modelling component.

This article introduces the development of METeorological data EXplorer or METEX (http://db.cger.nies.go.jp/metex/) and the techniques used to solve the problems of data storage and format. By connecting to a data server, the program can be used to calculate air trajectories of several decades in batch mode without storing meteorological data locally. Users do not have to know the format. The server is responsible for format conversion. For advanced users, they can use a scripting language to initialise calculation, to control program execution, and to make new input interfaces for data in different formats.

2. SOFTWARE AND SYSTEM DESIGN

2.1 SOAP Interface

To help reducing the burden of maintaining huge amount of meteorological data in local machine by end-user, we have chosen Simple Object Access Protocol or SOAP (http://www.w3.org/TR/soap/) for sending necessary data from server to client for trajectory calculation. SOAP is a protocol for exchanging XML-based messages over computer networks, normally using HTTP/HTTPS. Among many implementations of SOAP, we found that gSoap (http://www.cs.fsu.edu/~engelen/soap.html) suits our purpose best. The tool generates C/C++ code for SOAP/XML web services from C/C++ header file. The generated code also includes interfaces for serializing both primitive and structured types of C/C++ variable.

Generality often leads to complexity in software design. To balance the two opposite factors, we decided to implement SOAP interfaces for five types of C/C++ variable:

- 32-bits integer for data of absolute precision;
- 64-bits float for number with decimal points;
- Character string for descriptive information;
- Binary array for data array;
- Base64 encoded binary array.

The last two are generic arrays that can be used to transfer binary array of any type. The server and client have to predefine rules for what an array contains. The first array type was used in the early development of METEX and is kept for backward compatibility; and the second was introduced later to reduce the package size of serialized data array. The package size of serialized data produced by gSoap generated C/C++ code for the first array type is several times larger than the source data array. Using base64 encoding, serialized data size is not much larger than 4/3 of the original size.

The C/C++ header for generating SOAP interface for METEX by gSoap is listed in the appendix. Implementation details of the interface can be found in the C++ source code http://www.zegraph.com/z-script/src/metex2.cpp. Here we only outline two key points.

The client component makes request by calling the following C/C++ function generated by gSoap:

```
int soap_call_metex__Request(struct soap *soap, const char
*soap_endpoint, const char *soap_action, char *service, struct
metex__params *params, struct metex__output &r)
```

Which returns response status for the function parameters of

1. *soap* – gSoap object;
2. *soap_endpoint* – server URL;
3. *soap_action* – unused by this implementation;
4. *service* – used by this application to identify request type;
5. *params* – variable containing request parameters whose types are defined in the appendex;
6. *r* – results from server whose types are defined in the appendex.

The server component must implement this C/C++ function:

```
int metex__Request(struct soap *soap, char *service, struct
metex__params* params, struct metex__output &r)
```

Which will be called by the function that waits for client requests. The implementation fills the output variable *r* with data according the contents of *service* and *params* and return proper response status.

2.2 Input and Output Interfaces

A typical air trajectory calculation requires several meteorological variables in two- and three-dimensional, e.g., surface pressure, wind velocities, temperature, and etc. The data size could be in the order of tens of megabytes to hundreds of megabytes depending on the spatial and time resolutions of a dataset and the trajectory length. Hundreds of gigabytes to terabytes of storage space are needed for a research that uses historical meteorological data of several decades. Even if storage space is not a problem, downloading data can be time consuming. Meteorological data are subject to changes when new methods for improving data quality become available; therefore it is necessary to update data from time to time, which adds additional burden to data maintenance.

We identified that in practice most researches involving air trajectory are interested in regional problems and that in such cases only a small portion of meteorological data is needed for trajectory calculation, i.e., data at those grids along a trajectory. Figure 1 shows a 5-day back trajectory calculated for the greenhouse gas monitoring station at Hateruma Island, Japan. Because the uncertainty induced by meteorological data accumulates with the length of trajectory, which is equivalent to the length of weather prediction, much longer trajectories are considered not useful for the research of source-receptor relation. In addition, it does not take much longer to compute a trajectory from one dataset time to another than to read input data with moderate horizontal resolution of 1x1 degree and vertical levels of about 20.

Based on this analysis, we designed the input interface to get data grid-by-grid, instead of data of the whole 2-dimensional or 3-dimentional field. We further separated the IO interfaces of METEX from its trajectory models to maximize the flexibility. The program requests for data through a user object, which in turn tries to read data from a local machine or fetch data from a remote data server. For the output interface, the program either saves results in every hour or sends them to the user object for processing.

Because the modelling component does not have to know the format of input data, the program can be used with any type of data without re-compiling source code as long as an input interface is implemented. So far we have prepared input interface components for data from European Centre for Medium-Range Weather Forecasts (http://www.ecmwf.int/), National Centers for Environmental Prediction (http://www.ncep.noaa.gov/), and Japan Meteorological Agency (http://www.jma.go.jp/jma/indexe.html).

At initialisation, METEX requests for grid sizes and types for each dimension of a given dataset. During trajectory calculation, it requests for data at grids surrounding the current position of air parcel. This approach significantly reduces data traffic through the network.

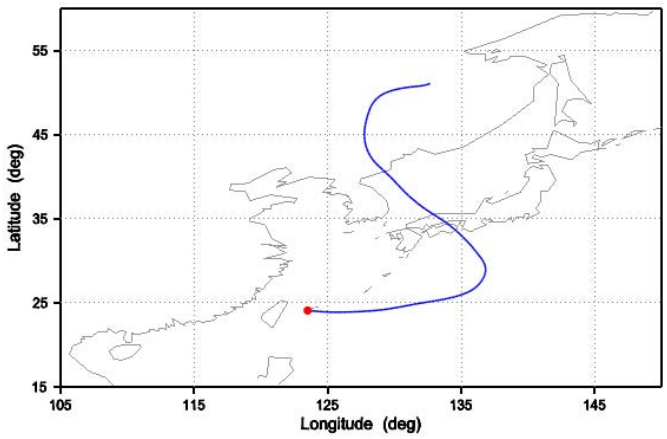


Figure 1. 5-day back trajectory initialised at Hateruma station (123º3'E, 24º49'N) with altitude of 500 m. The initial time is 0:00 UTC, January 1, 2007.

2.3 METEX as Software Library

In contrast to traditional design of environmental software that read initial and control parameters from initial files, METEX is designed as a dynamic link library to be used with a scripting language. The combination makes it even more flexible to handle input and output. The scripting language ZeScript (https://sourceforge.net/projects/zescript/) is used to control every aspect of METEX execution. The C/C++ code generated from gSoap is also bound with ZeScript to run as a service to supply data to METEX. Details for using the library can be found at http://db.cger.nies.go.jp/metex/interfaces.html.

2.4 Data Server

A data server has been set up at the Centre for Global Environmental Research, National Institute for Environmental Studies, Japan. When the URL http://db.cger.nies.go.jp/metex/soap/ is set to METEX, as shown in Figure 2, the program will get necessary data from the server for trajectory calculation. The server has data from January 1, 1950 to at least the previous month of the current date and updates its database from NCEP FTP server (ftp://ftp.cdc.noaa.gov/pub/Datasets/ncep.reanalysis/) at lest once per month.

Users can also download data from the server to their machine to do calculation when the number of calculation is large and the computing speed is critical to their applications.

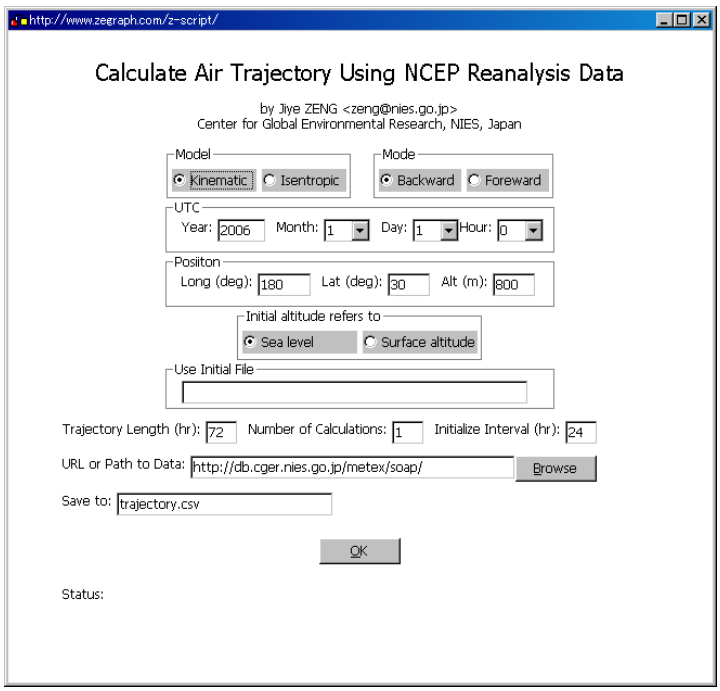


Figure 2. A user interface of METEX. The URL of a data may be set for fetching input data through the Internet. Alternatively, a path to data in local machine can also be set for reading input data locally.

2.5 Model Validation

In addition to visual and statistical comparisons of METEX results with those of other programs, we analysed the correlation between observed concentration change of methane in given time intervals and the end-point distances of backward trajectories that were initialised at observation times. The basic assumption is that the larger the distance, the higher the possibility that sampled air masses came from different regions with different source strength for methane. The results [see Zeng et al., 2003] show that although

trajectory quality varies with initial conditions and model assumptions, they reveal good source-receptor relations in general.

2.6 Test of Server Stability and Computing Speed

We have tested the reliability and stability of the server-client configuration with continuous executions of METEX for input data of January 1950 to December 2006. After solved some troubles at the initial stage, the server is now serving users all over the world. It sometimes accepted over 3 millions requests for data per day. No trouble has occurred in the past three months.

In a series of tests, we initialised the model every 6 hours, which is the time resolution of meteorological data, and calculated 3-day back trajectories over 56 years. The PC used to run tests has 3 GHz CPU, 1 GB memory, and Windows XP; and was located in the same LAN segment of the data server. It took about six days to calculate more than 80,000 trajectories when only one METEX was using the server. The speed is equivalent to 10 trajectory lines per minute. In comparison, the first version of METEX, which was designed in the traditional way to read input data of all grids at once, can calculate about 150 lines of trajectories of the same length with the same PC in one minute. Several similar tests for the server-client configuration with a home PC, located in the same city, shown twice slower speeds at the best condition.

3. CONCLUSION

Network computing used to be the territory of large and complex modelling and simulation. The ever-increasing accessibility and speed of Internet open a new way to develop environmental software for individual user that involves huge amount of data. Taking the advantage of Internet, we used SOAP to add an additional input interface to METEX so that users have the option of either reading data from local machine or request for data from a remote server for air trajectory calculation.

Although the computing speed of METEX running with the client-server configuration is a order of magnitude slower than running with data input from local machine, the performance is well acceptable because in most researches, such a software is used to calculate trajectory for a short period of time. For example, a laboratory in our institute uses it about once per month to obtain trajectories for air sample positions along ship routes; and a municipal government in Japan uses it one per month to archive trajectories for their monitoring stations. The request-for-data-on-demand approach helps to reduce their burden of maintaining meteorological data in local machine because downloading large amount of data can be time consuming and archived data needs updating when the data provider makes changes to improve data quality. With the client-server configuration, the separation of input component from modelling component makes the program transparent and easy to use for end-users.

ACKNOWLEDGMENTS

Dr. Robert A. van Engelen, the developer of gSoap, generously offered his help to use the code generator. This has greatly reduced the time to implement SOAP in METEX.

Appendix. C/C++ Variable Types for Serialization by gSoap

```
enum type_enum { METEX_ANY, METEX_INTEGER, METEX_REAL, METEX_STRING,
METEX_BINARY, METEX_BASE64 };

class metex__any
{ public:
  struct soap *soap;
  metex__any();
  virtual ~metex__any();
  virtual int type() const;
};
```

```
class metex__integer : public metex__any
{ public:
  int value;
  metex__integer();
  virtual ~metex__integer();
  virtual int type() const;
};

class metex__real : public metex__any
{ public:
  double value;
  metex__real();
  virtual ~metex__real();
  virtual int type() const;
};

class metex__string : public metex__any
{ public:
  char *__ptr;
  metex__string();
  virtual ~metex__string();
  virtual int type() const;
};

class metex__binary : public metex__any
{ public:
  int __size;
  unsigned char *__ptr;
  char *msg;
  metex__binary();
  virtual ~metex__binary();
  virtual int type() const;
};

class metex__base64 : public metex__any
{ public:
  char *__ptr;
  char *msg;
  metex__base64();
  virtual ~metex__base64();
  virtual int type() const;
};

struct metex__params{ int __size; metex__any **__ptr; };

struct metex__output{ metex__any *__ptr; };

int metex__Request(char *service, struct metex__params *params, struct
metex__output &r);
```

REFERENCES

Ashbaugh, L.L., A Statistical Trajectory Technique for Determining Air Pollution Source Regions, Journal of the Air Pollution Control Association, 33, 1096-1098, 1983.

Draxler, R.R. and Hess, G.D., Description of the Hysplit_4 Modeling System, NOAA Technical Memorandum ERL ARL-224, 25 pp, 2004.

Harris, J.M., Analysis of Day Midtropospheric Flow Patterns for the South Pole: 1985-1989, Tellus, 44, 409-412, 1992.

Lam, K.S., Wang, T.J., Chan, L.Y., Wang, T., and Harris, J., Flow Patterns Influencing the Seasonal Behaviour of Surface Ozone and Carbon Monoxide at a Coastal Site Near Hong Kong, Atmospheric Environment, 35, 3121-3135, 2001.

Miller, J.M., A Five-Year Climatology of Back Trajectories From the Mauna Loa Observatory, Hawaii, Atmospheric Environment, 15, 1553-1558, 1981.

Moody, J.L. and Galloway, J.N., Quantifying the Relationships Between Atmospheric Transport and the Chemical Composition on Bermuda, Tellus, 40, 463-497, 1988.

Mukai, H. and Suzuki, M., Using Air Trajectory to Analyse the Seasonal Variation of Aerosol Transport to the Oki Island, Atmospheric Environment, 30, 3971-3934, 1996.

Stohl, A., Trajectory Statistics – a New Method to Establish Source-Receptor Relationships of Air Pollutants and Its Application to the Transport of Particulate Sulphate in Europe. Atmospheric Environment, 30, 579-581, 1996.

Stohl, A., The FLEXTRA Trajectory Model Version 3.0, User's Guide, 41 pp, 1999.

Tsuang, B.J., Chen, C.L., Pan, R.C., and Liu, J.H., Quantification on Source/Receptor Relationship of Primary Pollutants and Secondary Aerosols From Ground Sources, Part I. Theory, Atmospheric Environment, 36, 411-419, 2002.

Zeng, J., Tohjima, Y., Fujinuma, Y., and Mukai, H., and Katsumoto, M., A Study of Trajectory Quality Using Methane Measurements from Hatetuma Island, Atmospheric Environment, 37, 1911-1919, 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

Web Components for Development of Computational Methods: Example with Fuzzy Logic Rules

G. Bellocchi†, F. Foscarini, M. Canonico, G. Van den Eede

European Commission Joint Research Centre, Institute for Health and Consumer Protection, Biotechnology and GMOs Unit, Ispra, Italy;
† *Corresponding author, gianni.bellocchi@jrc.it*

Abstract: Computational methods (either statistical or non-statistical) for experimental design, data analysis and optimisation are needed to provide information which can be used by a variety of decision makers and for a variety of purposes in the natural sciences research. It is crucial for scientists to keep pace with evolving computational approaches, to comply with specific needs emerging from applied research. Software tools are desirable not only to provide easy access to a large variety of data processing methods but also to grant for continually updating, adding more solutions as well as keeping all the parts up to date as computing software environments change and novel computational methods evolve. Based on that, the Data Analysis 'N' Assessment (DANA) project was approached and meant as platform-independent, integrated repository of statistical and soft-computing procedures for use in applied research. It consists of a core engine and a set of plug-ins sharing an array of functionalities. The core engine is written in Java and integrates a set of libraries (e.g., from the R language for statistical computing) accessible via web services. This is achieved through usage of the object-oriented paradigm and high-level utility functions, hiding much of the complexity of the underlying libraries. The core engine handles data logging, event handling, information and synchronisation between different libraries. All classes implement a defined interface that makes it possible to create derivative works by third parties. A Graphical Users' Interface is dynamically generated via PHP through a set of configuration XML files. The XML Schema Definition below undergoes validation. This paper illustrates the features of DANA through the components used for computation of fuzzy-logic based rules.

Keywords: DANA; Fuzzy logic; Java; Software component; XML.

1. INTRODUCTION

An ever-increasing number of researchers and software developers are attracted towards data processing techniques. Different reasons justify such continuing interest. On one hand, high level quality in scientific work is perceived when a minimum of sophistication in data processing is placed in. A tight correlation is also occurring between the problems that emerge in developing innovative analytical approaches and those typical of other research fields such as, for example, mathematical modelling. Moreover, today's scientific research easily results in the generation of considerable amounts of data. To ensure data quality, accurate analyses, and evaluation tend to be integral parts of the processes conducted by research projects. Creation of a platform to provide for effective management of the data and the associated analyses has therefore become a high priority. Such a platform should ultimately support data handling, processing, visualization, reporting, and knowledge dissemination tools, integrated with robust document tracking and control tools.
Basic and advanced statistical methods are routinely used to support a variety of studies. Complementary approaches are the so-called soft-computing techniques (including neural networks, fuzzy systems, evolutionary computations, etc.) which attempt to study, model

and analyse systems that remain intractable to conventional analytical methods. There is abundance of software that implements approaches to data processing, ranging from comprehensive packages (SAS, SPSS, etc.) to more limited in scope implementations, yet custom-built for specific purposes. From our team, examples of the latter are: IRENE-family tools [Fila et al., 2003a, b; Criscuolo et al., 2007] for model validation; ICARO [Silvestri et al., 2006] for assessment of environmental impact; AMPE [Acutis et al., 2007] for analytical methods evaluation; KeSTE [Paoletti et al., 2003] and SISSI [Confalonieri et al., 2007] for optimization of sampling strategies. All such approaches illustrate from different perspectives that there is actually a wealth of well-developed solutions to the basic problem of helping the scientists with their data collection, processing, and interpretation.

For scientists, it is desirable to retrieve analytical solutions and to process data from a free and extensible repository of alternative techniques. In order to improve the efficiency of use of resources and foster higher quality of analytical units, the component-based paradigm (which combines object-oriented and modular features) has become popular in the software business, based on the concept of encapsulating solutions to analytical problems in discrete, replaceable, and interchangeable units (components). Component-oriented designs actually represent a natural choice for building scalable, robust, large-scale applications, and to maximize the ease of maintenance in a variety of domains, including natural sciences data processing. The component development paradigm makes construction of software a case of plugging together independent components within an environment that implements the communication issues of components. The Java language has emerged with the aim to support inter-operability between different components and facilitating their integration. In this work, data-processing component-based software was designed in recognition of the need for the scientific communities to develop an organized approach to building the software that underlies data analyses. The aim was to comply with requirements of modularity, reusability/replaceability, extensibility, transferability, web accessibility, and open-source technology manageability.

The features of the software component are introduced and discussed in this paper to demonstrate that great utility in data processing can be gained through the implementation and use of object-oriented software components. The solutions implemented for application of fuzzy-logic based rules are also illustrated.

2. SOFTWARE COMPONENT STRUCTURE

The software component DANA (Data Analysis 'N' Assessment) contains functions for basic and advanced data processing procedures. Each procedure is implemented as a class. Simple procedures are implemented as independent classes while complex procedures are the result of an association of classes. The simplified conceptual view of Figure 1 shows iterations in the user interface-service-core communication flow.

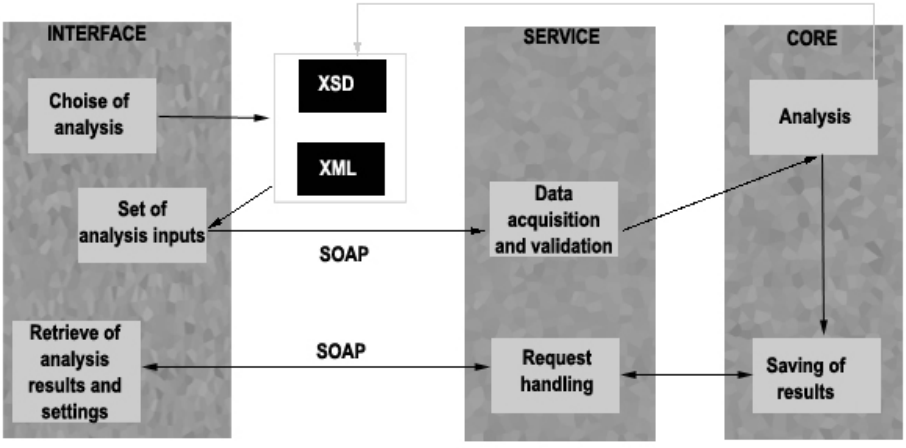


Figure 1. Conceptual view of software iterations.

Black arrows indicate interactions, either direct or SOAP [Simple Object Access Protocol]-mediated, generated at the execution time; grey arrow indicates a predefined dependency relationship. XML Schema Definition (XSD) is used to validate and ensure the interface compatibility with the system, since the Graphical User Interface (GUI) is automatically created through a generic XML file provided by each procedure available in the application. Figure 2 illustrates a portion of the XSD file used to validate any generic XML file implementing the GUI.

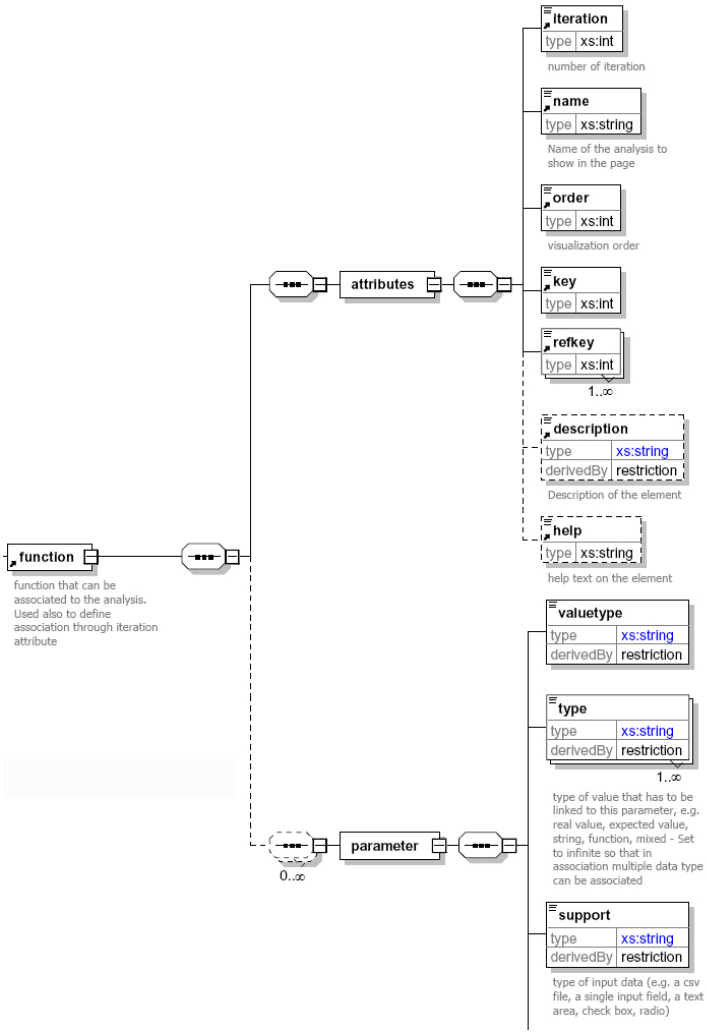


Figure 2. Portion of XSD schema used to validate a generic XML file.

The development of a reusable structure was based on object-oriented design. The application of this design philosophy led to the development of a hierarchy of software components that are manual controller independent and also have a standardized interface. The key design requirements for this architecture were: open-system, reusablility, application independent, extensive error-handling and safety checking, applicability to real-time data consistency and control, and reduced development time to implement new functions and analyses. Ease of maintenance is granted, also providing functionalities such as the test of input data versus their definition prior to computing any procedure. The component can be extended independently by third parties, adding new functionalities (extensibility) without the need of rewriting and/or re-compiling old code (modularity) if complying with the interface requirements. Moreover, external mathematical open source routines (e.g., classes from the R language statistical computing) were integrated to DANA

environment, in order to increase the computing solutions of the application. In order to fulfil the interface requirements, the application is dynamically linked to the functions available in the system at a given moment.

Based on the above modular approach, the implementation of DANA grants the possibility to replace existing functions (modules) with new ones (replaceability) as well as the usability of modules within other applications (reusability). The use of Java language assures the migration of the system to different platforms (transferability). For rapid development process and improvement of code readability, “design patterns” (e.g., Gamma et al., 1995) are used as effective solutions to describe how objects communicate without knowing each other’s data and methods. The design patterns “strategy” allows offering to the user different algorithms to achieve the same goal, called strategies. This solution emphasizes the issues of modularity and reusability. Web accessibility is granted by complying with the Web Accessibility Initiative strategies and guidelines developed by the World Wide Web Consortium (W3C), the requirements of consistency of behaviour and layout being assured through the use of templates in conjunction with Cascading Style Sheets. A Graphical Users’ Interface (GUI) is dynamically generated via PHP through a set of XML configuration files. Through these XML files the GUI will be dynamically modified with respect to specific analysis’ requirements (examples in Figure 3 and Figure 4). The component is either meant for users or developers interested in developing component-oriented solutions to data processing. The component is deployable and reusable in any application run over alternative operating systems.

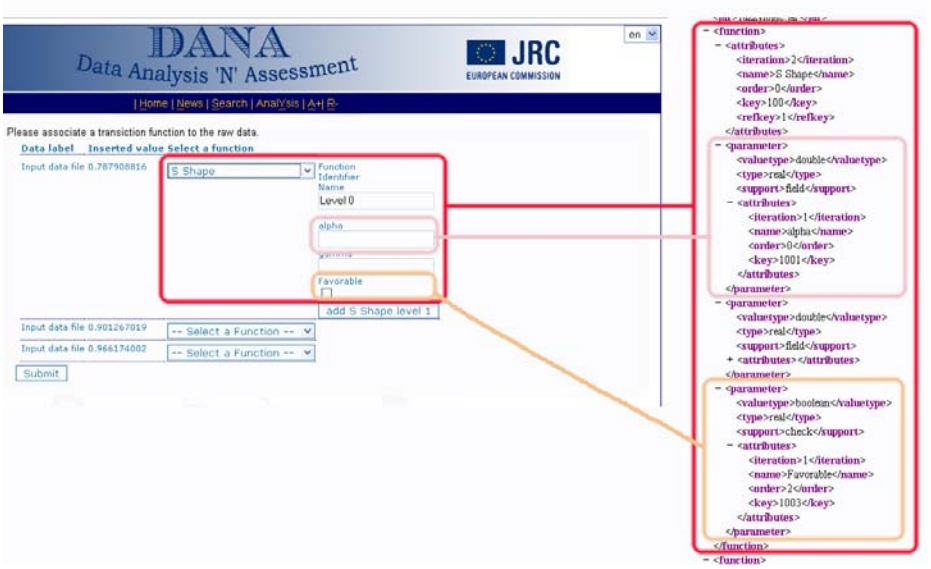


Figure 3. Example of GUI and the portion of XML file used to define it. In this example, S shape function and its parameters are defined in the XML file and automatically generated in the GUI.

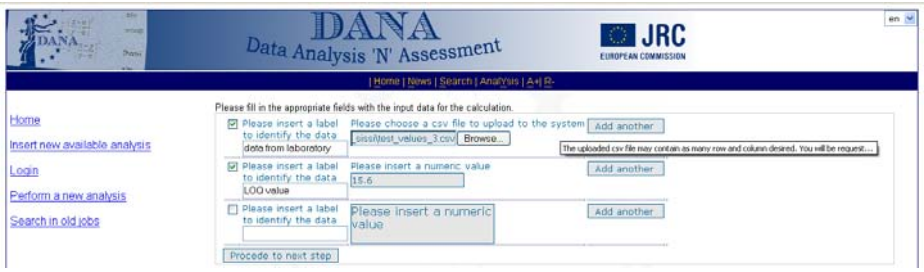


Figure 4. Example of input data screen, dynamically created from the XML file associated to a specific analysis.

3. COMPONENTS FOR FUZZY-LOGIC BASED RULES

Fuzzy logic derives from the fuzzy set theory originally introduced by Zadeh (1965). Fuzzy-logic based techniques are robust on uncertain and imprecise data such as subjective judgements and preferences [Ross, 2004]. For this reason, their use has been proposed [Hall and Kandel, 1991] and successfully applied to design expert systems in a variety of fields, including environmental modelling [e.g., van der Werf and Zimmer, 1998; Cornelissen et al., 2001; Ferraro et al., 2003; Rivington et al., 2005; Ocampo-Duque et al., 2006; Hamels et al., 2007]. Fuzzy logic uses sets of values with unclear boundaries, and can be used for mapping inputs to appropriate outputs. Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. This mapping provides a basis from which decisions can be made, or patterns discerned. Generally, the fuzzy inference process involves three basic concepts: membership functions, fuzzy set operations, and inference rules. A membership function is a curve that defines how each point in the input space is mapped to a membership value between 0 (no membership) and 1 (full membership). The standard fuzzy set operations are: union, intersection and additive complement. The inference rule is based on *if-then* statements, e.g., if *x* is A (antecedent) than *y* is B (consequent). A fuzzy inference system can be divided into four parts: fuzzification, weighting, evaluation of inference rules, and defuzzification. The fuzzification process involves the definition of inputs, outputs, and their respective membership functions that transform the numerical value of a variable into a membership grade to a fuzzy set. Since not all input variables may have the same importance, it is necessary to establish a way to weigh up the influence of each variable in the final score.

Dedicated standalone software tools are available where fuzzy-logic based rules are implemented for use in specific areas [e.g., Fila et al., 2003a, b; Silvestri et al., 2006; Acutis et al., 2007; Criscuolo et al., 2007; Sylaios et al., 2008]. A variety of approaches are available to generate fuzzy-logic rules [Ross, 2004] that are also exploited to aggregate variables of different nature into synthetic values [e.g., Bellocchi et al., 2002]. For example, monotonic membership functions – symmetric: e.g., linear, S shaped; asymmetric: e.g., logistic - are largely adopted to model the transition of a variable from one to another set. For situations where observational evidence indicates that the transition process is more complex, alternative functions (such as trapezoidal or bell shaped) are adopted for mapping the membership of a given input. Easy access to alternative approaches, amalgamated into one extensible tool, is desirable in order to tailor the approaches to purpose-specific projects and assess them in a comparative fashion.

In DANA, a set of components guide the fuzzification-aggregation-defuzzification process through multiple-choice dynamic steps granting the user high flexibility in operating alternative membership functions and weighing systems, and unlimited aggregation levels. The fuzzy logic computations are presently handled by two components, one based on S shape membership functions and one supplying a variety of alternative functions. The screenshots below illustrate what fuzzy-logic procedure displays on the screen at both an intermediate stage of aggregation (Figure 5) and at its cessation (Figure 6). In both figures, four exemplary numerical values (user-defined ‘Laboratory data’ and labelled from 0 to 3) are entered for aggregation. This step follows the selection of a membership function of the type illustrated in Figure 3. The user is required to both select a function and assign its values to a favourable or an unfavourable set. In Figure 5, data labelled 1 and 2, differently weighed (0.6 and 0.4, respectively) are first aggregated (Level 1, named by the user ‘2nd association level’). A second-level aggregation (Level 2, named by the user ‘final association’), based on the same principle, is also illustrated. A summary of the aggregations performed is shown in Figure 6 that is preliminary to defuzzification (not shown).

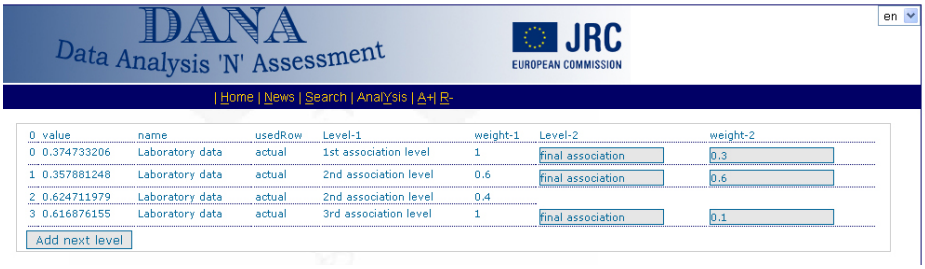


Figure 5. Fuzzy-logic based inference: exemplary second-level aggregation.

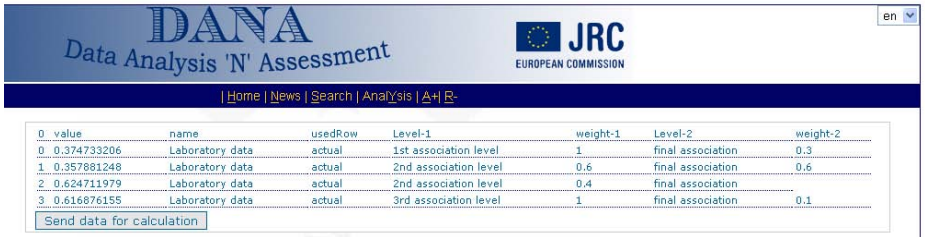


Figure 6. Fuzzy-logic based inference: exemplary cessation of aggregation.

4. DISCUSSION

Performing thorough experimental data analyses may appear as a tiresome process for different reasons. On one hand, people responsible for experimental design and data collection may not have sufficient background to deal with appropriate, computing-intensive data processing. Activities associated with data collection and data quality assurance tend to be performed by multiple groups working with a variety of software applications. Moreover, when novel analytical approaches come into light there is the need to have them implemented into ready-to-use, accessible software. Often, time and budgetary constraints do not allow accurate and complete analyses. The consequences of using insufficient data analyses inevitably are not to gain the insights from the structural/functional components of data variability, and also losing the relationships potentially existing between the variables investigated. Computer-aided tools are helpful to manage the quality of research in terms of definition of experimental requirements and criteria for analysis. To meet the substantial research quality challenges, it is necessary to improve the processing tools and technologies currently available, and their cost-benefit characterizations. The emergence of new technologies has fostered debate on the reuse and extension of analytical procedures within a variety of applications and platforms. A large number of statistical and soft-computing procedures have been implemented as software that cannot be easily maintained or reused, except by their authors, and cannot be transported to other platforms. Object-oriented programming has become associated with a rather easy methodology of prototyping and evolutionary software development that can be exploited in order to include legacy code into newly developed applications for data processing. The focus put on the use of design patterns encourages usability, reusability and cross-language compatibility, thus facilitating model development, integration, documentation and maintenance.

Keeping the above requirements in view, we have approached the DANA project aimed at designing a tool for encapsulation of solutions to data processing. Making it available for web use, DANA may serve as a convenient means to support collaborative studies among a large network of scientists. DANA was first and foremost developed in the context of policy-support and enforcement of the European Commission (EC) legislation for genetically modified organisms (GMO). For this reason, the project start-up has been the thorough implementation of methodologies (i.e., fuzzy-logic approaches) supporting collaborative validations of GMO analytical methods [Bellocchi et al., 2008], managed by

the Community Reference Laboratory for GM Food and Feed (http://gmo-crl.jrc.it) in accordance with Regulation (EC) No 1829/2003 [2003] and Regulation (EC) No 641/2004 [2004] and international guidelines (European Network of GMO Laboratories, http://engl.jrc.it). Statistical methodologies for support of sampling strategies along the food and feed chain [e.g., Bellocchi et al., 2007] will be next implemented as realization of Recommendation 2004/787/EC [2004]. The fuzzy-logic capabilities will be also expanded by combining fuzzy-based rules with additional tools for sensitivity analysis and optimization (neural networks).

Its generic features and capabilities, choice of a variety of rules and designs, display of multiple integrated analyses, make DANA an attractive and interesting tool for use in a variety of fields. The developers are committed to enrich its offerings, to ensure an evolving statistical methodology, and to grant easy access and communication. Frequent interaction with the end-users has been, and will continue to be, a priority for the development team.

5. AVAILABILITY AND DOCUMENTATION

DANA is meant to be made available for internet access and use. The application will be extensively documented through both an online help and a user manual. State-of-the-art software documentation technique is being used to widely explain key concepts and usage of the application. At the present stage, the application is under development and a preliminary version is running on local servers for testing and further development of its capabilities. Interested users may want to contact the corresponding author for information on the development status and availability.

6. CONCLUSIONS

The goals of DANA development are to extend easy access to computational methods to multiple users, and provide architecture to ensure reuse and extension of coded computing resources. DANA attempts to overcome some of the technical challenges that to date have limited the development of reusable computing (either general or specific-purpose) capabilities. DANA is also a way to share data processing knowledge in an extensible way.

ACKNOWLEDGEMENTS

This research has been developed within the framework of European Commission Joint Research Centre Action n° 15009 (Support to the Implementation of the GMO Legislation and Operation of CRLs), Action Plan 2008, Objective 07.00, Deliverables 07.01, 07.02.

REFERENCES

Acutis, M., P. Trevisiol, R. Confalonieri, G. Bellocchi, E. Grazioli, G. Van den Eede, and C. Paoletti AMPE: a software tool for analytical method validation, *Journal of AOAC International*, 90, 1432-1438, 2007.

Bellocchi, G., M. Acutis, G. Fila, and M. Donatelli, An indicator of solar radiation model performance based on a fuzzy expert system, *Agronomy Journal*, 94, 1222-1233, 2002.

Bellocchi, G., M. Acutis, C. Paoletti, R. Confalonieri, P. Trevisiol, E. Grazioli, C. Delobel, C. Savini, M. Mazzara, and G. Van den Eede, Expanding horizons in the validation of GMO analytical methods: fuzzy-based expert systems, *Food Analytical Methods*, DOI 10.1007/s12161-008-9021-8, 2008.

Bellocchi, G., C. Paoletti, E. Grazioli, R. Confalonieri, and G. Van den Eede, Complementary use of tools (KeSTE and SISSI) for sample size determination in GM kernel lots. In: Costa, J.F.C.L., and Koppe J.C. (Eds.), *Proceedings of the 3rd World Conference on Sampling and Blending*, October 23-25, Porto Alegre, Brazil, 469-478, 2007.

Confalonieri, R., M. Acutis, G. Bellocchi, and G. Genovese, Resampling-based software for estimating optimal sample size, *Environmental Modelling & Software*, 22, 1796-1800, 2007.

Cornelissen, A.M.G., J. Van den Berg, W.J. Koops, M. Grossman, and H.M.J. Udo, Assessment of the contribution of sustainability indicators to sustainable development: a novel approach using fuzzy set theory, *Agriculture, Ecosystem & Environment*, 86, 173-185.

Criscuolo, L., M. Donatelli, G. Bellocchi, and M. Acutis, Component and software application for model output evaluation. In: Donatelli, M., J. Hatfield, and A. Rizzoli (Eds.), *Proceedings of the International Symposium on Methodologies on Integrated Analysis on Farm Production System*, September 10-12, Catania, Italy, 2, 211-212, 2007.

Ferraro, D.O., C.M. Ghersa, and G.A. Sznaider, Evaluation of environmental impact indicators using fuzzy logic to assess the mixed cropping systems of the Inland Pampa, Argentina, *Agriculture, Ecosystem & Environment*, 96, 1-18, 2003.

Fila, G., G. Bellocchi, M. Acutis, and M. Donatelli, IRENE: a software to evaluate model performance, *European Journal of Agronomy*, 18, 369-372, 2003.

Fila, G., G. Bellocchi, M. Donatelli, and M. Acutis, IRENE_DLL: A class library for evaluating numerical estimates, *Agronomy Journal*, 95, 1330-1333, 2003.

Gamma, E., R. Helm, R. Johnson, J.M. Vlissides, Design patterns: elements of reusable object-oriented software, Addison-Wesley Professional Computing Series, Upper Saddle River, NJ, USA, 1995.

Hall, L.O., and A. Kandel, The evolution from expert systems to fuzzy expert systems. In: Kandel, A. (Ed.), *Fuzzy expert system theory*, CRC Press, Boca Raton, USA, 3-21, 1991.

Hamels, S., S. Leimanis, M. Mazzara, G. Bellocchi, N. Foti, W. Moens, J. Remacle, and G. Van den Eede, Microarray method for the screening of EU approved GMOs by identification of their genetic elements, Report of validation, Community Reference Laboratory for GM Food and Feed of the Joint Research Centre, Institute for Health and Consumer Protection, Biotechnology and GMOs Unit, 2007.

Ocampo-Duque, W., N. Ferré-Huguet, J.L. Domingo, and M. Schuhmacher, Assessing water quality in rivers with fuzzy inference systems: A case study, *Environment International*, 32, 733-742, 2006.

Paoletti, C., M. Donatelli, S. Kay, and G. Van den Eede, Simulating kernel lot sampling: the effect of heterogeneity on the detection of GMO contaminations, *Seed Science and Technology*, 31, 629-638, 2003.

Recommendation 2004/787/EC, On technical guidance for sampling and detection of genetically modified organisms and material produced form genetically modified organisms as or in products in the context of Regulation (EC) No. 1830/2003, 2004.

Regulation (EC) No 1829/2003 of the European Parliament and Council of 22 September 2003 on genetically modified food and feed, *Official Journal of the European Union*, 18.10.2003 L268, 01-23, 2003.

Regulation (EC) No 641/2004 of 6 April 2004 on detailed rules for the implementation of Regulation (EC) No 1829/2003 of the European Parliament and of the Council as regards the application for the authorization of new genetically modified food and feed, the notification of existing products and adventitious or technically unavoidable presence of genetically modified material which has benefited from a favourable risk evaluation, *Official Journal of the European Union*, 07.04.2004 L102, 14-25, 2004.

Rivington, M., G. Bellocchi, K.B. Matthews, and K. Buchan, Evaluation of three model estimations of solar radiation at 24 UK stations, *Agricultural and Forest Meteorology*, 135, 228-243, 2005.

Ross, TJ., Fuzzy logic with engineering applications, John Wiley & Sons, Hoboken, NJ, USA, 2004.

Silvestri, N., G. Fila, G. Bellocchi, and E. Bonari, An indicator to evaluate the environmental impact of olive oil waste water's shedding on cultivated fields, *Italian Journal of Agronomy*, 1, 243-257, 2006.

Sylaios, G.K., N. Gitsakis, T. Koutroumanidis, and V.A. Tsihrintzis, CHLfuzzy: a spreadsheet tool for the fuzzy modeling of chlorophyll concentrations in coastal lagoons, *Hydrobiologia*, DOI 10.1007/s10750-008-9358-4, 2008.

Van der Werf, H.M.G., and C. Zimmer, An indicator of pesticide environmental impact based on a fuzzy expert system, *Chemosphere*, 36, 2225-2249, 1998.
Zadeh, L.A., Fuzzy sets, *Information and Control*, 8, 338-353, 1965.

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

OpenMI Compliant Import of Initial and Boundary Data into a Numerical 3D Model

P. Schade[a], G. Lang[a], J. Jürges[a]

[a] Federal Waterways Engineering and Research Institute (BAW), Coastal Department, Wedeler Landstr. 157, 22559 Hamburg, Germany (peter.schade@baw.de, guenther.lang@baw.de, jens.juerges@baw.de)

Abstract: Two- or three dimensional hydronumerical models require complex initial and boundary data. Water level elevation data along an open ocean boundary varies in space as well as in time and is typically applied to an estuarine model. Reading model specific files is the classic method for importing the data. If several independent models are used in parallel, each one will require its own file format or programmable interface. Although these multimodel forecasts may deliver more reliable results, the additional effort raises the threshold for accomplishing the work in a reasonable amount of time. The OpenMI compliant approach provides a standardized data import. Software components collaborate in an integrated modelling system. One component reads the data, whereas another component, the numerical engine, imports it. The Reader GEIWrapper contains a Generic Engine Interface (GEI), enabling the user to access various types of data formats with a single set of methods. The format specific methods are encapsulated by the GEI. The component WLDelft Wrapper includes the Delft3D numerical engine. Both wrappers have been migrated from existing software. The module accessing the initial data has also been integrated in the numerical model UnTRIM. This enables Delft3D and UnTRIM to import from identical files and to employ the same methods thus simplifying their use in multimodel applications.
The import into Delft3D works with a great variety of data and has turned out to be markedly comfortable. For example, spatial interpolation between the locations of the input files and the model grid can be done automatically. The OpenMI environment supports temporal interpolation as well.

Keywords: OpenMI; Delft3D; estuarine model; multimodel application

1. INTRODUCTION

The processes simulated in estuarine hydronumerical models include water level elevation, flow velocities, waves and salinity as well as sediment transport. The models require initial data e.g. the distribution of salinity in the entire model area. Furthermore they import boundary data, which may vary in space and time. Typical examples are the water level elevation varying along an ocean boundary and total discharge at an upstream boundary. The preparation of this data and its adaption to different numerical models may require considerable effort.

The Open Modeling Interface (OpenMI) [Gregersen et al., 2007] is based upon the idea of an integrated modelling system. Several models are connected by a unified linking mechanism. Although the linkage should work for all time based simulations, it aims at usage in the water and possibly in the environmental domain. The OpenMI standard interface specification [Gijsbers et al., 2005] prescribes the concrete attributes. A software component that fulfils these requirements is called a Linkable Component.

Quantity and elementset are further central OpenMI terms. A quantity defines what has been exchanged between models. The elementset offers information where or at which locations these values have been exposed.

It is one objective of OpenMI to restrict the modellers' freedom as little as possible. This has lead to a flexible model definition. According to the standard a model is an entity that provides and / or asks for data. Therefore, a reading component with data base functionality is a model as well. Furthermore OpenMI aims at user friendliness. Some utilities and tools have been added to the OpenMI environment. The OpenMI Configuration Editor, e.g. enables the generation and running of model configurations by mouse-click.

2. INTEGRATION OF DELFT3D INTO THE WORK ENVIRONMENT

2.1 Workflow of hydronumerical models

The conclusion of a hydro-engineering expertise is not based on the model run alone, but on input data and the interaction of the simulation itself with its pre- and postprocessing (Figure 1). A complete work environment for the numerical model UnTRIM [Casulli et al., 2002] already existed at our institute, the BAW. When integrating the numerical model Delft3D [Lesser et al., 2004] into this environment, several alternatives for the import of initial and boundary data were taken into consideration. An offline conversion into Delft3D specific file formats would have meant redundant storage of identical data. A new reader of proprietary BAW data specifically for Delft3D would not have worked for other numerical models. The BAW had a more generic approach in mind, which should ensure the re-use of already existing and newly developed software components. Furthermore, a great variety of data should be processable. These requirements were the main reasons for selecting the OpenMI. Secondly, the Reader could be used for any OpenMI compliant numerical model as indicated by the rectangle designated "…" in the figure. All numerical models at the BAW use unified postprocessing methods which analyse complete time-series in the proprietary Binary Data File (BDF) format.

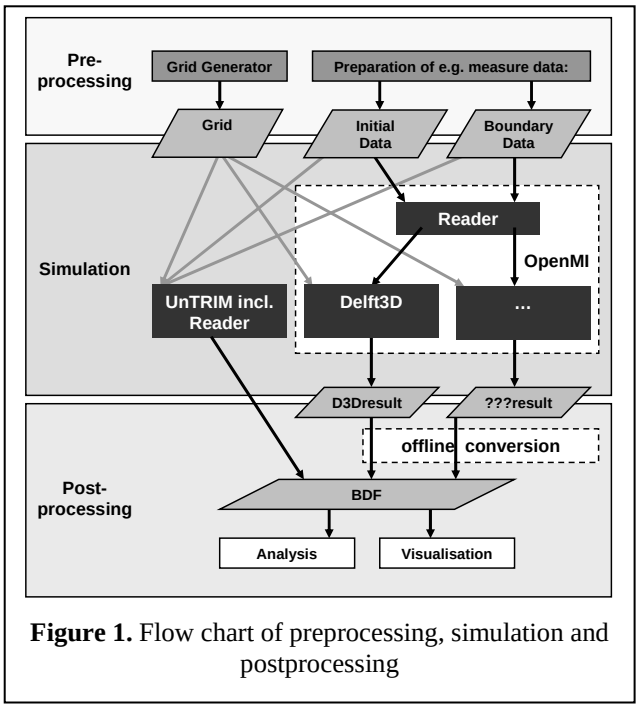


Figure 1. Flow chart of preprocessing, simulation and postprocessing

2.2 Import by the Linkable Components GEIWrapper and WLDelft Wrapper

The OpenMI compliant import of the initial and boundary data into Delft3D occurs during the simulation. Figure 2 displays the content of the dashed OpenMI rectangle from Figure 1 in more detail. Two Linkable Components have been developed. The GEIWrapper reads the input data, whereas the WLDelft Wrapper accepts this data as an input for Delft3D. The data flow from the first to the second component occurs by means of the methods provided by the OpenMI environment.

The development of the WLDelft Wrapper by WL | Delft Hydraulics (Deltares) and the GEIWrapper by the BAW resulted from a close collaboration of both institutes. Existing Fortran 90 code has been migrated into OpenMI compliant components

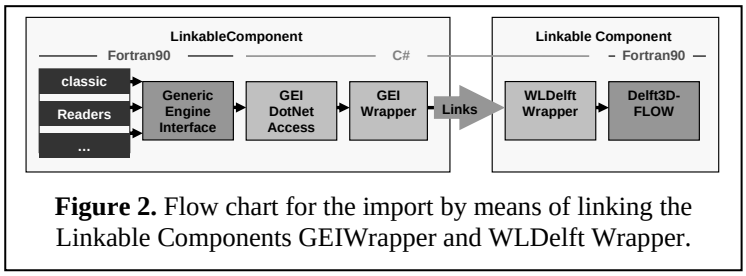


Figure 2. Flow chart for the import by means of linking the Linkable Components GEIWrapper and WLDelft Wrapper.

The Generic Engine Interface (GEI) is the core of the GEIWrapper. It enables a generic access to several file formats. Data in the BDF format, the SELAFIN format of the Telemac model or in the netCDF format is on the lowest level accessed with specific classic Readers. Since these components are encapsulated by the GEI, the external access works with a single set of standardized methods. The user calls, for example, only one central GetValues-method for "pulling" the data into the model. Implementing individual Fortran GetValues-methods for each of the three formats mentioned above was a sizeable task. But the increased complexity is more due to BAW specific requirements than to OpenMI compliance. The internal structure of the Fortran 90 code was rearranged and extended. Internally object-oriented modules represent the basic interfaces defined in the standard. These modules can be reused for further OpenMI compliant Fortran 90 engines. The GEI itself is more specific. Its methods show parameters of intrinsic data types, simplifying the embedding in components written in other languages. The final wrapping of the GEI with a C# component was comparatively easy. Again, object-oriented techniques lead to a set of only 17 methods, which cover the required functionality. Though the final linking process is controlled by the .NET technology (C#), the internal engine runs with unmanaged Fortran 90 code. The bridge between both worlds is stable.

All requirements of this project were fulfilled by the OpenMI standard. Scalar and vector values can be processed. The technique accepts ID-based as well as geo-referenced data. For the latter OpenMI provides methods for spatial interpolation. The interface standard allows geo-referenced data to be one-, two- or three dimensional. Though the GEIWrapper mainly offers 2D data, it is used to run Delf3D in 2D as well as in 3D mode. The WLDelft Wrapper has meanwhile been integrated into the Delft3D main development line.

Verification is conducted in three stages including unit testing. On the highest level the complete import sequence is tested in the configuration editor. 40 test scenarios show that the OpenMI compliant import is able to process a great variety of data.

3. CATEGORIES OF IMPORTED DATA

3.1 Initial Data

Initial data for the estuarine models at the BAW is stored in the proprietary format ipds.dat (Initial Physical DataSet). The io_ipds software package processes the values which can be assigned to regions or to sampling points. The spatial interpolation to all vertices of the underlying grid can be done with either of two interpolation methods. The first method looks for the nearest sampling points in various sectors. The second employs a Delauny triangulation. The OpenMI compliant import benefits from integrating these methods by migration.

The following list displays the many faces of the imported initial quantities:

- water level elevation
- salinity
- temperature
- constituent

- depth
- roughness
- erosion / deposition parameters
- sediment mass at the bottom

The initial salinity distribution in the Ems Estuary (Figure 3) is based on former simulations. The OpenMI Configuration Editor controls the communication between GEI- and WLDelft Wrapper. The GEIWrapper assigns the salinity values to the vertices of the underlying initial data grid. If the user selects in the data accepting WLDelft Wrapper an elementset designated as "<provided>", the interpolation to the Delft3D grid will be done automatically.

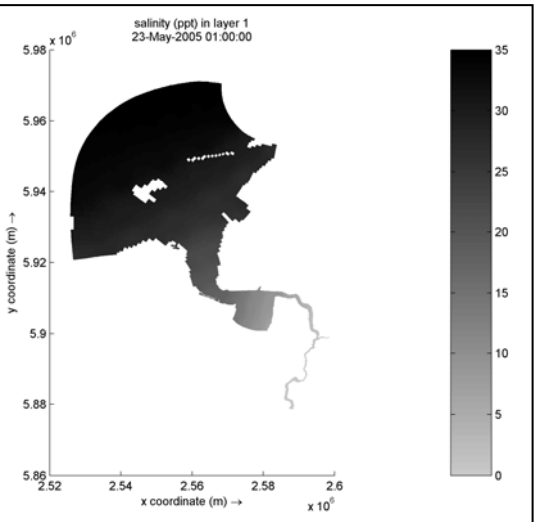


Figure 3. Initial salinity in the Ems Estuary.

3.2 Open Boundary Data

Open boundary data is time dependent. If necessary, a linear interpolation of the provided date and time will be done. The GEIWrapper exposes the values on elementsets of type XYPoint, which are vertices with x- and y-coordinates. The elementset of the WLDelft Wrapper can be of a different element type, e.g. line, since the GEIWrapper takes care of the type conversion. The model of the Ems Estuary is limited by two open boundaries to the North Sea (Figure 4). Delft3D requires water level elevation and salinity as boundary conditions at these elementsets "XYPoint-East" and "XYPoint-North". The ElementSet "AllLocations" delivers all open boundary vertices on the data providing side. By establishing a link from "AllLocations" to "XYPoint-East" each XYPoint is assigned to the nearest XYPoint in "AllLocations". The same procedure is done for "XYPoint-North". A mapping matrix is generated according to which the water level elevation is transferred during the computational phase. The user has to perform few mouse-clicks in order to assign the data. This process has turned out to be markedly comfortable, whereas the classic manual assignment can be time-consuming and error-prone.

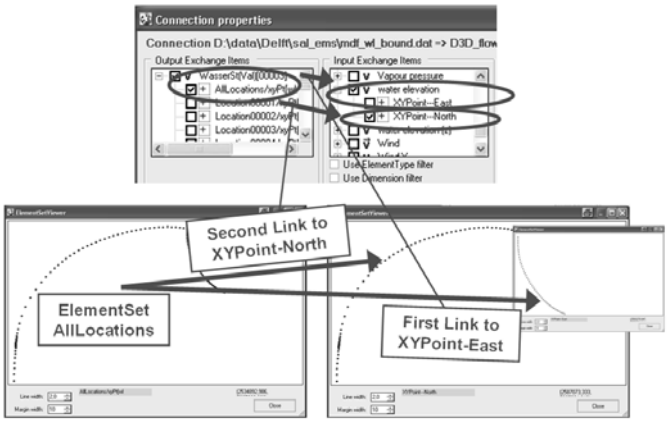


Figure 4. Import of complex water elevation boundary conditions by establishing two links in the OmiEd (top).

Furthermore, the classic import is currently restricted to a file size of 4 Gigabyte. If the import is done OpenMI compliant on the fly, only input data of one or two time steps will be needed at a time. Thus OpenMI enables long-term simulations with periods of one year or longer.

3.3 Free Surface Boundary Data

Meteorological quantities at the free surface of a water body influence the hydrodynamics. Most of the quantities are globally constant. However, pressure and wind are exposed on 2D fields at XYPoints. The transient wind impact becomes important for storm surge forecasts. The BAW runs a model of the North Sea (Figure 5), which reproduces the high water values in the German Bight during the All Saints' Day storm surge of 2006. The wind field is stored on a coarse grid. WLDelft Wrapper interpolates the data to the vertices of a Delft3D grid with a higher resolution. Again, it is a matter of establishing one link in the OpenMI Configuration Editor.

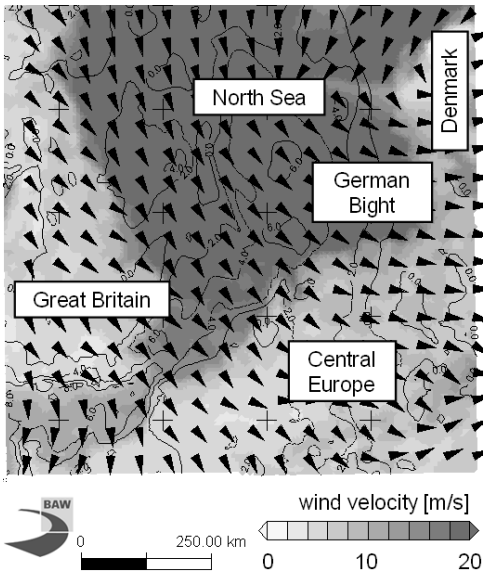


Figure 5: 2D vector wind field over the North Sea during the storm surge of 11/01/2006. By courtesy of DWD.

3. 4 Sources and Sinks

The Reader GEIWrapper offers sources and sinks on elementsets of type XYPoint. The WLDelft Wrapper offers sources and sinks on elementsets of three different types. If the user selects "<provided>" on the data accepting side, the discharges defined in the GEI-Wrapper will be implanted at the correct position in the Delft3D grid. Two cooling water outlets of a power plant at the Elbe Estuary have been generated this way (Figure 6). BAW simulations rarely include sources and if so, it is usually only a single source. In this case the efficiency of the classic import via Delft3D specific files and the OpenMI compliant import is nearly the same. OpenMI shows its excellence in more complex applications.

4. MULTIMODEL APPLICATIONS

4.1 Introduction

The reliability of computational results is an intensively discussed issue. Although a model has been accurately calibrated, some uncertainties remain. Is the quality of the input data high enough? Do the equations reproduce the physical processes? Is the parameterization correct? It is difficult to answer these questions with a single simulation. Multimodel applications do not deliver a single solution but an ensemble of results. Analysing the ensemble with stochastic methods may result in more reliable forecasts. The term multimodel can refer to the use of

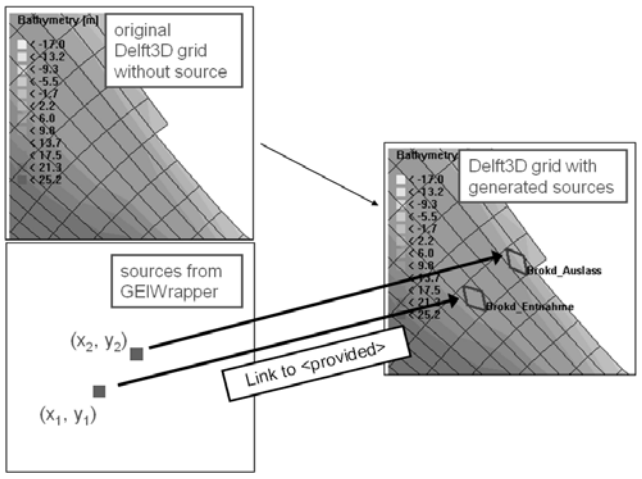


Figure 6. Implantation of two sources into a Delft3D grid by establishing a single link

various independent models with their individual approaches. It may as well mean that only one numerical system is used but that input data and parameterization vary. Combining both types is also possible. Originally, multimodel ensemble forecasts were used in the meteorological and climate domain [Palmer et al., 2004]. The idea spread and ensembles were amongst others used in morphodynamic simulations [Plüß et al., 2007].

Though this approach can provide the modeller a better insight, it does require a sizeable effort. Especially running two or more independent models in parallel often means an individual pre- and postprocessing. With the standardized OpenMI interface the preprocessing has to be done only once.

4.2 OpenMI and multimodel applications

We describe a hypothetical example in the following. The numerical engines UnTRIM and Delft3D have been used according to this approach. UnTRIM imports its initial and boundary data directly from proprietary files. OpenMI and the GEIWrapper enable Delft3D to access the identical initial and boundary data files. Furthermore, UnTRIM and Delft3D use the same interpolation methods for initial data, which eliminates the import as a possible cause for different results. An UnTRIM and a Delft3D model of the Ems Estuary give comparable results. The impact of a man-made salinity discharge is analysed. Both models run in 3D mode since salinity stratification may be possible. In order to raise the comparability, the UnTRIM grid is a clone of the Delft3D grid. It has been converted from the original Delft3D grid. The actual state of both models without source has been calibrated. Subsequently the salinity discharge with 2.78 m^3/s and a concentration of 300 ppm has been set in both models. According to Figure 1, the results of all four runs undergo the same postprocessing. A multitude of tidal characteristic numbers [Lang, 2008] can be computed. The Figures 7 and 8 display tidal characteristic numbers of salinity on Ems profiles. The maximum salinity diagram in Figure 7 shows that eight grey curves for the eight tides are calculated within the period of data analysis. They form a band with the black curve as the mean of these maximum salinity values in their midst. The Δ (maximum salinity) diagram summarises the forecast for the impact of the source on the mean maximum salinity. UnTRIM and Delft3D predict similar source induced changes along the navigation channel.

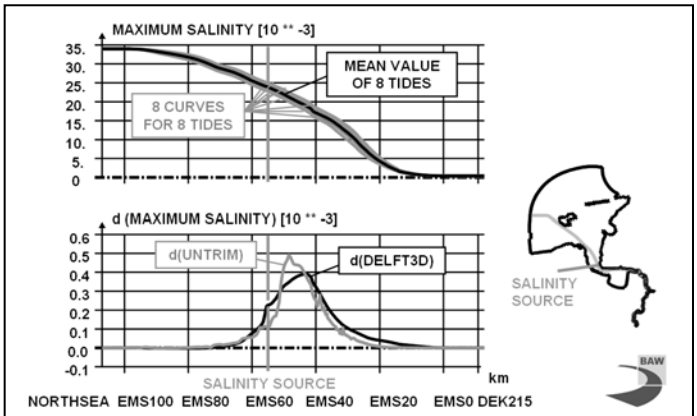


Figure 7. The maximum salinity (top) and the Δ (maximum salinity) (bottom) along the Ems navigation channel.

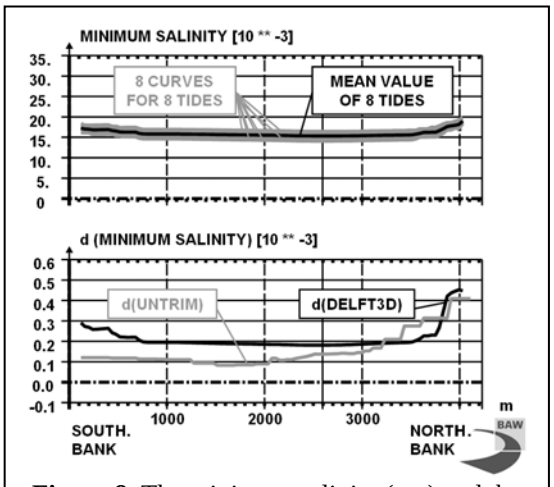


Figure 8. The minimum salinity (top) and the Δ (minimum salinity) on a cross profile located 6km upstream from the salinity source.

The minimum salinity has been analysed on a cross profile located six km upstream from the discharge (Figure 8). The absolute Delft3D

values in the upper diagram show only little variation in the actual state. The lower diagram indicates that the impact of the salinity source, located at the northern bank, is for both models stronger at this bank. A systematic deviation becomes visible towards the southern bank. Delft3D predicts a higher difference between the values with and without salinity source than UnTRIM. This means that water with higher and lower salinities mixes faster in the Delft3D model. The parameter for horizontal diffusivity has most probably an impact on this behaviour. More runs with changed parameterisation could support this hypothesis, which would imply an additional multimodel ensemble.

5. CONCLUSIONS AND RECOMMENDATIONS

Multimodel applications evaluate the quality of different models by comparing their results. Reading from the same files has two effects. First of all, running two models in parallel becomes easier. Secondly, the research into the functionality of a model is simplified, since different results can not have been caused by different input data.

The methods for spatial and temporal interpolation, which are offered by the OpenMI environment, have made the interactive linking markedly user-friendly. Simple processes, such as, e.g., importing data from a single source, do not benefit greatly from this technique. We recommend these methods for data, which is transient and exposed at many locations. At the BAW this mainly applies to data at open and free surface boundaries.

Importing data into a complex 3D model demonstrated the versatility of the OpenMI standard. Several file formats and data of different categories can be accessed with one component. The classic way of integrating numerical models into the workflow costs a sizeable effort for each input format and each model. On the other hand, making a model or a reader OpenMI compliant implies an overhead. But large portions of this work have to be done only once. GEIWrappers as well as WLDelft Wrappers can be connected to other OpenMI compliant models. Future applications will benefit from the knowledge acquired by developers and users.

ACKNOWLEDGEMENTS

The authors would like to thank the associates of the Deltares (formerly WL | Delft Hydraulics) software group for making Delft3D OpenMI compliant. Their support during the development of the GEIWrapper and the verification was very helpful. The project also benefits from the detailed documentation and the training courses offered by the members of the OpenMI Association. Andreas Plüß and Ulrich Vierfuß (BAW) were so kind to supply us with models of the North Sea and the Elbe Estuary.

REFERENCES

Casulli, V. and Zanolli, P., Semi-Implicit Numerical Modelling of Non-Hydrostatic Free-Surface Flows for Environmental Problems, *Mathematical and Computer Modelling, 36*, 1131 – 1149, 2002.

Gijsbers, P.J.A., Brinkman, R., Gregersen, J.B., Hummel. S. and Westen, S.J.P., The OpenMI Document Series: Part C, The org.OpenMI.Standard Interface Specification (version 1.0). *Available at http://www.OpenMI.org*, 2005.

Gregersen, J.B., Gijsbers, P.J.A., and Westen, S.J.P., OpenMI: Open Modeling Interface, *Journal of Hydroinformatics, 9*(3), 175-191, 2007.

Lang, G., Analysis of Calculated Results; Tidal Characteristic Numbers, *http://www.baw.de/vip/en/departments/department_k/methods/kenn/kenn-en.htm*, 2008.

Lesser, G.R., Roelvink, J.A., van Kester, J.A.T.M. and Stelling, G.S., Development and validation of a three-dimensional morphological model, *Coastal Engineering, 51*(8-9), 883-915, 2004.

Palmer, T.N., Alessandri, A., Andersen, U., Cantelaube, P., Davey, M., Délécluse, P., Déqué, M., Díez, H., Doblas-Reyes, F.J., Feddersen, H., Graham, R., Gualdi, S., Guérémy, J.-F., Hagedorn, R., Hoshen, M., Keenlyside, N., Latif, M., Lazar, A., Maisonnave, E., Marletto, V., Morse, A.P., Orfila, B., Rogel, P., Terres, J.-M. and

Thomson, M.C., Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (Demeter), *Bulletin of the American Meteorological Society*, *85*, 853-872, 2004.

Plüß, A. D., Heyer, H., Morphodynamic Multi-Model Approach for the Elbe Estuary, *River, Coastal and Estuarine Morphodynamics: 5th IAHR Symposium (RCEM 2007), Proceedings*, 113-117, 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

Demonstration of Integrated Modelling using the OpenMI in the Scheldt River Basin

Neel Devroede[a], Yves Ronse[a], Johan Van Assel[b] and Hans Vereecken[c]

[a] *Vlaamse Milieumaatschappij (VMM), A. Van de Maelestraat 96, 9320 Erembodegem, Belgium (n.devroede@vmm.be, y.ronse@vmm.be)*

[b] *Aquafin, Dijkstraat 8, 2630 Aartselaar, Belgium (johan.vanassel@aquafin.be)*

[c] *Flanders Hydraulics Research, Berchemlei 115, 2140 Antwerpen, Belgium (hans.vereecken@mow.vlaanderen.be)*

Abstract: Water management in the transnational Scheldt River Basin (Belgium, France, the Netherlands) is scattered among many different authorities and operators. For several years, most of them have adopted modelling as a technology for optimising investment and operation strategies for the part of the water system that is under their responsibility. As the European Water Framework Directive imposes water management to be integrated across both authorities and water domains, there is a clear need to streamline and integrate the various modelling efforts. However, many of those models have been developed completely independently from each other, with inconsistent spatial boundaries, and using different approaches and objectives. Hence, integrating these models is far from straightforward. The development and release of the OpenMI (Open Modelling Interface) standard in 2005 offered a potential solution to linking models of various origins and concepts, and the challenge was taken up to try and apply this new standard at full scale on real operational models. In the frame of the demonstration project OpenMI-Life, four use cases were defined within the Scheldt basin, in which various aspects of model linking will be tested. By the end of the project, it is hoped that water managers will have better insights of how interactions between water systems may affect strategic decisions.

Keywords: OpenMI, model linking, integrated water management

1. INTRODUCTION

Due to its transnational character and its geographic location in the heart of one of the most densely populated areas in Europe, the Scheldt River Basin is a good example of how complex and challenging integrated water management is in practice.
From an institutional point of view, the International Scheldt Commission co-ordinates the policies of no less than six member states/regions (France, the federal state and three regions of Belgium, and the Netherlands). But even within these six entities, there is a variety of authorities and operators involved in the different water domains and water industry sectors.
From a water management point of view, there are important issues about flood protection and water quality. Furthermore, the economic role of the river basin, which contains the Antwerp, Ghent and Zeebrugge ports along with many heavily used canals, is not to be underestimated.

During the preparation of the OpenMI-Life project, four use cases were identified (Vits and Devroede [2007]), which were thought to give a good picture of the possible interactions between the main large scale modelling programs that have been running for the last ten years. These use cases are :

- linking a sewer model with a hydraulic river model
- linking two different hydraulic river models for navigable and non-navigable water courses
- linking a river quality model with two hydraulic river models
- linking a 1D-river model with a 2D-estuary and coast model.

Details of the four use cases are described further.

The aim of the OpenMI-Life project is to demonstrate that the OpenMI standard (Gregersen et al. [2007]) and its implementations can provide a technical solution for the different linking problems in the four use cases, and that the OpenMI Association as a support organisation can come up with solutions for those areas where further improvement and development to the Standard would still be required. Besides the use cases in the Scheldt River Basin, tests are conducted in the Greek Pinios River Basin.
The project is supported by the European Life Programme and is co-ordinated by the UK Centre for Ecology and Hydrology (CEH). It started in October 2006 and lasts until January 2010. More information on the project can be found at www.openmi-life.org.

2. DESCRIPTION OF THE USE CASES

2.1 Use case A : sewer-river interactions in the drainage area of Leuven

In the first use case a (hydraulic) sewer model is linked to a hydraulic river model. The models respectively describe the urban drainage area around the town of Leuven (appr. pop. 120000, appr. area 120 km²) and the river Dijle with its main tributaries between the Walloon/Flemish regional border and the confluence with the river Demer (appr. length 40 km, appr. area 300 km²) (Figure 1). Partners involved in this use case are Aquafin (the company responsible for building and operating the wastewater treatment plants and main trunk sewers in Flanders) and the Division Operational Water Management of VMM (Vlaamse Milieumaatschappij – Flemish Environmental Agency), responsible for the non-navigable watercourses in Flanders. The models have been built in InfoWorks CS (sewers) and InfoWorks RS (rivers) (Wallingford Software, UK).

Simulating the models in linked mode is expected to lead to an improved forecast of flooding, both in the sewer system and in the river, and will provide new insights and opportunities for optimising the investment schemes and operational management for both the sewer and the river system.

Under normal conditions two types of interactions can be defined. Firstly, the sewer system discharges into the river system at various locations, such as permanent outfalls, overflows and at the waste water treatment plant. Secondly, high water levels in the river system can prevent free discharging from the sewer system. Flows may occasionally revert in these cases where outfalls are not protected by a flap valve.
Under flood conditions additional exchange of water can occur between sewer and river system, as the river may flood certain sewer manholes, causing the river water to enter the sewer system, or flooded manholes may spill (diluted) sewage into river flood areas.

In OpenMI terminology this means that the quantities exchanged are flows (from sewer to river model) and water levels (from river to sewer model). Although this does not look particularly complicated, it is the high number of links (more than 100 in normal

conditions and probably an even higher number for flood conditions) and the fact that all links are bidirectional, which makes up the technical challenge for this use case. All these exchanges will lead to a continuous and dynamic flow redistribution between both models, which would never be achievable using predefined boundary conditions.

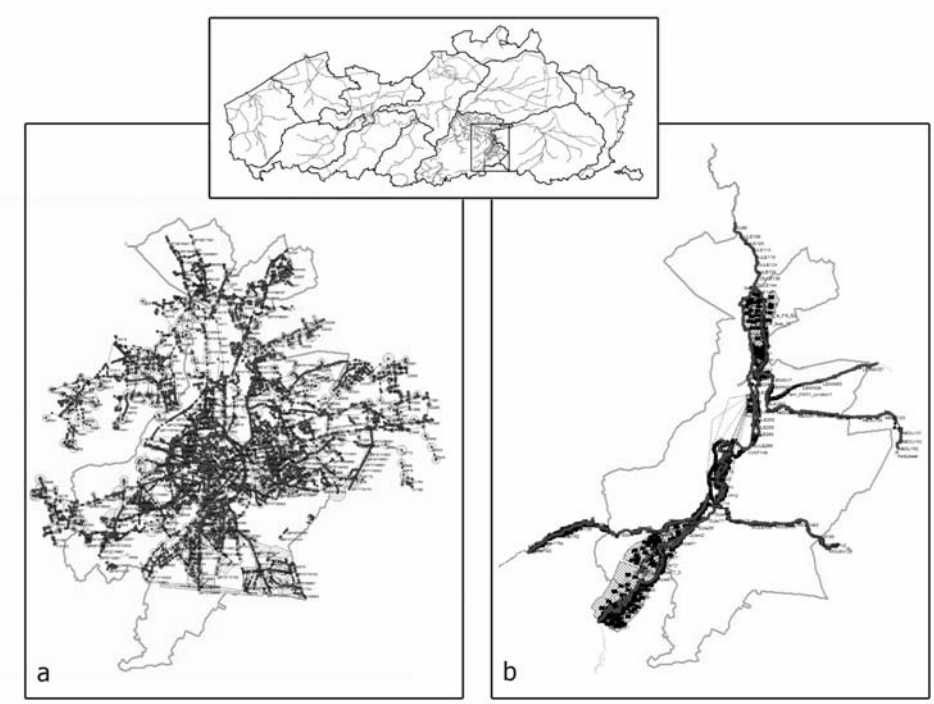


Figure 1. Geographic setting (top) and models for use case A : InfoWorks CS sewer model of Leuven (left, a) and InfoWorks RS Dijle river model (right, b)

2.2 Use case B : linking Scheldt and Dijle river sub-basins using two different hydraulic river models

The second use case comprises the linking of two independently built hydraulic river models. The first one describes the subbasin of the river Dijle and its tributaries, upstream from the confluence with the river Demer (the same model as used in use case A). The second one describes the tidal part of the Scheldt river and its tributaries, including the river Dijle downstream from the confluence with the river Demer (Figure 2).
The end of the tidal and navigable zone, which forms the split point between the two models, also delineates the competence area of both partners in this use case : on the one hand the Division Operational Water Management of VMM; on the other hand the Divison Flanders Hydraulics Research of the Flemish Ministry of Mobility and Public Works. The models have been built in InfoWorks RS (Wallingford Software, UK) and Mike11 (Danish Hydraulic Institute, Denmark).

Simulating the models in linked mode, thus avoiding the need for setting up appropriate and reliable upstream and donwstream boundary conditions, is expected to improve the accuracy of flood forecasting in both models. By linking the models both competent authorities can take into account the impact of operational flood management (such as the use of retention basins) in each other's parts of the river basin.

As for use case A, the quantities exchanged are flows (from Dijle model to Scheldt model) and levels (from Scheldt model to Dijle model), again defined as bidirectional links. The number of exchange points however is very limited, even with the occurrence of mazed tributaries in the boundary area between the two models. As both original models have an

overlapping zone (in order to dampen the immediate impact of boundary conditions), alternative scenarios for the definition of the links will be investigated (e.g. with the flow exchange not necessarily occurring at the same location as the water level exchange). When looking at the models in flood conditions, flow and level exchange will be applied not only on the main river channel, but also in the flood zones.

Finally, a clearly different technical challenge, as opposed to use case A is the fact that use case B deals with models from different suppliers.

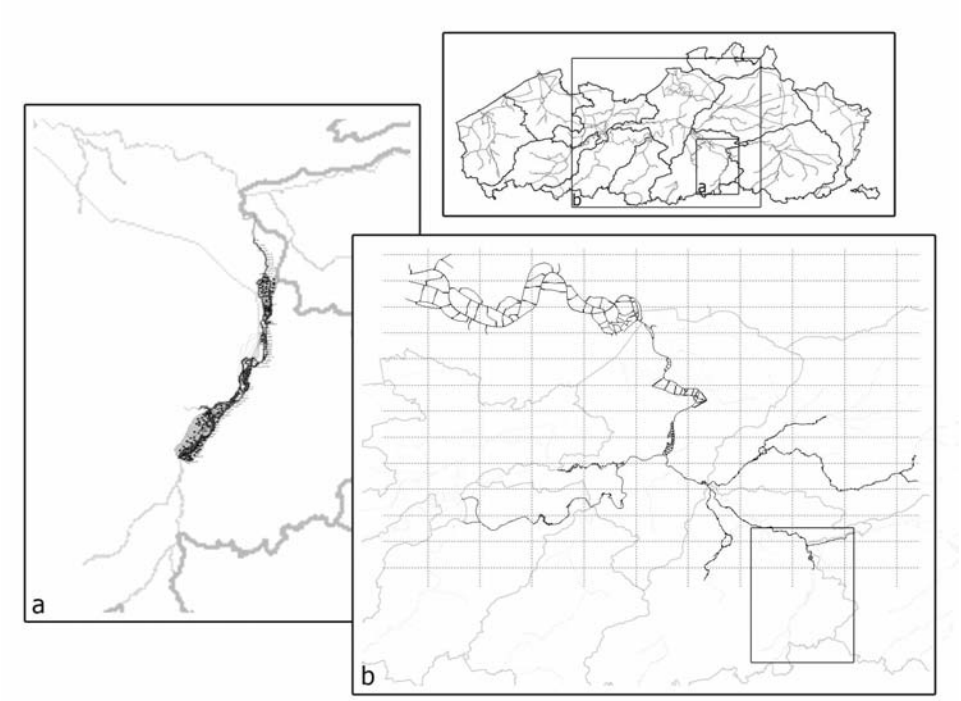


Figure 2. Geographic setting (top) and models for use case B : InfoWorks RS model of the river Dijle (left, a) and Mike11-HD model of the river Scheldt and tributaries (right, b)

2.3 Use case C : linking a river quality model with two different hydraulic river models in the Dijle and Demer sub-basins

In use case C parts of the two aforementioned hydraulic river models will be linked (one at a time) with a river quality model, which describes the whole of the river Dijle and river Demer basins (including the Walloon part of the river Dijle) (Figure 3).
The river quality model PEGASE was developed by the University of Liege, Belgium and is used by the Division Water Quality Management of VMM in view of developing its surface water quality management plans. It has a built-in hydrological module, which – based upon flow observations from river gauges- can produce flow patterns for the river branches. By linking the PEGASE model with the hydraulic river models InfoWorks RS and Mike11, it is expected that the flow calculations will become much more accurate compared to the ones produced by the built-in hydrological module. This in turn will improve the accuracy of the river quality calculations, as these are obviously very dependent from the velocities. In areas without a InfoWorks or Mike11 feed, PEGASE will continue to use its own flow calculations. Point inflows to the river (waste water treatment plants, industries) will continue to be taken from the PEGASE input database.

Besides the expected improvement of the water quality calculations, the linking of the models will enable water quality and river managers to account for the expected quality of the flood water in the decision process of the construction and operation of flood zones.

The technical aspect of the linking is different from the two aforementioned use cases in so far that the models are sharing the same geographical area. This means that the linking is to be seen as a global overlay rather than as a point-to-point link as in use cases A and B. Inconsistencies in the details of the river schematisations in both models form a specific point of attention when applying this.
Quantities exchanged are water depths, flows and velocities (all from the hydraulic river models to the river quality models).

Contrary to the other use cases, it is also to be mentioned that the PEGASE model was not yet OpenMI compliant before the start of the OpenMI-Life project. Hence, the process of migrating a model is another element in demonstrating the applicability of the OpenMI.

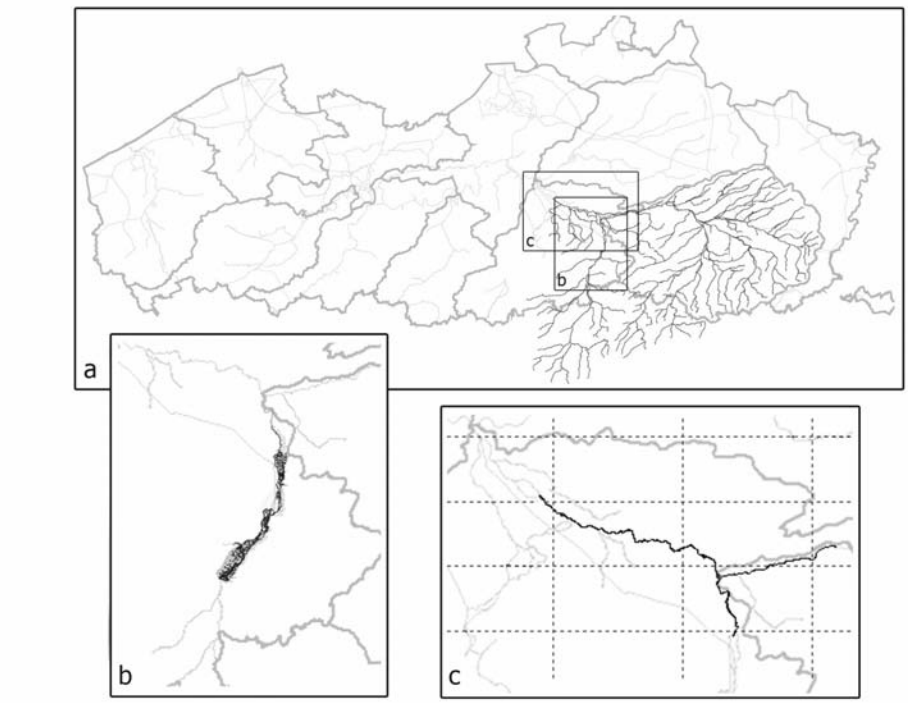


Figure 3. Geographic setting (top, a) and models used for use case C : PEGASE river quality model for Dijle and Demer (top, a), InfoWorks RS hydraulic river model for Dijle (left, b) and Mike11-HD hydraulic river model for Demer (right, c)

2.4 Use case D : linking 1D-river model to 2D-estuary models in the Dender sub-basin and main Scheldt basin.

The fourth and final use case describes the linking of a 1-dimensional hydraulic river model to a 2-dimensional estuary and coastal model. The first one is the model of the Flemish part of the river Dender basin (built by Flanders Hydraulics Research). In a later stage of the project it is hoped that this could be replaced by the full tidal Scheldt model. For the second one, two different options will be investigated: on the one hand the "Kustzuid" model, on the other hand the "Zeekennis" model (Figure 4).
"Kustzuid" is an operational model from the Dutch authority Rijkswaterstaat, built with the WAQUA software (currently maintained by Deltares). It covers the whole Scheldt estuary and a large part of the North Sea. For the purpose of the OpenMI-Life project it was extended to the confluence of the Dender and Scheldt rivers. "Zeekennis" is a morphology oriented model, built in Delft3D, covering a smaller area than "Kustzuid".

Linking the 1- and 2-D models is expected to increase the accuracy of flood prediction (especially in a later phase when the Dender model would be replaced with the full 1D tidal Scheldt model), and to improve the forecasting of the accessibility for large vessels of the port of Antwerp.

As for use cases A and B, quantities exchanged are flow (from Dender model to estuary models) and water level (from estuary models to Dender model) in a bidirectional link. Special attention is required in this use case to the transformation of the quantities. Not only do the models have different dimensions (1D to 2D), but due to the different national altitude references, a linear conversion in the water level has to be applied as well.

Waqua and Delft3D were also not yet OpenMI-compliant at the start of the project. And as for use case B and C, this use case also deals with models from different suppliers.

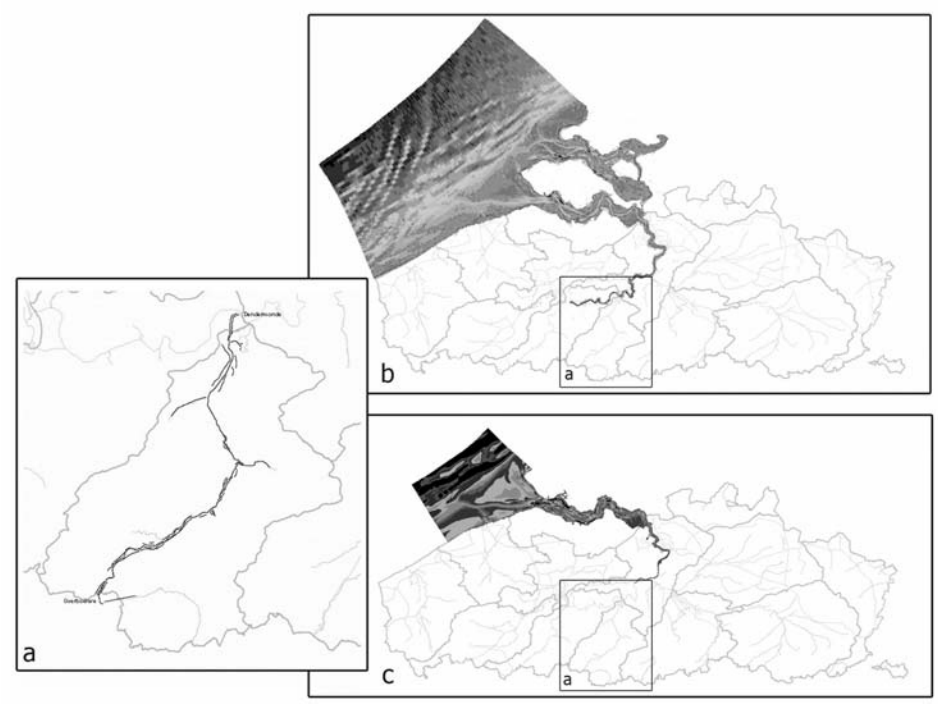


Figure 4. Geographic setting (top, b) and models used for use case D : Mike11-HD model for the Dender sub-basin (left, a), Waqua-model "Kustzuid" (top, b) and Delft3D model "Zeekennis" (right, c).

3. CURRENT PROGRESS

3.1 General

The timing of the OpenMI-Life project foresaw four major phases in the elaboration of the use cases :

- Definition phase describing the objectives, technical details and expected problems of each use case (October 2006 – March 2007)
- Trial phase, during which all aspects of the linking are being tested and problems identified and solved. This phase includes migration for those models that were not yet OpenMI-compliant (April 2007 – September 2008)
- Operational phase, during which the models will be run in linked mode on a full operational scale, i.e. performing all the types of simulations they would normally be used for in stand alone mode (October 2008 – March 2009)

- Evaluation phase. In this latest phase the use cases will evaluate the results of the operational phase (benefits compared to stand alone modelling, benefits in view of water management policy) (March 2009 – September 2009).

All use cases have produced a definition report, from which it appeared that the current version of the OpenMI standard was not a limiting factor with a view to the linking operations that were envisaged. In some cases however, there was a need for revising certain elements of the way OpenMI had been implemented in the then already compliant softwares.

3.2 Use case A

The linking of sewer and river model under normal conditions seems to work fine, even with many bidirectional links. The exchange of flows under flood conditions is currently still being tested. Other items that need resolving are the simulation of predefined series of events in linked mode, and further decisions need to be made about inconsistencies in rainfall used in both models.
None of these is expected to endanger the timing of the use case.

3.3 Use case B

Linking the upstream RS river model to the downstream Mike11 river model at a single exchange point proved to be very straightforward. In the near future it will be tested if linking at multiple exchange points is workable. Afterwards coupling at an overlapping set of exchange points will be tried out and it will be looked at how this can be applied to overlapping flooding areas.

3.4 Use case C

The main present technical achievement in this use case is the migration of the river quality model PEGASE from the UNIX environment into the Windows environment. The migration to make this model OpenMI compliant is ongoing. Several stand alone simulations have been performed with the river flow models InfoWorks RS and MIKE11 on the river Dijle in order to feed the PEGASE model after the links between the models have been defined. Appointments for defining the links have also been made.

3.5 Use case D

Waqua and Delft3D are now both OpenMI compliant. The river Dender model was coupled to the “Zeekennis” model and a correct data exchange between the two models was achieved. Due to the fact that the “Zeekennis” model was originally built for a smaller area not reaching Dendermonde, it would take too much time to extend and calibrate the “Zeekennis” model. But it was shown that the linking between the two models from a technical perspective worked correctly. The coupling between the river Dender model and “Kustzuid” is currently being tested.
The progress of the use case is in agreement with the foreseen planning.

4. FUTURE ISSUES

From the case studies, the largest concern about OpenMI is not its technical implementation, but its user-friendliness. If OpenMI based modelling is really to tear down the barriers for practical co-operation between authorities, then the use cases will

have to prove that it is possible to conceive, set up and run linked simulations with little more technical and organisational effort than what authorities are experiencing now in their normal modelling practice.
One of the key expectations that were raised by all partners involved in the use cases was the possibility of remote linking. Local linking means that all models have to run simultaneously on one machine and often requires additional licenses for software that would otherwise not be necessary. Creating links between models running at different locations and under each user's own licence, would improve the perception of complexity and practicability of integrated modelling. Although recent developments (Curn [2007]) seem promising, it cannot be guaranteed that all models will have implemented this option by the end of the operational phase.
Linked modelling will also introduce additional problems of quality assurance procedures. The risk of wrong sets of models being linked to one another is a real concern, and could lead to a deterioration of the quality of the calculations, which would be dramatic for the credibility of linked modelling.
All these issues will have to be carefully considered during the last phase of the project so as to produce a well funded evaluation of the use of OpenMI in integrated water management.

5. CONCLUSIONS

The first trials of the application of the OpenMI on real scale models in the Scheldt River Basin indicate that there is a clear potential for its use in integrated water management. So far there were no real technical obstacles that could not be handled with the current version of the OpenMI. Improved implementations for specific applications may be necessary however.
It is expected that the use cases will continue to make good progress and that operational simulations can be run at the time foreseen in the OpenMI-Life project schedule. At that time (end of 2009) it will become clear if the expected technical benefits of linked modelling and the improved way of co-operating between the different authorities will stimulate real integrated water management within Europe and beyond.

ACKNOWLEDGEMENTS

The OpenMI-Life project is made possible through the financial funding of the Life Programme of DG Environment of the European Commission (contract n° LIFE06 ENV/UK/000409). Acknowledgement also goes to the town of Leuven for making their sewer model available for the project.

REFERENCES

Curn, J., Distribution for Open Modelling Interface and Environment, Master thesis, Department of Software Engineering, Charles University Prague, Czech Republic, 2007.

Gregersen, J.B., Gijsbers, P.J.A., and Westen, S.J.P., OpenMI : Open Modelling Interface, *Journal of Hydroinformatics*, *9*(3), 175-191, 2007.

Vits, S., and Devroede, N., Interacties tussen modellen; het openMI-LIFE project, *Water*, *32*, 15-18, 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

The Migration of the UTHBAL Hydrologic Model into OpenMI

A. Loukas[a], K. Kokkinos[b], L. Vasiliades[a], and A. Liakopoulos[a]

[a] *Department of Civil Engineering, University of Thessaly, 38334 Volos, Greece (aloukas/lvassil/aliakop@civ.uth.gr)*

[b] *Department of Information Technology and Telecommunications, TEI, Larissa, Greece (k_kokkinos@teilar.gr)*

Abstract: This paper deals with the implementation and the migration into OpenMI of a monthly conceptual hydrological model, called UTHBAL. The model has been developed by Loukas et al. [2003, 2007] and now is implemented under the Microsoft Visual Studio .NET C# IDE. The integrated decision support framework provided consists of an autonomous model engine, a graph editor, an optimization/calibration server and a migration wrapper. The software allows input data polymorphism and extensibility of model integration, which is lumped in the present form. The infrastructure for incorporating various optimization libraries has also been built. On line data exchange is feasible since the model has been migrated into OpenMI through a dynamically configurable wrapper. We found that the model characteristics relevant to coupling with other hydrological models lie on the space, time, structure and optimization compatibilities. Thus other models which analyze similar phenomena can seamlessly be coupled with ours.

Keywords: Hydrological Modelling; Migration; Rainfall-Runoff Models; OpenMI.

1. INTRODUCTION

Hydrological modelling and decision support frameworks are intended to be used by water resource managers and modelers in water resources management. Unfortunately, the existing frameworks only attempt to combine their sub-module codes into a single program, requiring the translation of different source codes with different complier versions. Such a process has been standard practice within the hydrologic community, however, it is costly and time consuming. The idea of external interaction of hydrological software has recently been established through the development of the Open Modeling Interface (OpenMI) - standard [Gregersen et al., 2007; Moore and Tindall, 2005]. The use of the OpenMI-standard not only address the coupling of individual catchment processes but also guarantees minimal programming effort for the spatial and temporal synchronization and data interaction between compatible models. The above methodology, however, can be applied only by implementing the necessary infrastructure in order to achieve compliance with OpenMI. This means that every model must be accompanied by a migration wrapper in order to interact with external models through the OpenMI standard [Gregersen et al., 2005].

This study describes the migration process of the monthly conceptual hydrological model UTHBAL for water balance modeling [Loukas et al., 2003; Loukas et al., 2007]. Lumped water balance models have been developed at various time scales (e.g. hourly, daily, monthly and yearly) and to varying degrees of complexity. Monthly water balance models were first developed in the 1940s by Thornthwaite [1948] and have been adopted, modified, and applied to a wide spectrum of hydrological problems. Recently, these hydrological models have been employed to explore the impact of climatic change. They

also have been utilized for long-range streamflow forecasting. Although such applications may use hourly or daily models, these models are, however, more data intensive and have more parameters than the monthly models. A complete review of water balance model applications could be found in Xu and Singh [1998].

Before the migration process, the model was implemented under the Microsoft Visual Studio .NET C# IDE. The framework consists of four basic components: (a) the data incorporation component, (b) the optimization-verification component, (c) the graphical component and (d) the migration component. The above modules have been implemented to be portable and externally executable as independent dynamic link libraries (dlls). For the creation of the appropriate OpenMI wrapper engine, a semiautomatic approach is used: the user is asked to identify the preferable hydrologic components to be exchanged in the interaction process by formulating various .omi files. The migration process is semi-dynamic, thus helping the modeler to provide various OpenMI migration wrappers according to the coupling characteristics and dimensionalities with other models. This migration process has been verified under the NUnit test framework for all created .omi files that can be produced with all possible combinations of interchanging data sets.

The migration process correctness and the coupling ability of our application has been verified for a specific case study involving the coupling of the UTHBAL framework with a groundwater model for simulating surface runoff and groundwater recharge. In the following sections we provide: (a) a description of the UTHBAL model, (b) a software engineering point-of-view explanation of the developed framework summarizing its functional requirements, (c) a rigorous algorithm for the migration wrapper development according to OpenMI standards and (d) a thorough proof of concept of the data interchange process for the above referred case study.

2. THE UTHBAL MODEL

The model has been developed by Loukas et al. [2003] and updated in its present form for simulation of hydrologic cycle components [Loukas et al., 2007]. The UTHBAL model has been successfully applied to watersheds in Cyprus, in the regions of Crete and Thessaly of Greece and in the Greek-Bulgarian transboundary Nestos/Mesta River basin.

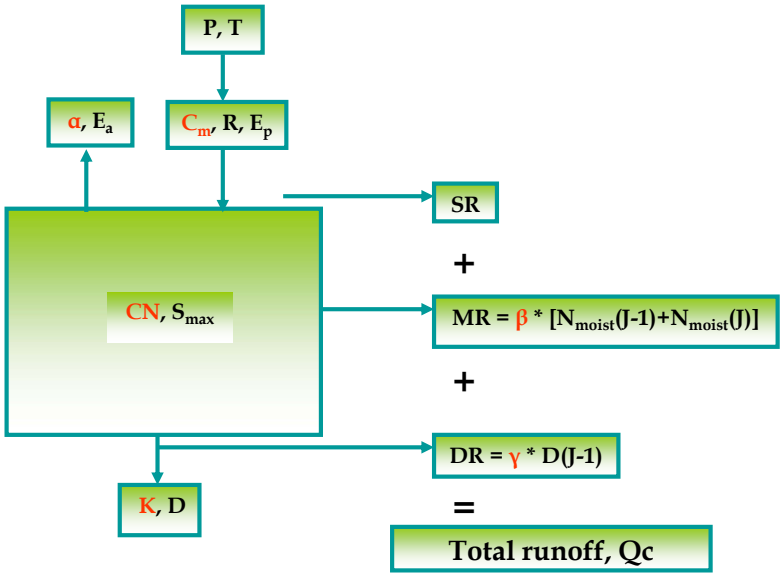


Figure 1. UTHBAL flowchart and model parameters.

The UTHBAL model requires monthly values of mean temperature, precipitation, and potential evapotranspiration and produces values for actual evapotranspiration, soil moisture, groundwater and surface runoff. The model separates the total precipitation into rainfall and snowfall, because the correct division of precipitation is essential for modelling

the mass balance of seasonal snow covers and for accurate runoff simulation [WMO, 1986]. The rain-snow percentage is estimated using a logistic relationship based on mean monthly temperature [Knight et al., 2001]:

$$\begin{aligned} \%S &= 0 && \text{if} \quad T \geq 12.22\,^{0}C \\ \%S &= \frac{1}{(1.35^{T} * 1.61) + 1} && \text{if} \quad -10\,^{0}C \leq T \leq 12.22\,^{0}C \\ \%S &= 1 && \text{if} \quad T \leq -10\,^{0}C \end{aligned} \tag{1}$$

where, *%S* is the monthly percentage of precipitation which is falling in the form of snow and *T* the mean monthly temperature. The snowmelt of month *J*, *SM(J)* is estimated using the simple degree-day method [Semadeni-Davies, 1997]:

$$SM(J) = C_m \cdot T(J) \tag{2}$$

where, *T* is the mean monthly temperature and C_m is the monthly melt rate factor (mm/°C per month).

The snow water equivalent of the accumulated snowpack, SWE_{sp} is estimated by:

$$SWE_{sp}(J) = SWE_{sp}(J-1) + S(J) - SM(J) \tag{3}$$

where, *S(J)* is the snow fallen during month *J* equals to:

$$S(J) = \%S \cdot P(J) \tag{4}$$

where, *P(J)* is the total precipitation of month *J*.

The model divides the total watershed runoff into three components: the surface runoff, the interflow, and the baseflow using a soil moisture mechanism (Figure 1). The first priority of the model is to fulfill the actual evapotranspiration. The monthly actual evapotranspiration depends on the available soil moisture and the average surface potential evapotranspiration for that month. The model estimates and gives as output, apart from the actual evapotranspiration, the available soil moisture, the deep recharge to groundwater, and the surface runoff. The detailed algorithm of UTHBAL could be found in a recent paper [Loukas et al., 2007]

3. THE UTHBAL IN THE OPENMI FRAMEWORK

The UTHBAL model was implemented under the Microsoft Visual Studio .NET C# IDE (Figure 2), in an integrated environment that accepts time series data based on different geographical locations. Currently the manipulation is implemented in lumped mode. The extension into the distributed mode is an ongoing process. This will incorporate interaction and data interchange with GIS applications running these applications inside the UTHBAL framework. The framework is constructed in four basic modules each of those integrating respectively: (a) the XML-format storage of the applications, (b) the calibration and optimization of the model, (c) the visualization of the output data and (d) the automated migration and export into OpenMI. The following subsections elaborate on the above components. An architectural schematic of UTHBAL within the OpenMI framework is depicted in Figure 3.

3.1 Storage Component

Case studies based on the UTHBAL can be stored, loaded, and run separately by the framework. Each case study is stored in a .ubm file, an XML-type format file that includes: (1) a standard geographic location for the lumped mode or the XYZ-polygon coordinates

for the distributed, (2) temporal discretization of the case study (day, month, etc.) and (3) relative paths to the appropriate dlls. The `IEngine` object is then programmed to read from that XML-file and is prepared to create the set of input/output data for exchange. We also wrap in a .dll the object that does all the intermediate calculations for the simulation of the time series. That makes the running of any UTHBAL - OpenMI simulations fully portable and independent of the physical presence of the UTHBAL application. In the future, we plan to create a web service so users can run their case scenarios using the UTHBAL remotely.

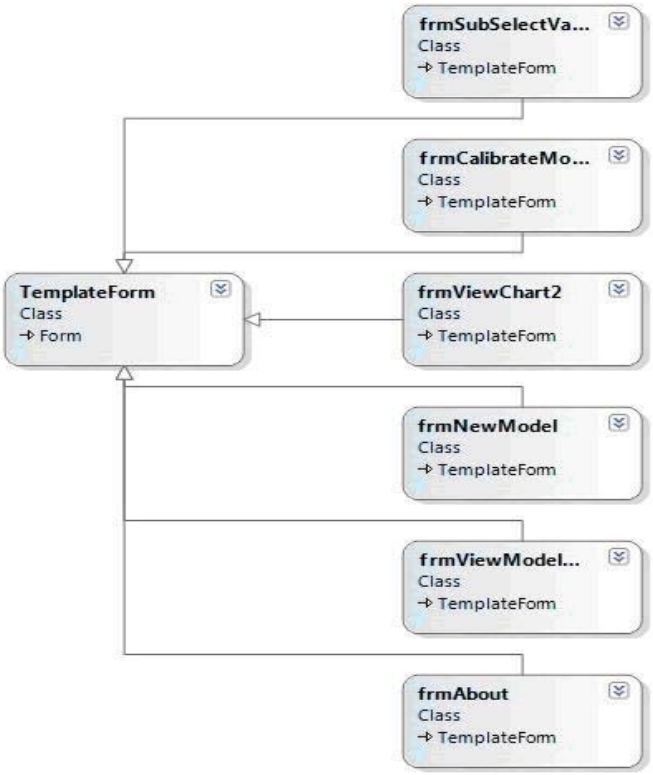


Figure 2. A class diagram of the UTHBAL model.

3.2 Calibration, Optimization and Verification

The UTHBAL model has six parameters to be optimized in order to estimate watershed runoff and hydrological cycle components, namely, Cm, CN, K, α, β, and γ. The framework allows the user to calibrate the model by dynamically selecting any preferable time period for calibration. Currently only the Generalized Reduced Gradient method [Lasdon and Smith, 1992] is used for optimizing model parameters and Model Efficiency (*Eff*) [Nash and Sutcliffe, 1970] was used as the objective function between the observed and simulated runoff. Apart of the *Eff*, two more statistical measures of the quality of runoff simulation are used: a) the percentage runoff volume difference (*DV%*), and b) the coefficient of determination (R^2) between the observed and simulated runoff. The inclusion of several other optimization methods to the framework is an ongoing process also. Our intention is to extend the current module into an optimization server, thus making the framework suitable for multi-threaded applications, serving multiple users concurrently.

3.3 Output Data Visualization

The UTHBAL framework applies a unique visualization method of input and output data implemented using C# Crystal Reports and the ZedGraph library [ZedGraph, 2008]. Data

can be visualized in histograms by having the user dynamically select which quantities to depict (Figure 4). The on line manipulation such as zooming and the storage of the graphs in various formats (jpg, pdf etc.) is also available (Figure 5). The stored graphs are tightly coupled to the individual case studies by inserting their storage path into the .ubm XML-structure for portability.

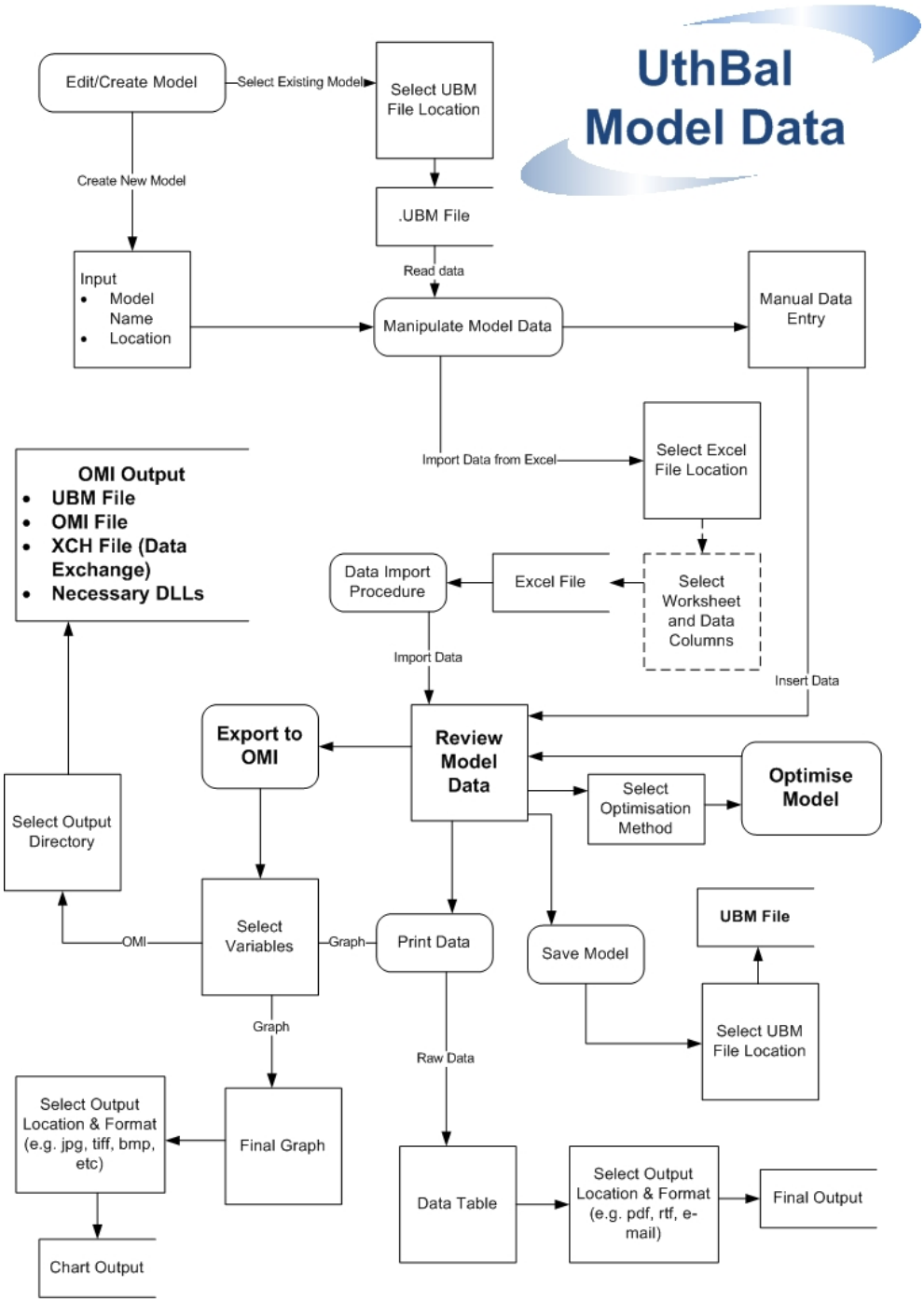


Figure 3. The architecture of UTHBAL.

3.4 Model Migration to OpenMI

The provision of data exchange of the UTHBAL framework with other hydrologic models is based on the OpenMI paradigm. However, special effort has been made to extend this

practice by establishing a metadata connection of the model to the OpenMI editor through the use of .ubm files. An .ubm file is simply an XML file carrying out the model inputs and outputs along with the time series specifications. The location of this file for each simulation run is passed to the `IEngine` component of OpenMI as a parameter to the OMI file. Since any OMI file would need access to the UTHBAL classes for the nodes available, this effectively means that, for our case these class libraries are also ported. The result is that, any user is able to export a UTHBAL case study to an OMI model in one workstation and run the produced OpenMI simulation in another. The migration process accomplishes that by physically copying the necessary libraries from the UTHBAL directory to the target OpenMI directory.

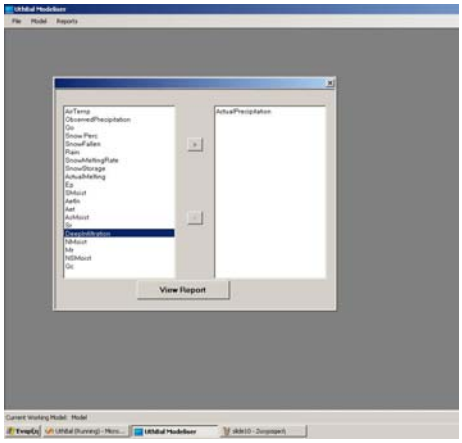


Figure 4. Dynamic selection for depicting intermediate output data.

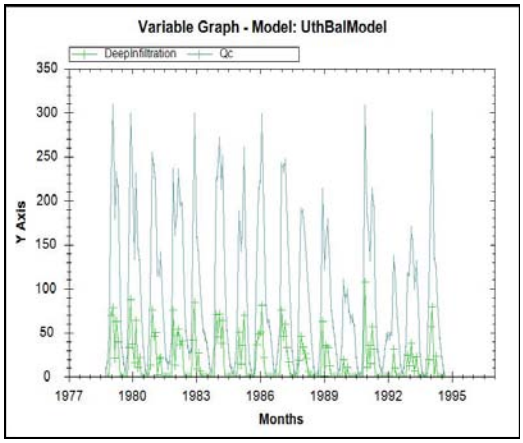


Figure 5. Resulting graphs.

Another user friendly feature of the framework is the automatic built up of the OMI file for each case study. This is accomplished by scripting the standard XML tag information including the locations of the `IEngine`, the `LinkableEngine` and the `LinkableComponent` of the application.

We added to the functional requirements of the framework the ability of the user to select which model variables should be available through the OpenMI data exchange. Since UTHBAL model can be migrated to OpenMI in order to exchange data with other models, it is practical to always show all intermediate output data variables. This is solved by saving the variable names in a binary file in the same directory as the OMI file. The file needs to be present so that the `UthBalModelOMIEngine` knows the variables to show the user. The `LinkableEngine` (called `UthBalModelOMIInterface`) instantiates a new `UthBalModelOMIEngine` which is simply an implementation of the OpenMI's `IEngine` interface. The `UthBalModelOMIEngine` class looks for a path of the .ubm file so that it can instantiate a new UTHBAL model. If none is found, the UTHBAL model class automatically looks for .ubm files in its own directory. We cannot have a `UthBalModelOMIEngine` object without the corresponding .ubm file, since the simulation cannot run without data. The `UthBalModelOMIEngine` class handles all communication from the UTHBAL model to the OpenMI through the procedures of the `IEngine` interface. More specifically, the time horizon of the model is defined as an `ITimeSpan` variable (OpenMI's time period representation) which spans from the first day of the month of the first node to the first day of the month of the last node of the model. The `outputExchangeItems` array is filled according to the previously mentioned binary file, while the model description (given by using the `GetModelDescription` method) is set to show the model's name. Also the `GetValues` method is set to use a method in the UTHBAL framework that reads a variable from a given node according to the variable name. This method was placed in the node class specifically for facilitating the migration.

4. A TESTING CASE OF UTHBAL MIGRATION

The migration process correctness and the coupling ability of our application has been verified for a specific case study involving the coupling of the UTHBAL framework with the groundwater regional model produced by the OpenMI Association. Changes in the argument section of the original RegionalGWModel.omi file included only the setting of simulation starting and ending days. We considered the general case scenario where UTHBAL possibly interchanges all calculated variables with compatible models. Quantity passing succeeded for the entire temporal spread as expected. The promotion of UTHBAL to the distributed phase relatively to its OpenMI linkage is also pre-tested with the inclusion of an array of ILinkableComponents passing only the single value of deep infiltration or groundwater recharge. Figure 6 illustrates the NUnit testing and Figure 7 the coupling testing process.

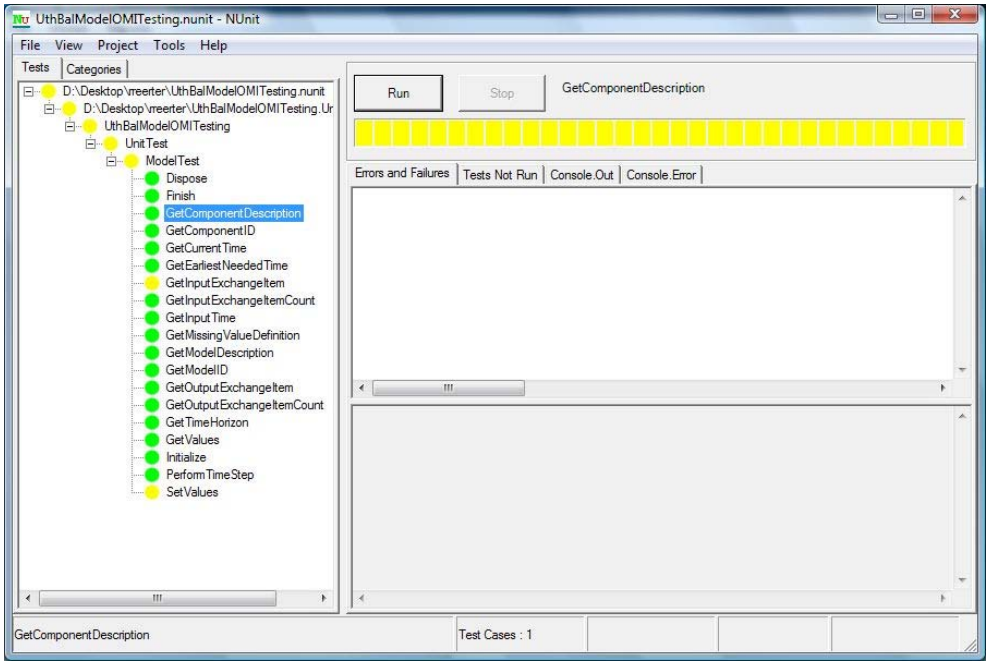


Figure 6. NUnit testing of UTHBAL wrapper.

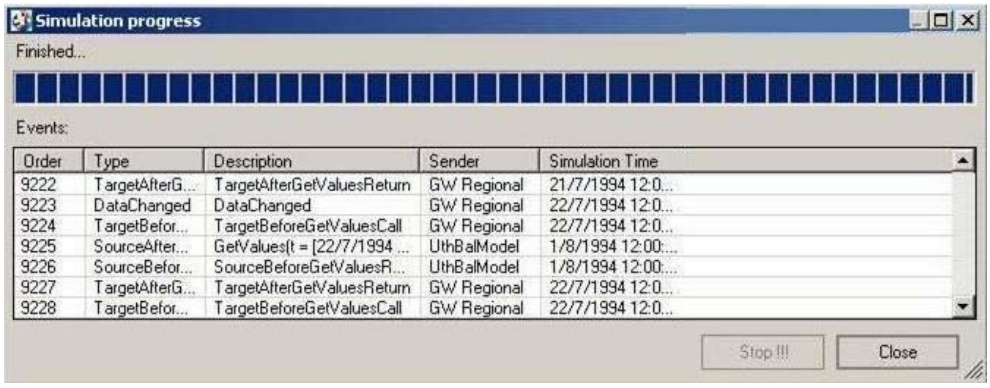


Figure 7. Simulation run of the coupling test.

5. CONCLUSIONS

In this paper, the migration process of a monthly conceptual hydrological model in the OpenMI framework has been presented. The hydrological model is the UTHBAL model

developed in the University of Thessaly. Before the migration process, the model was implemented under the Microsoft Visual Studio .NET C# IDE. The framework consists of four basic components: (a) the data incorporation component, (b) the graphical component, (c) the optimization-verification component and (d) the migration component. More specifically, (a) the data incorporation component involves the input of data in text or Microsoft Excel© format, as well as the storage process of case studies in separate XML format type files, (b) the graphical component depicts the calculated time series graphs of the components of hydrologic cycle (i.e. rainfall, snowfall, snowmelt, actual evapotranspiration, soil moisture, surface and groundwater runoff), (c) the optimization component involves the techniques of the model parameter optimization based on Generalized Reduced Gradient method and finally (d) the migration component that implements the model engine wrapping. The above modules have been implemented to be portable and externally executable as independent dynamic link libraries (dlls). For the creation of the appropriate OpenMI wrapper engine, a semiautomatic approach is used: the user is asked to identify the preferable hydrologic components to be exchanged in the interaction process by formulating various .omi files. The migration process has been verified under the NUnit Testing framework for all created .omi files for a specific case study and for the communication of UTHBAL with a groundwater model.

ACKNOWLEDGEMENTS

This research is partially funded by EC under the LIFE Programme and contributing to the implementation of the thematic component LIFE-Environment under the policy area "Sustainable management of ground water and surface water management". Contract no: LIFE06 ENV/UK/000409.

REFERENCES

Gregersen, J.B., P.J.A. Gijsbers, and S.J.P. Westen, OpenMI: Open Modelling Interface, *Journal of Hydroinformatics*, 9(3), 175-191, 2007.

Gregersen, J.B., P.J.A. Gijsbers, S.J.P. Westen, and M. Blind, OpenMI: the essential concepts and their implications for legacy software, *Advances in Geosciences*, 4, 37-44, 2005.

Knight, C.G., H. Chang, P.M. Staneva, and D. Kostov, A simplified basin model for simulating runoff: the Struma River GIS, *Professional Geographer*, 53(4), 533-545, 2001.

Lasdon, L.S., and S. Smith, Solving sparse nonlinear programs using GRG, *ORSA Journal on Computing*, 4(1), 2-15, 1992.

Loukas, A., L. Vasiliades, and N. Mpastrogiannis, Hydrologic simulation of surface water balance in Yermasoyia watershed, Cyprus, *In Sidiropoulos E., Iacovides I., (eds) Research of water resources in Cyprus*, Nicosia, Cyprus, 83-114, 2003 (in Greek).

Loukas, A., N. Mylopoulos, and L. Vasiliades, A modeling system for the evaluation of water resources management strategies in Thessaly, Greece, *Water Resources Management*, 21(10), 1673-1702, 2007.

Moore, R., and I. Tindall, An overview of the open modelling interface and environment (the OpenMI), *Environmental Science & Policy*, 8, 279-286, 2005.

Nash, J.E., and J.V. Sutcliffe, River flow forecasting through conceptual models. Part-1: A discussion of principles, *Journal of Hydrology*, 10(3), 282-290, 1970.

Semadeni-Davies, A., Monthly snowmelt modelling for large-scale climate change studies using the degree day approach, *Ecological Modelling*, 101, 303-323, 1997.

Thornthwaite, C.W., An approach toward a radiation classification of climate, *Geogr. Rev.* 38(1), 55-94, 1948.

World Meteorological Organization, (WMO), Intercomparison of models of snowmelt runoff, *Operational Hydrology No. 23 (WMO No. 646)*, Geneva, 1986.

Xu, C.-Y., and V.P. Singh, A review on monthly water balance models for water resources investigations, *Water Resources Management*, 12(1), 20-50, 1998.

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 Application of Data Mining to PM_{10} Level Medium-Term Prediction

E. Pasero[a], A. Montuori[b], W. Moniaci[a] and G. Raimondo[a]

[a]Electronics Department and [b]Department of Architectural and Industrial Design, Polytechnic of Turin,. C.so Duca degli Abruzzi, 24 10129 Torino, Italy
(eros.pasero/alfonso.montuori/walter.moniaci/giovanni.raimondo@polito.it)
INFN, Sez Torino, I-10125 Turin, Italy

Abstract: The study described in this paper, analyzed the urban air pollution principal causes and identified the best subset of features (meteorological data and air pollutants concentrations) for each air pollutant in order to predict its medium-term concentration (in particular for the PM_{10}). An information theoretic approach to feature selection has been applied in order to determine the best subset of features by means of a proper backward selection algorithm. The final aim of the research is the implementation of a prognostic tool able to reduce the risk for the air pollutants concentrations to be above the alarm thresholds fixed by the law. The implementation of this tool will be carried out using machine learning methods based on some of the most wide-spread statistical data driven techniques (Artificial Neural Networks, ANN, and Support Vector Machines, SVM).

Keywords: Machine learning methods; feature selection; air-pollution time series analysis and prediction.

1. INTRODUCTION

One of the biggest problems of urban areas is air pollution. Air pollution arises from the adverse effects on the environment of a variety of contaminants emitted into the atmosphere by natural and man-made processes such as industrial emissions, fixed combustions and vehicular traffic. The respect of the European laws concerning urban and suburban air pollution requires the analysis and implementation of automatic operating procedures in order to prevent the risk for the principal air pollutants to be above the alarm thresholds. The aim of the analysis is the medium-term forecasting of the air pollutants mean and maximum values by means of meteorological actual and forecasted data. Critical air pollution events frequently occur where the geographical and meteorological conditions do not permit an easy circulation of air and a large part of the population moves frequently between distant places of a city.

The analysis described in this paper was performed on the hourly data of the principal air pollutants (Sulphur Dioxide SO_2, Nitrogen Dioxide NO_2, Nitrogen Oxides NO_x, Carbon Monoxide CO, Ozone O_3 and Particulate Matter PM_{10}) and meteorological parameters (air temperature, relative humidity, wind velocity and direction, atmospheric pressure and solar radiation) measured by a station located in the urban area of the city of Turin (Italy). All the measurements data are relative to the time period 01/06÷10/07, Agenzia Regionale per la Protezione Ambientale del Piemonte (the Piedmont Regional Environmental Protection Agency), ARPA Piemonte, Aria Web.

The Turin urban area is located at the western side of the Po Valley, the most industrialized and populated district in Italy. It is characterized by complex terrains with mountains rising over 2700 m on the north west and a range of hills reaching 700 m located just east of the

city. Despite its position at mountain feet, the city of Turin is frequently affected by low winds, making its climatology and exposure to severe pollution episodes similar to other Po Valley sites. Apart from terrain features, the large urban area with its micrometeorological footprint gives rise to complex local atmospheric flow patterns that, particularly in the lowest layers, show relevant differences from large scale circulation (Calori et al. [2006]). From the beginning of the 70's policies for the reduction of chemical agents release in the air have been adopted. According to the Piedmont Regional Environmental Protection Agency such policies (as for example improved formulations of fossil oils for industrial activities, large adoption of methane for residential heatings, and the renewal of the fleet of circulating vehicles) gave good results allowing significant decreases in SO_2, lead (Pb) and CO emissions. Nowadays air pollution problems arise from NO_2, PM_{10}, and O_3 the critical impact on public health of which was recognized only recently. The most severe health issue is now constituted by the high levels of PM_{10}, a category of pollutants including solid and liquid particles having an aerodynamic diameter of up to 10 μm. In 2000 the yearly average of PM_{10} in Turin was equal to 71 $\mu g/m^3$ and the number of exceedances of the daily limit value for the protection of human health (fixed by the law at 50 $\mu g/m^3$ according to EU Directive 99/30/CE) was equal to 214 (while it should be less than 35 per civil year). PM_{10} can be a health hazard for several reasons: it can harm lung tissues and throat, aggravate asthma and increase respiratory illness. Indeed, high PM_{10} levels have been correlated to increase in hospital admissions for lung and heart disease (Ostro et al. [1999]). These events require drastic measures such as the closing of the schools and factories and the restriction of vehicular traffic. The forecasting of such phenomena with up to two days in advance would allow to take more efficient countermeasures to safeguard citizens health.

In all the cases in which we can assume that the air pollutants emission and dispersion processes are stationary, it is possible to solve this problem by means of statistical learning algorithms that do not require the use of an explicit prediction model. The definition of a prognostic dispersion model is necessary when the stationarity conditions are not verified. It may happen for example when it is needed to forecast the air pollutant concentration variation due to a large variation of the emission of a source or to the presence of a new source, or when it is needed to evaluate a prediction in an area where there are no measurement points.

The research activity described in this paper concerns the feasibility of applying machine learning methods to forecast air pollution. The analysis carries on the work already developed by the NeMeFo (Neural Meteo Forecasting) research project for meteorological data short-term forecasting, Pasero et al. [2004]. The Artificial Neural Networks (ANN) and the Support Vector Machines (SVM) have been often used as a prognostic tool for air pollution, Benvenuto and Marani [2000], Perez et al. [2000], Canu and Rakotomamonjy [2001], Božnar et al. [2004], Cecchetti et al. [2004], Slini et al. [2004]. Even if we refer to these approaches as black-box methods, in as much as they are not based on an explicit model, they have generalization capabilities that make possibile their application to non-stationary situations.

The first step for the implementation of a prognostic neural network or SVM is the selection of the best subset of features that are going to be used as the input to the forecasting tool. The potential benefits of the features selection process are many: facilitating data visualization and understanding, reducing the measurement and storage requirements, reducing training and utilization times, defying the curse of dimensionality to improve prediction or classification performance. It is important to highlight that the selection of the best subset of features useful for the design of a good predictor is not equivalent to the problem of ranking all the potentially relevant features. In fact the problem of features ranking, on the basis of the correlation between each of them and the target, is sub-optimum with respect to features selection especially if some features are redundant or unnecessary. On the contrary a subset of variables useful for the prediction can count out a certain number of relevant features because they are redundant, Guyon and Elisseeff [2003]. Depending on the way the searching phase is combined with the prediction, there are three main classes of feature selection algorithms: filters, wrappers and embedded. A filter is defined as a feature selection algorithm using a performance metric based entirely on the training data, without reference to the prediction algorithm for which

the features are to be selected. Wrapper algorithms include the prediction algorithm in the performance metric. The name is derived from the notion that the feature selection algorithm is inextricable from the end prediction system, and is wrapped around it. Finally, embedded methods, perform the selection of the features during the training procedure and they are specific of the particular learning algorithm.

In this work the method used for features selection is a filter. More precisely a selection algorithm with backward eliminations was used. The criterion used to eliminate the features is based on the notion of relative entropy (also known as the Kullback-Leibler divergence), inferred by the information theory.

2. FEATURE SELECTION ALGORITHM

The first step of the analysis was the selection of the most useful features for the prediction of each of the targets relative to the air-pollutants concentrations. For each air pollutant the target was chosen to be the mean value over 24 hours, measured every 4 hours (corresponding to 6 daily intervals a day). The complete set of features on which was made the selection, for each of the available parameters (air pollutants, air temperature, relative humidity, atmospheric pressure, solar radiation, wind speed and direction), consisted of the maximum and minimum values and the daily averages of the previous three days to which the measurement hour and the reference to the week day were added. Thus the initial set of features, for each air-pollutant, included 112 features. From this analysis an apposite set of data was excluded; such set was used as the test set.
The algorithm developed by Koller and Sahami [1996] was used to select an optimal subset of features from the set of features described above. In the following the formalism of the authors to describe the theoretical framework of the algorithm will be used. Let $\boldsymbol{F}=(F_1, F_2,..., F_N)$ be the set of structural features and let $\boldsymbol{Q}=(Q_1, Q_2,..., Q_M)$ be the set of the chosen target. For each assignment of values $\boldsymbol{f} =(f_1, f_2,..., f_N)$ to $\boldsymbol{F}$ we have a probability distribution $P(\boldsymbol{Q} \mid \boldsymbol{F} = \boldsymbol{f})$ on the different possible classes, $\boldsymbol{Q}$. We want to select an optimal subset $\boldsymbol{G}$ of $\boldsymbol{F}$ which fully determines the appropriate classification. We can use a probability distribution to model the classification function. More precisely, for each assignment of values $\boldsymbol{g}=(g_1, g_2,..., g_P)$ to $\boldsymbol{G}$ we have a probability distribution $P(\boldsymbol{Q} \mid \boldsymbol{G} = \boldsymbol{g})$ on the different possible classes, $\boldsymbol{Q}$. Given an instance $\boldsymbol{f}=(f_1, f_2,..., f_N)$ of $\boldsymbol{F}$, let $\boldsymbol{f_G}$ be the projection of $\boldsymbol{f}$ onto the variables in $\boldsymbol{G}$. The goal of the Koller-Sahami algorithm is to select $\boldsymbol{G}$ so that the probability distribution $P(\boldsymbol{Q} \mid \boldsymbol{F} = \boldsymbol{f})$ is as close as possible to the probability distribution $P(\boldsymbol{Q} \mid \boldsymbol{G} = \boldsymbol{f_G})$. To select $\boldsymbol{G}$ the algorithm uses a backward elimination procedure, where at each step the feature F_i which has the best Markov blanket approximation $\boldsymbol{M_i}$ is eliminated, Pearl [1988]. Formally, we say that a subset $\boldsymbol{M_i}$ of $\boldsymbol{F}$ which does not contain F_i is a Markov blanket for F_i if F_i is conditionally independent of $\boldsymbol{F} - \boldsymbol{M_i} - \{F_i\}$ given $\boldsymbol{M_i}$. If $\boldsymbol{M_i}$ is a Markov blanket of F_i then it is also the case that the classes in $\boldsymbol{Q}$ are conditionally independent of the feature F_i given $\boldsymbol{M_i}$. The mean value of the relative entropy between the distributions $P(\boldsymbol{Q} \mid \boldsymbol{M_i}=\boldsymbol{f_{Mi}}, F_i= f_i)$ and $P(\boldsymbol{Q} \mid \boldsymbol{M_i}=\boldsymbol{f_{Mi}})$ is used to understand how close $\boldsymbol{M_i}$ is to being a Markov blanket for F_i:

$$\delta_G(\boldsymbol{F_i} \mid \boldsymbol{M_i}) = \sum_{f_{M_i}, f_i} P(\boldsymbol{M_i} = f_{M_i}, \boldsymbol{F_i} = f_i) \cdot \sum_{Q_i \in Q} P(\boldsymbol{Q_i} \mid \boldsymbol{M_i} = f_{M_i}, \boldsymbol{F_i} = f_i) \cdot \log \frac{P(\boldsymbol{Q_i} \mid \boldsymbol{M_i} = f_{M_i}, \boldsymbol{F_i} = f_i)}{P(\boldsymbol{Q_i} \mid \boldsymbol{M_i} = f_{M_i})} \quad (1)$$

At each step the feature F_i for which $\delta_G(\boldsymbol{F_i} \mid \boldsymbol{M_i})$ is minimum is eliminated. The computational complexity of this algorithm is exponential only in the size of the Markov blanket, which is small. For the above reason we could quickly estimate the probability distributions $P(\boldsymbol{Q} \mid \boldsymbol{M_i}=\boldsymbol{f_{Mi}}, F_i= f_i)$ and $P(\boldsymbol{Q} \mid \boldsymbol{M_i}=\boldsymbol{f_{Mi}})$ for each assignment of values $\boldsymbol{f_{Mi}}$ and f_i to $\boldsymbol{M_i}$ and F_i. The estimate of the probability density was made by using the Parzen method as described in Parzen [1962] and Costa et al. [2003]. We modified the Koller-

Sahami algorithm in the way a candidate Markov blanket $\boldsymbol{M_i}$ for the feature F_i is selected. Instead of selecting a candidate Markov blanket $\boldsymbol{M_i}$ of size k for the features F_i by using the set of the k features most correlated to F_i we selected the k features F_j which minimize the mean value of relative entropy between the distributions $P(\boldsymbol{Q} \mid F_i = f_i, F_j = f_j)$ and $P(\boldsymbol{Q} \mid F_j = f_j)$. After a trial and error procedure the Markov Blanket size was set equal to 3 considering the trade-off between the best approximation of a Markov blanket (large sizes) and minimum fragmentation of the training set (small sizes), Koller and Sahami [1996].
In particular this method was applied to the selection of the best subset of features useful for the prediction of the average daily concentration of PM_{10} in the city of Turin that was often above the limit value (threshold) for the safeguard of human health (50 $\mu g/m^3$). The best subset of 16 features turned out to be the following:

- Average concentration of PM_{10} in the previous day.
- Maximum hourly value of the ozone concentration one, two and three days in advance.
- Average concentration of the ozone one, two and three days in advance.
- Minimum hourly value of the air temperature one, two and three days in advance.
- Maximum hourly value of the air temperature one day in advance.
- Average value of the solar radiation one day in advance.
- Average concentration of NO_2 one day in advance.
- Minimum hourly value of the humidity one and two days in advance.
- Minimum hourly value of the wind direction one day in advance.

The results can be explained considering that PM_{10} is partly primary, directly emitted in the atmosphere, and partly secondary, that is produced by chemical/physical transformations that involve different substances as SO_2, NO_x, VOCs, NH_3 at specific meteoriological conditions (Quaderno Tecnico Arpa [2002]). The features most correlated with the target and still present in the best subset of 8 features are the following ones (reported with the corresponding correlations): Average concentration of PM_{10} in the previous day (0.77), Average concentration of NO_2 one day in advance (0.49), Average value of the solar radiation one day in advance (0.35) and Maximum hourly value of the air temperature one day in advance (0.34).

3. FORECASTING WHEN THE CONCENTRATIONS EXCEED THE LIMIT VALUE FOR THE PROTECTION OF HUMAN HEALTH

A set of feed-forward neural networks with the same topology was used. Each network had three layers with 1 neuron in the output layer and a certain number of neurons in the hidden layer (varying in a range between 3 and 20). The hyperbolic tangent function was used as transfer function.

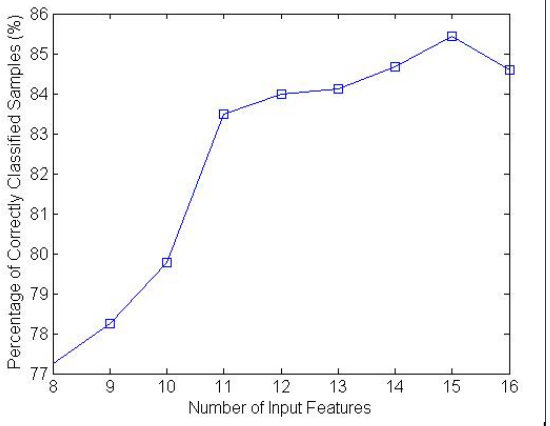


Figure 1. ANN Performance as a function of Input Features (samples above the threshold).

The back-propagation rule, Werbos [1974], was used to adjust the weights of each network and the Levenberg-Marquardt algorithm, Marquardt [1963], to proceed smoothly between the extremes of the inverse-Hessian method and the steepest descent method. The Matlab Neural Network Toolbox, Demuth and Beale [2005], was used to implement the neural networks'set.

An SVM with an ε-insensitive loss function, Vapnik [1995], was also used. The Gaussian function was used as kernel function of the SVM. The principal parameters of the SVM

were the regularized constant C determining the trade-off between the training error and model flatness, the width value σ of the Gaussian kernel, and the width ε of the tube around the solution. The SVM performance was optimized choosing the proper values for such parameters. An active set method, Fletcher [1987], was used as optimization algorithm for the training of the SVM. The SVM was implemented using the "SVM and Kernel Methods Matlab Toolbox", Canu et al. [2005].

Furthermore in order to evaluate the global performance of the machine learning methods we calculated the Index of Agreement (*IA*), following the suggestions of Wilmott et al. [1985]:

$$IA = 1 - \frac{\sum_i |f_i - s_i|^2}{\sum_i \left(|f_i - \bar{s}| + |s_i - \bar{s}| \right)^2} \quad \textbf{(2)}$$

The value of the IA is from 0 to 1 and is a measure of the agreement between forecasts (f_i) and observations samples (s_i). The best performance corresponds to the value 1.

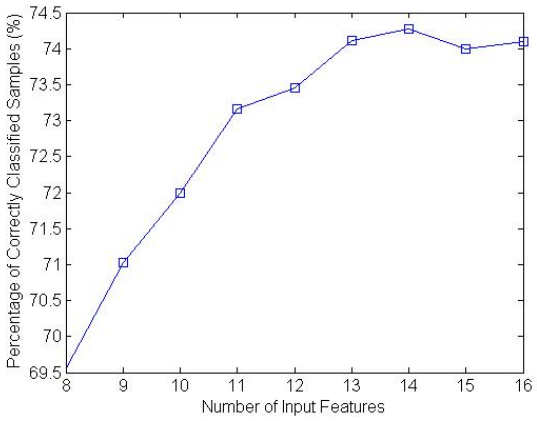


Figure 2. ANN Performance as a function of Input Features (samples below the threshold).

The neural networks were trained on a subset of the data used for the features selection. The training set was chosen to be adaptive with a constant-width sliding window. The training window width was set equal to half the number of the available data (row1). The window was slided in such a way that to forecast the (row1+1)st sample we trained the ANN on the previous row1 samples (from the 1st to the row1st), to forecast the (row1+2)nd sample we used the previous row1 samples (from the 2nd to the (row1+1)st) and so on. In this way we obtained a subset of data temporally close enough to the sample to be predicted, in order to enhance the forecasting accuracy. The test set consisted of one out of five data not used for the features selection algorithm in order to speed up computation times. Since the numbers of the training samples above and below the maximum threshold for the PM_{10} concentration were different, the training of the networks was performed weighting more the kind of samples present a fewer number of times.

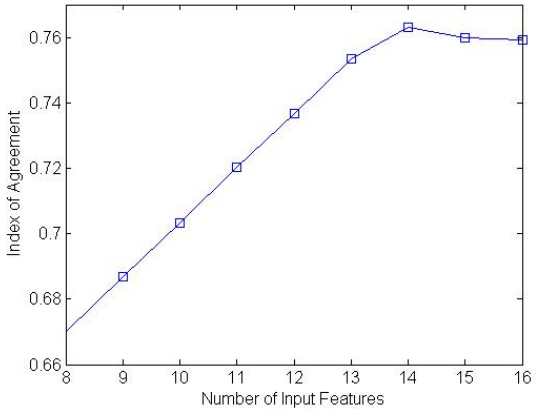


Figure 3. ANN Performance: Index of Agreement

The ANN performance, both for the samples under the threshold and for the samples above the threshold, reaches a maximum at a given number of input features (optimal value) and then tends to flatten when the number of input features increase. In Figure 1 and Figure 2

are shown the performances of the ANN with 18 hidden neurons, which gave the best forecasting results, for the samples below and above the threshold. This behaviour shows the effectiveness of the features selection algorithm; in fact the performance flattens adding features above an optimal number because the non-linearities involved in the generation and dispersion of the PM_{10} are already tracked by the optimal number of features and adding extra features may not provide further relevant information for the forecasting. More precisely the performance increased meaningfully from 8 to 14/15 input features and tended to flatten when the size of the input vector was greater than 14. The results obtained with 55 samples of days under the threshold and 102 samples of days above the threshold are the following: *FA*, defined as the ratio between the number of false alarms and the number of totally predicted exceedances, is low (14%) while the capability to forecast when the concentrations are above the threshold is about 85%. These results are comparable to those, relative to similarly polluted areas, that can be found in the literature as in Cecchetti et al. [2004]. The index of agreement, shown in Figure 3, follows the same trend of the performance as a function of the input features and approaches 0.76 as its maximum. In fact models are trained to return a binary classification (exceedance/not exceedance) rather than to forecast the expected concentration; hence, their performances are optimal in terms of exceedances detection but are still good in terms of prediction capability.

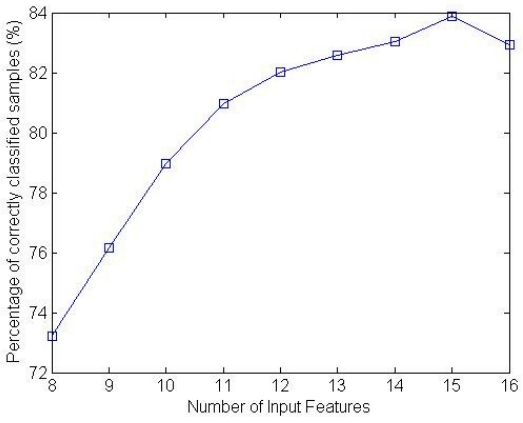


Figure 4. SVM Performance as a function of Input Features (samples above the threshold)

Different assignment for SVM parameters ε, σ and C, were tried in order to find the optimum configuration with the highest performance. When ε and C were kept constant (ε=0.001 and C=1000), the SVM performances depended on σ and reached a maximum when σ=1, corresponding to an optimum trade-off between SVM generalization capability (large values of σ) and model accuracy with respect to the training data (small values of σ). When σ and C were kept constant (σ=1 and C=1000), the best performances were achieved when ε was close to 0 and the allowed training error was minimized. From this observation, by abductive reasoning we could conclude that the input noise level was low. In accordance with such a behavior the performance of the network improved when the parameter C increased from 1 to 1000. Since the results tended to flatten for values of C greater than 1000, the parameter C was set equal to 1000. The best performance of the SVM corresponding to ε=0.001, σ =1 and C=1000 is shown in Figure 4 and Figure 5 for the samples above and below the threshold. *FA* was low (14.7%) while the capability to forecast when the concentrations were above the threshold was about 84%. The best performance of the SVM was achieved using as input features the best subset of 16 features. In Table 1 it is shown a comparison of

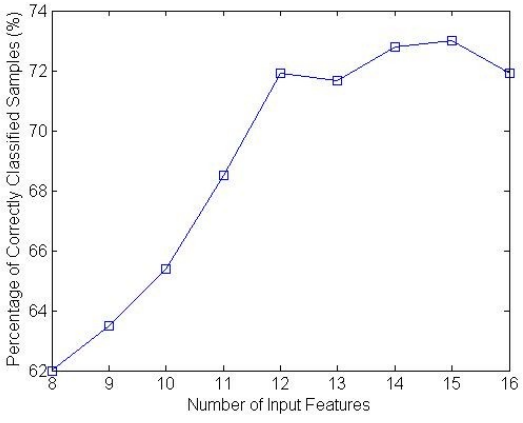


Figure 5. SVM Performance as a function of Input Features (samples below the threshold)

the performances of the SVM (ε=0.001, C=1000 and σ equal to 1) and the ANN (18 neurons in the hidden layer): performance indexes include *FA*; furthermore performances are also assessed in terms of average prediction ability by means of *IA*, the true/predicted correlation ρ and by the mean absolute error *MAE*.

Table 1. Best Performances of the ANN and SVM.

Machine Learning Method	% correct forecasting above the threshold	% correct forecasting below the threshold	*FA*	*MAE* ($\mu g/m^3$)	ρ	*IA*
ANN (18 Hidden Neurons; 15 Input Features)	85.4%	74%	14%	10.25	0.8	0.755
SVM (ε=0.001, C=1000 and σ=1; 15 Input Features)	83.8%	73.5%	14.4%	10.75	0.785	0.727

4. CONCLUSIONS

The training of the ANN and SVM will be improved with a proper data-deseasonalization (Cecchetti et al. [2004]) and with stacking techniques (Wolpert [1992]) using the measurements and forecasted values of the selected features as inputs. The stacking approach consists of iterating a procedure that combines measurements data and data which are obtained by means of prediction algorithms, in order to use them all as the input to a new prediction algorithm. Since for some pollutants the meteorological conditions are very important in the generation process, different neural networks will be trained for each geopotential condition, Benichou [1995]. The analysis will be completed extending the forecasting capability of the machine learning algorithm to areas where there are no measurement points, by means of the optimization of a multi-source gaussian dispersion model. Finally it could be interesting to carry out the same kind of analysis described in this paper for PM_{10} also for the other air pollutants.

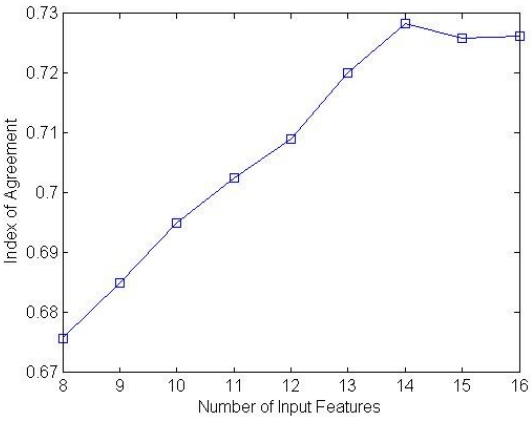


Figure 6. SVM Performance: Index of Agreement

ACKNOWLEDGEMENTS

We would like to thank the authors of the programs used for the scientific research related to this paper. We acknowledge S. Canu, Y. Grandvalet, V. Guigue, and A. Rakotomamonjy for the "SVM and Kernel Methods Matlab Toolbox ".

REFERENCES

ARPA Piemonte, Aria Web, https://secure.regione.piemonte.it/ambiente/aria/rilev/ariaweb/index.htm

ARPA Piemonte, Campagna di Rilevamento della Qualita' dell'aria con Utilizzo del Laboratorio Mobile nel Comune di Villastellone, Relazione Conclusiva. Aprile 2004 – Febbraio 2005.

Benichou, P., Classification automatique de configurations meteorologiques sur l'europe occidentale. Technical report. Meteo-France Monographie, 1995.

Benvenuto, F., Marani, A., Neural networks for environmental problems: data quality control and air pollution nowcasting. *Global NEST: The International Journal 2(3),* 281-292, 2000.

Božnar, M. Z., Mlakar, P., Grašič , B., Neural Networks Based Ozone Forecasting. *Proceeding of 9th Int. Conf. on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Garmisch-Partenkirchen,* Germany, 2004.

Calori G., Clemente M., De Maria R., Finardia S., Lollobrigida F. and Tinarelli G., Air quality integrated modelling in Turin urban area, *Environmental Modelling & Software*, 21:4, 468-476, April 2006.

Canu, S., Grandvalet, Y., Guigue, V. and Rakotomamonjy, A., SVM and Kernel Methods Matlab Toolbox. Perception Systèmes et Information, INSA de Rouen, Rouen, France. Available: http://asi.insarouen.fr/~arakotom/toolbox/index.html, 2005.

Canu, S., Rakotomamonjy, A., Ozone peak and pollution forecasting using Support Vectors. *IFAC Workshop on environmental modelling,* Yokohama, 2001.

Cecchetti M., Corani G., Guariso G., Artificial Neural Networks Prediction of PM_{10} in the Milan Area, *Proc. Of IEMSs 2004*, University of Osnabrück, Germany, 14-17 June 2004.

Costa M., Moniaci W., and Pasero E., INFO: an artificial neural system to forecast ice formation on the road, *Proc. of IEEE International Symposium on Computational Intelligence for Measurement Systems and Applications,* 216–221, 29-31 July 2003.

Demuth H., Beale M., Neural Network Toolbox User's Guide, The MathWorks, Inc., 2005.

Fletcher R., Practical Methods of Optimization, John Wiley & Sons, NY, 2nd ed., 1987.

Guyon I., Elisseeff A., An introduction to variable and feature selection, *The Journal of Machine Learning Research*, 3, 1157-1182, January 2003.

Koller, D., Sahami, M., Toward optimal feature selection. *Proc. of 13th International Conference on Machine Learning* (ICML), Bari, Italy, 284-292, 1996.

Marquardt D., An algorithm for least squares estimation of nonlinear parameters, *SIAM J. Appl. Math.*, vol. 11, 431-441, 1963.

Ostro, B., Chestnut L., Vichit-Vadakan N., and Laixuthai A.. The impact of particulate matter on daily mortality in Bangkok, Thailand. *Journal of Air and Waste Management Association*, 49, 100–107, 1999.

Pasero, E., Moniaci, W., Meindl, T., Montuori, A., NEMEFO: NEural MEteorological Forecast. *Proceeding of SIRWEC 2004, 12th International Road Weather Conference,* Bingen, 2004.

Parzen E., On Estimation of a Probability Density Function and Mode, *Annals of Math. Statistics,* vol. 33, 1065-1076, Sept. 1962.

Pearl J., Probabilistic Reasoning in Intelligent Systems, Morgan Kaufmann, San Mateo, CA, 1988.

Perez, P., Trier, A., Reyes, J., Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile, *Atmospheric Environment*, 34, 1189-1196, 2000.

Quaderno Tecnico ARPA (Emilia Romagna) - SMR n°10-2002.

Slini T., Kaprara A., Karatzas K., and Moussiopoulos N., PM_{10} forecasting for Thessaloniki, Greece, *Environmental Modelling & Software Volume 21, Issue 4*, 559-565, April 2006.

Vapnik V., The Nature of Statistical Learning Theory, New York, Springer-Verlag, 1995.

Werbos P., Beyond regression: New tools for Prediction and Analysis in the Behavioural Sciences, Ph.D. Dissertation, Committee on Appl. Math., Harvard Univ. Cambridge, MA, Nov. 1974.

Wilmott C.J., Davis R.E., Feddema J.J., Klink K.M., Legates D.R., Rowe C.M., Ackleson S.G. and O'Donell J., Statistics for the evaluation and comparison of models, *Journal Geophysics Research 90,* 8995-9005, 1985.

Wolpert D., Stacked generalization, *Neural Networks*, 5, 241-259, 1992.

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

Neural Network based method for predicting regional visitor attendance levels in recreational areas

T. Räsänen and M. Kolehmainen

Research Group of Environmental Informatics, Department of Environmental sciences, University of Kuopio, P.O. Box 1627, FIN-70211 Kuopio, Finland (teemu.rasanen@uku.fi, mikko.kolehmainen@uku.fi)

Abstract: The number of people in a certain place in desired time period is one of the main questions in many monitoring and management applications. Such information is needed also in tourism which has been one of most growing business areas in recent years. Increasing volume forces development of new practises for the sustainable and efficient management of recreational areas and tourist centres. Regional visitor monitoring provides general tools for managers and decision-makers to handle multidimensional growth and the development of tourism business. Our study aims to develop predictive approach for solving continuously regional visitor attendance levels using inferred data and computationally intelligent methods. In opposite to other known visitor monitoring systems, we used mobile telecommunications network to provide data for creation of estimate for number of people in certain region. Furthermore, regional weather conditions, traffic density, restaurant sales and the use of accommodation facilities were coupled together with mobile telecommunication event data. The Self-organizing Map (SOM) was used to integrate these variables into a combined regional attendance index, and the multilayer perceptron (MLP) was used to create the short-term predictions of visitor attendance levels. Finally the proposed continuous modelling system consisted online data gathering, server based modelling core and web-based user interface for information sharing. The system was tested and validated using real data gathered from the recreational area of Tahko. The regional visitor attendance level model and predictions were validated against expert opinions and regional freshwater consumption data. The results showed in general that the method is suitable for describing a real regional situation and seasonal variations in visitor attendance levels. Moreover, the results indicated that mobile telecommunication data improves predictions of daily visitors. Nevertheless, feedback from the possible end-users showed that presented method has potential in applications also in many other fields than tourism.

Keywords: positioning of masses of people, on-line monitoring, predicting, self-organizing map, multilayer perceptron

1. INTRODUCTION

The number of people in a certain place in desired time period is one of the main questions in many monitoring and management applications. Tourism and outdoor activities have been growing in recent years and have become a more important area of the service sector [Lim and McAleer, 2005]. Moreover, the popularity of outdoor sports, leisure activities and nature tourism has been growing in recent years. Increasing use of the landscape and natural forests has aroused severe problems, including local conflicts between commercial activities and the use of nature [Krämer and Roth, 2002]. Thus, effective and sustainable

management and systematic visitor monitoring is needed. Regional visitor monitoring can be used as a general tool for managers and decision-makers to handle multidimensional growth and the development of tourism business. In more detail, sustainable management encompass activities like minimizing raw material loss, efficient waste management and optimizing public transport, needed staff or opening hours.

In past studies, Cessford, Cockburn and Douglas [2002] has classified traditional visitor monitoring techniques into three types: (1) direct observations using staff observers or camera recordings at sites, (2) on-site counters or other devices to record visitor presence, and (3) inferred counts, i.e. the use of other data to obtain on-site estimates. Furthermore, combinations of these approaches are often used. Tourist centres and recreational areas have typically plenty of activities linked to its surroundings; visitors are typically very mobile and spread quite wide area inside the centre. Therefore it is not practically possible to construct satisfying visitor monitoring system using on-site counters.

However, the main aim of our study was to develop a new continuous and predictive approach for estimating regional visitor attendance levels using inferred data sources and computationally intelligent methods. The presented method was tested with created predictive on-line modelling system using real data gathered from the recreational area of Tahko [Tahko, 2008] which is the biggest combined downhill, cross-country ski and leisure resort in the southern half of Finland.

2. MATERIALS AND METHODS

The data based approach for predicting regional visitor attendance was based on time-series data gathered from the following sources in the Tahko area. The data used contained mobile telecommunications event data from the networks of the mobile operator Finnet Ltd., accommodation data from Tahkovuori Ltd., traffic density data from the National Road Administration's local measuring point, daily restaurant sales data from TahkoChalet Ltd., and weather data from the Finnish Meteorological Institute's local measuring station. The variables used in creation of visitor attendance model and predictions are presented in Table 1.

Table 1. Variables used in modelling.

yd_sin	Day of year (1-365) (transformed into continuous form using the sine-function)
yd_cos	Day of year (1-365) (transformed into continuous form using the cosine function)
holnum_1	Binary variable for a normal workday
holnum_2	Binary variable for a Saturday or eve of a feast-day
holnum_3	Binary variable for a Sunday or feast-day
sin_hour	Sine of hour
cos_hour	Cosine of hour
temp	Temperature
hum	Humidity
calls	Number of outgoing calls
rcalls	Number of calls received
grps	Number of grps data transfer events
sms	Number of sms messages
adults	Daily number of adults with accommodation bookings
childs	Daily number of children with accommodation bookings
restaurant	Daily sale in euros
Traffic_sum	Number of light vehicles

Typically cell-based positioning is used to determine the geographical location of individual mobile user. However, in this case the cell-based information was used, in a

novel way, to determine the number of people in a certain region at a given time. Simplified demonstration of positioning of masses of people is illustrated in Figure 1. The mobile telecommunication data were collected from seven base transceiver stations (BTS) encompassing the whole area of Tahko ski resort. These stations were are part of the mobile (GSM) network of Finnet Ltd., which has an approximately 20% market share in the Finnish mobile phone sector.

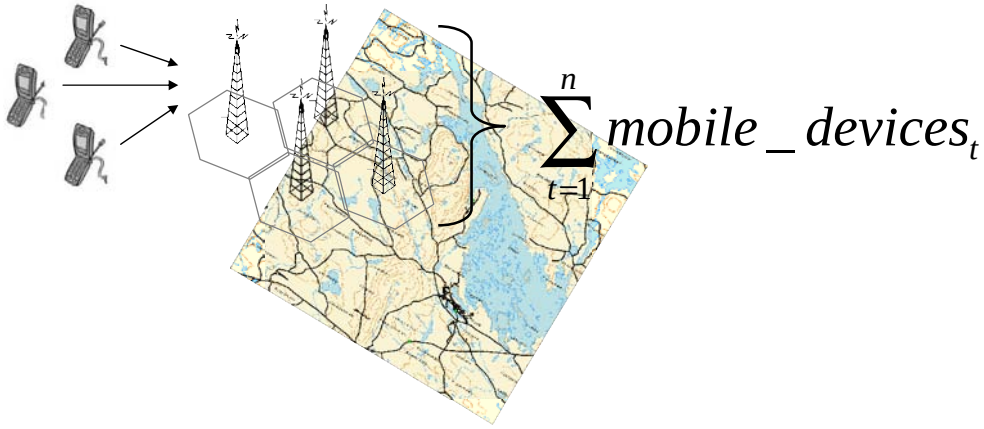


Figure 1. Simplified demonstration of using mobile telecommunication network to locate masses of people in a certain region. The number of mobile devices or the number of telecommunication events in each network cell are used to estimate number of people.

The modelling and predicting regional visitor attendance levels was based on automated data processing chain which is summarized in Figure 2. The data collection was done via internet, by email and using Java Web Services interfaces, depending on the source data management system. System was updated every morning after the data was collected and estimate for the following seven days were presented in web interface. Additionally, all the historical information was also available in web application and comparison between different time periods could be done easily.

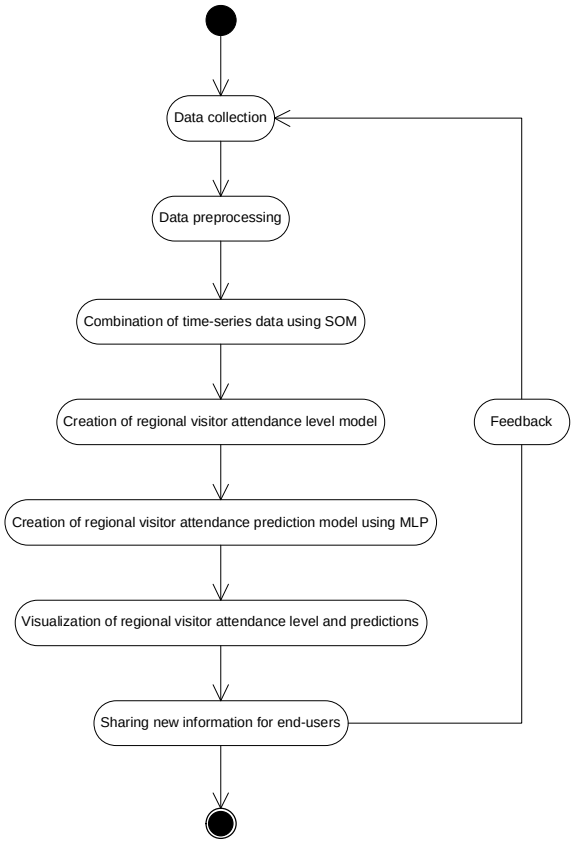


Figure 2. The simplified schema of on-line data processing chain for predicting regional visitor attendance levels.

The regional model for attendance level was constructed using the Self-Organizing Map (SOM), one of the best-known unsupervised neural learning algorithms [Kohonen, 1997; Kolehmainen et al., 2001]. The SOM was trained using dataset containing variables described in Table 1. After that the attributes of its prototype vectors was used to construct the combined regional visitor attendance level model. All the variables correlated with the actual number of visitors in the area are chosen as key variables, in this case the number of telecommunication events, number of people with booked accommodation, restaurant sales and traffic density. The scaled values for these key variables on each prototype vector were summarized and the neurons of the SOM were ordered according to their sum values. The neuron which had the highest sum of scores for the key variables contains the time periods when the area reached its maximum visitor attendance level, and conversely, the neuron which had the minimum sum of scores for the key variables contains the time periods when the area had its minimum attendance level. Finally, all the neurons were fitted between the maximum and minimum values and a percentage model was created.

The Multilayer Perceptron (MLP) [Haykin, 1999; Hecht-Nielsen, 1991; Gardner and Dorling, 1998] was used to create a short-term regional visitor attendance level prediction for the next seven days. The multi-layer perceptron (MLP) is the most commonly used type of feed-forward neural network, with a structure consisting of processing elements and connections. The processing elements, called neurons, are arranged in layers, comprising an input layer, one or more hidden layers and an output layer. The input layer serves as a buffer that distributes input signals to the next layer, which is a hidden layer. Each unit in the hidden layer sums its input processes with a transfer function and distributes the result to the output layer. It is also possible for several hidden layers to be connected in the same fashion. The units in the output layer compute their output in a similar manner. For a more thorough review, the reader is referred to Haykin [1999].

In this case the parameters of MLP were following; the number of neurons in hidden layers was 15, the learning algorithm used was resilient backpropagation (trainrp) and the performance function of mean square error was applied. The transfer functions selected were hyperbolic sigmoid tangent for the hidden layer and linear for the output layer.

Statistical indicators were used to provide a numerical description of the goodness of the model. Indicators like index of agreement (IA), root mean square error (RMSE) and coefficient of determination (R^2) were calculated according to Equations 1-3. The use of different operational performance indicators for evaluating models has been discussed at a more detailed level in Willmot [1982].

$$IA = 1 - \left[\frac{\sum_{i=1}^{N} (P_i - O_i)^2}{\sum_{i=1}^{N} (\left|P_i - \overline{O}\right| + \left|O_i - \overline{O}\right|)^2} \right] \quad (1)$$

$$RMSE = \left(\frac{1}{N} \sum_{i=1}^{N} [P_i - O_1]^2 \right)^{\frac{1}{2}} \quad (2)$$

$$R^2 = \frac{\sum_{i=1}^{N} (P_i - \overline{O})^2}{\sum_{i=1}^{N} (O_i - \overline{O})^2} \quad (3)$$

3. RESULTS

The MLP model was used to perform short-term predictions of visitor attendance levels, produced using SOM, for weekly periods. The continuous system produced forecast of next seven days every morning at seven o'clock. The results were promising and daily time-series comparison between model (solid line) and predicted visitor attendance level (dashed line) is presented in Figure 3. This plot applies to the time period between 5th March and 23th October 2005. The goodness of MLP prediction model was validated using scatter plots (Figure 4) and statistical indicators and the results of calculations are shown in Table 2.

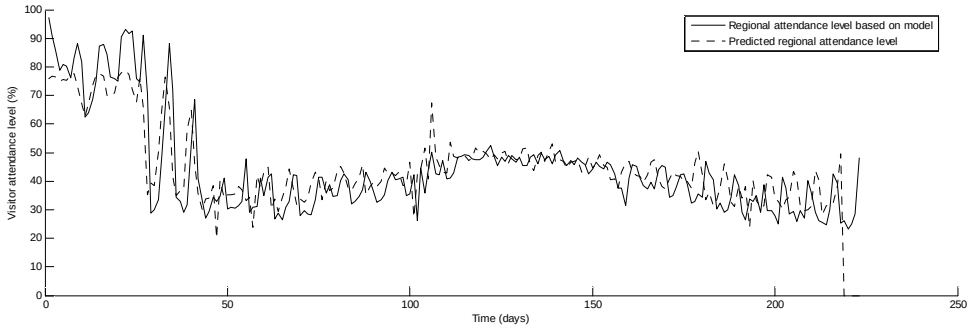


Figure 3. The regional visitor attendance level model (solid line) versus predicted situation (dashed line) for Tahko area. The resolution of the data was 24 hours and used time period was 5th March to 23th October 2005.

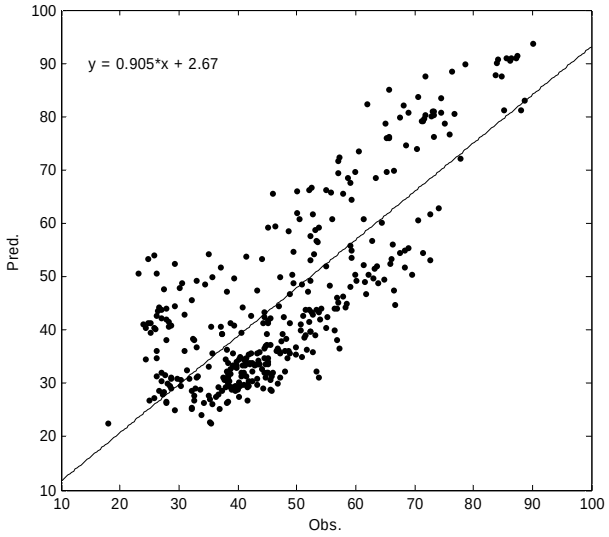


Figure 4. A Scatter plot of the predicted (+7 days) versus the actual value of the visitors' attendance. The least-squares fitting line is presented with solid line.

Table 2. The goodness of the predicted visitor attendance model by statistical indicators.

Statistical indicator	**+1 day**	**+2 day**	**+3 day**	**+4 day**	**+5 day**	**+6 day**	**+7 day**
Index of agreement	0.93	0.91	0.90	0.90	0.91	0.89	0.88
Root mean square error	11.51	12.99	14.00	14.21	13.10	12.82	10.04
Coefficient of determination	1.01	1.02	1.02	1.01	1.01	0.99	0.76

4. CONCLUSIONS

Efficient and sustainable management of tourist centres demands, not only detailed and comprehensive information about characteristics of visitors and activities, but also up to date information about number of visitors in certain time period. Presented approach can be used continuously with dynamic on-line data acquisition for modelling and prediction tasks. In this way visitor attendance level information and predictions are always available to support different tasks of decision making. The multi-phased data processing chain was used for regional visitor attendance level modelling, the self-organizing map was employed to construct the combined regional visitor attendance model and the multilayer perceptron was used to create short-term predictions.

Moreover, the mobile telecommunications data were used in novel way, as this time-series information on the numbers of customer telecommunication events in the region was used to locate a mass of subscribers and to predict the level of visitor attendance in the area. Results showed that inferred data and these computationally intelligent methods provide useful predictive on-line tools for improving sustainable decision-making and business planning in the field of tourism. In practise, presented web-based modeling system were applied to optimizing number of operating taxi vehicles, planning opening hours and times of grocery shop and rescheduling the collection of municipal wastes.

On the other hand, validation of the modelling results was difficult because there is no appropriate basis for validation which describes total regional visitor attendance levels or the total numbers of visitors. Expert opinions and other comparable data (i.e. fresh water consumption) therefore constituted the firmest basis for validation.

The presented prediction approach was part of the bigger information system which could be used also in many other application fields than tourism. This kind of intelligent system which produces continuous near real-time information about number of people in certain

place has great potential for example in air quality exposure detection, traffic census, urban town planning, disaster or emergency management or planning routing of hazardous cargo. In conclusion, living in an urban environment affects citizen health and quality of life in many ways. Therefore, in environmental impact or risk assessments, it is crucial to notify in more detailed level how many citizen are actually exposed possible undesirable phenomenons like terrible air quality or other emissions. Such information, could be linked to other environmental modelling systems like OSCAR [Sokhi et el., 2008], EXPAND [Kousa et al., 2002] or AirQUIS [Slordal et al., 2008]. Modern technology and data based approaches enables to calculate number of people in certain place using even smaller than hour time resolution.

ACKNOWLEDGEMENTS

The authors wish to thank the Finnish Funding Agency for Technology and Innovation and Finnet Association for the financial support of the project.

REFERENCES

Cessford, G., Cockburn, S. and Douglas, M. Developing New Visitor Counters and their Applications for Management. International Conference on Monitoring and Management of Visitor Flows in Recreational and Protected Areas proc. (MMV). 14-20. 2002.

Gardner, M.W. and Dorling, S.R. Artificial neural networks (the multi-layer perceptron) - a review of applications in the atmospheric sciences. *Atmospheric Environment 32*, 2627-2636. 1998.

Haykin, S. *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, Englewood Cliffs, NJ. 1999.

Hecth-Nielsen, R. *Neurocomputing. Reprinted with corrections*. Addison-Wesley, Reading, MA. 1991.

Kohonen, T. *Self-Organizing Maps, 2nd Edition*. Springer-Verlag Berlin Heidelberg. Germany. 1997.

Kolehmainen, M., Martikainen, H. and Ruuskanen, J. Neural networks and periodic components used in air quality forecasting. *Atmospheric Environment, 35*, 815-825. 2001.

Kousa, A., Kukkonen, J., Karppinen, A., Aarnio, P. Koskentalo, T. A model for evaluating the population exposure to ambient air pollution in an urban area. *Atmospheric Environment 36*, 2109-2119. 2002.

Lim, C. and McAleer, M. Ecologically sustainable tourism management. *Environmental Modeling & Software, 20*, 1431-1438. 2005.

Sokhi, R.,S. et al. An integrated multi-model approach for air quality assessment: Development and evaluation of the OSCAR Air Quality Assessment System. *Environmental Modeling & Software, 23*, 268-281. 2008.

Slordal, L.H., Mc Innes, H., Krognes, T. The Air Quality Information System AirQUIS. *Information Technologies in Environmental Engineering, Special Issue, January 2008.*

Tahko. Website of Tahko Ski Resort. http://www.tahko.com/. 2008.

Willmot, C. Some comments on evaluation of the model performance. *Bulleting of American Meteorological Society, No 63*, pp. 1309-1313. 1982.

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

Neural Networks and Co-Kriging techniques to Forecast Ozone Concentrations in Urban Areas

C. Carnevale, G. Finzi, E. Pisoni*, V. Singh and M. Volta

DEA – University of Brescia, Via Branze 38, 25123, Brescia, Italy
({carneval, finzi, enrico.pisoni, vikas.singh, lvolta} @ ing.unibs.it)

Abstract: An integrated forecasting system, consisting of Neural Network (NNs) models and co-kriging techniques, has been developed to forecast maximum eight hours ozone (max8h) concentration, two days in advance, over an urban domain including Milan area of northern Italy. Total numbers of available measurement stations falling within the domain are 23. NNs perform the forecasting at each measurement location and the co-kriging algorithm interpolates the forecasting data all over the domain. NNs have been identified for the period of 2000-2006. Leave-One-Out Cross Validation (LOOCV) has been performed to validate the results of NNs. To perform spatial interpolation of the forecasted maximum daily eight hour (max8h) ozone, co-kriging has been used. For validation of the proposed forecasting system, 5 out of 23 stations have been selected. Year 2004 has been chosen as a test case year to perform the overall forecast. The validation results show good agreement in terms of statistical indexes. The proposed forecasting methodology represents a fast and reliable way to provide decision makers and general public with ozone forecasting data over an urban area.

Keywords: Ozone forecasting; Neural Networks; co-kriging, Spatial Interpolation.

1. INTRODUCTION

Since last decade tropospheric ozone episodes have become more and more critical over Europe, mainly in Southern metropolitan regions during summer months, due to the sun radiation significant role played in photochemical transformations of urban and industrial NO_x and VOC emissions. Because of the environmental risk of ozone exposition, the EU Directive 2002/3/EC, following the WHO guidelines, prescribes air quality standards for ozone in terms of threshold values for health protection, population information and warning. In order to prevent critical episodes and to inform the population, proper real time alarm modelling systems have to be set up.
This paper proposes an integrated method, harmonizing Neural Networks [Corani (2005), Agirre-Basurko et al., (2006), Schlink et al., (2006), Sousa et al., (2007)] and co-kriging algorithms [Isaaks et al. (1990), Clayton et al. (1997)], to forecast the max8h ozone over an urban domain which includes Milan metropolitan area (Northern Italy). NNs are used to provide the forecast at the measurement locations, co-kriging to perform spatial interpolation of the forecasted data.

2. METHODOLOGY

Developed forecasting system consists of two parts:

1. Stochastic modeling system (Neural Network system) to forecast, two days in advance, daily max8h of ozone over each measurement station within the study domain.
2. Interpolation System to perform forecasted ozone maps over whole domain.

2.1 Neural Network system

Elman neural networks (Elman, 1990) have been identified to perform the forecasting for each measurement station. Elman NN implements a vectorial function $f : R^Q \rightarrow R^L$, where Q

and *L* are the dimension of the input and output vectors of the net respectively. The *l-th* element of the vector function *f* for the *n-th* pattern is defined as:

$$f_l(v^n) = \log sig(\sum_{m=1}^{M}(OW_{l,m} \cdot a_m^n) + g_t)$$

Where:

$$a_m^n = \tan sig(\sum_{q=1}^{Q}(IW_{m,q} \cdot v_q^n) + \sum_{w=1}^{W}(FW_{m,w} \cdot a_w^{n-1}) + b_m)$$

$$\tan sig(x) = \frac{2}{1+\exp(-2x)} - 1$$

and M is the number of the neurons in the hidden layer.
The matrices *IW (M×Q)*, *OW (L×M)* and *FW (M×M)* are the input, output and feedback weight matrix respectively, and *b (M×1)* and *g (L×1)* vectors are the bias terms. NNs weights (IW, OW, FW, b and g) are tuned on a training dataset by means of a back-propagation algorithm (Mathworks, 2006).
Ozone measurement time series have been divided in identification and validation data sets; identification data set spans from 2000 to 2006 leaving out one year each time for validation.

2.2 Interpolation System

Interpolation System involves three steps:

1. Experimental semi-variogram and cross-variogram calculation;
2. Fitting and modeling of variograms using LCM (Linear Model of Coregionalization) [Pardo-Iguzquiza et al., 2002];
3. Estimation of max8h ozone over unmeasured locations using co-kriging method.

2.2.1 Experimental Semi-variogram and Cross-variogram Calculation

Variogram characterizes the spatial continuity or roughness of a data set and represents the variance of the increments. Empirical semivariogram *γ(h)*, computed as half of variogram, is a measure of the relation between pairs of points:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} \left[z(u_\alpha) - z(u_\alpha + h)\right]^2$$

where $z(u_\alpha)$, $z(u_\alpha + h)$ are the measurements in the points u_α, $u_\alpha + h$; *N(h)* is the number of pairs of points, the distance of which is *h*.
When primary data is not sufficient to represent the spatial structure, a more intensely sampled data, which is correlated with the primary data, is used to obtain the spatial structure of primary variable. Cross-variogram $\gamma_{zy}(h)$ represents the spatial relationship between primary $z(u_\alpha)$ and secondary variable $y(u_\alpha)$ and is defined as:

$$\gamma_{zy}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} \left[z(u_\alpha) - z(u_\alpha + h)\right]\left[y(u_\alpha) - y(u_\alpha + h)\right]$$

In present case, the primary variable is max8h ozone forecast at measurement locations over the domain and TCAM model [Carnevale et al., 2008] simulated ozone maps for summer 2003, calculated in the framework of an APAT (Italian Environmental Protection Agency, www.apat.gov.it/) project, has been used as secondary variable.

2.2.2 Fitting of Variograms and Linear Coregionalization Model

The empirical variograms are fitted by an analitycal funcion (model). In this work the linear coregionalization model (LCM) [Pardo-Iguzquiza et al., 2002] has been applied. It is a sum of two or more proportional covariance (semi-variogram) models. A proportional covariance model is the simplest multivariate model used in geostatistics and is one in which all the semi-variograms $\gamma_{ij}(h)$ are proportional to a single semi-variogram $\gamma(h)$ function:

$$\gamma_{ij}(h) = b_{ij}\gamma(h)$$

where, b_{ij} are symmetric coefficients, that define a positive definite matrix $B=[b_{ij}]$.
Fitting a LCM comprises the following steps:

a) All direct semi-variograms and cross-semi-variograms are estimated for the same number of lags and the same lag distances *h*.
b) The number and types of elementary models and their ranges are postulated.
c) The sills (coregionalization matrix) are fitted by optimization technique (WSS [Pardo-Iguzquiza et al., 2002]).

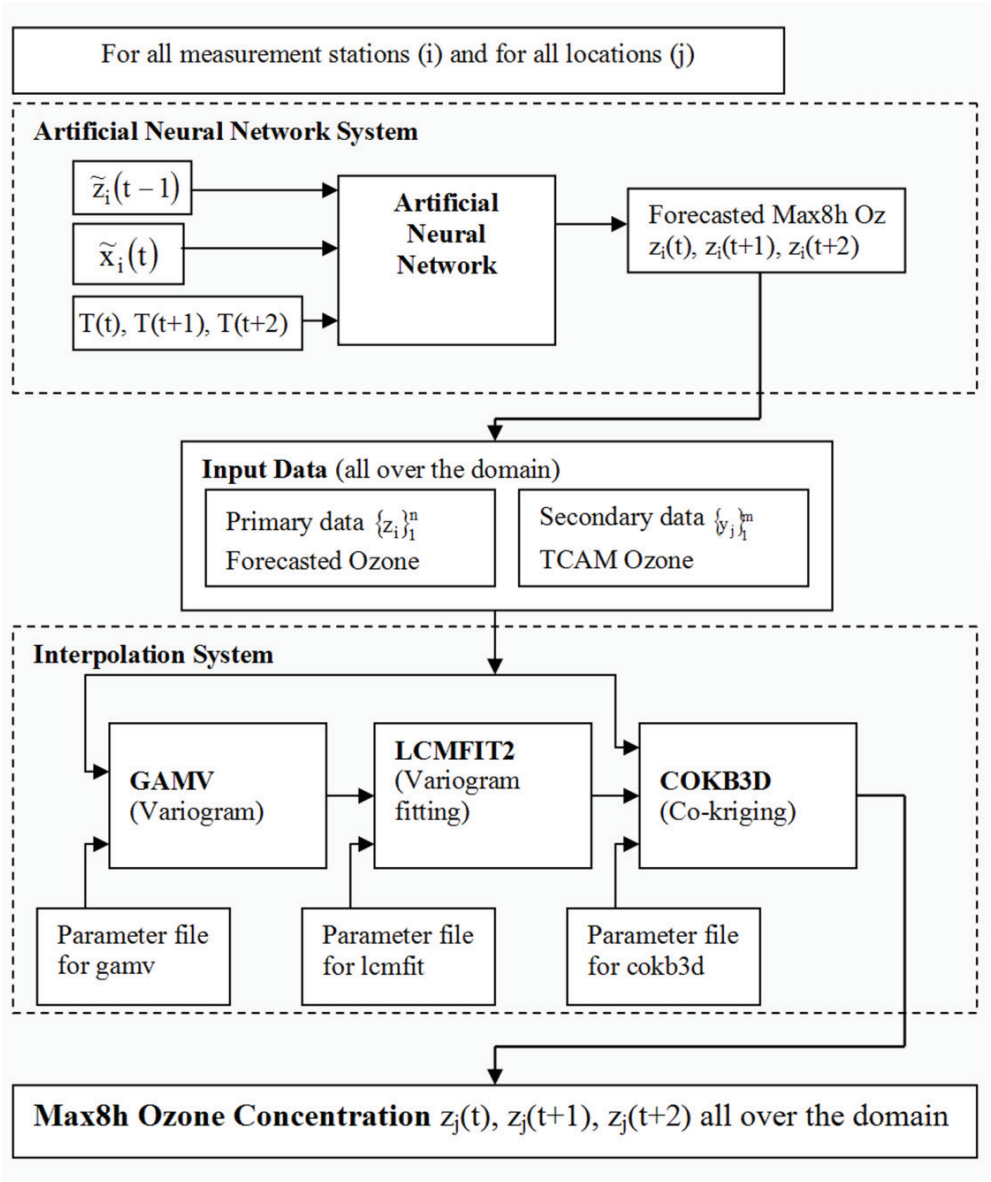


Figure 1. Block diagram of developed forecasting system

2.2.3 Co-kriging

Co-kriging is an interpolation technique that allows one to better estimate primary variable if the distribution of a secondary variable is sampled more intensely than the primary variable. The co-kriging estimate is a linear combination of both primary and secondary data values and is given by

$$Z_0 = \sum_{i=1}^{n} a_i.z_i + \sum_{j=1}^{m} b_j.y_j$$

where Z_0 is the estimated max8h ozone at location *0*; $\{z_i\}_1^n$ are the forecasted, by NNs, max8h ozone at n nearby locations and $\{y_j\}_1^m$ are the TCAM ozone at m nearby locations;

$\{a_i\}_1^n$ and $\{b_i\}_1^m$ are the co-kriging weights. These weights are calculated by inverting covariance matrices of each distance pair and multiplying it with the covariance matrix of the distance pair from the estimation point. The covariance matrix are generated using the semi-variogram and cross-variogram models.
Fortran 77 GSLIB libraries [http://www.statios.com/, http://www.gslib.com/] have been used for co-kriging system. GAMV code has been used to calculate experimental Semi-variogram and Cross-variogram and COKB3D code has been used to calculate the max8h ozone estimates. For variogram fitting LCMFIT2 program [Pardo-Iguzquiza et al., 2002] has been used. Block diagram of the developed system is shown in Figure 1, where $\tilde{z}_i(t-1)$ are the max8h ozone measurements for the previous day; $\tilde{x}_i(t)$ are the maximum first twelve hours ozone concentration for t; T(t), T(t+1), T(t+2) are the mean temperature; $z_i(t)$, $z_i(t+1)$, $z_i(t+2)$ are the NNs forecasted Max8h Ozone for each station and $z_j(t)$, $z_j(t+1)$, $z_j(t+2)$ are the forecasted Max8h Ozone concentration all over the domain for the corresponding days.
The developed methodology has been tested over a densely inhabited and industrialized area located in the Northern Italy domain and including Milan.

3. RESULTS

3.1 Case study domain

Test case has been performed to forecast max8h ozone, for the entire year 2004, over Northern Italy domain, including Milan metropolitan area. The domain has a dimension of 60x60 km^2, divided in 144 cells of 5x5 km^2 (Figure 2). The domain has 23 measurement stations (Figure 2), and each measurement station has max8h ozone measured from 2000 to 2006. Test case and validation of the results have been performed in two steps. In the first step, NNs Leave-One-Out Cross Validation (LOOCV) has been performed to forecast max8h ozone and in second step, performance of whole system, including co-kriging system, has been tested and validated for 2004.

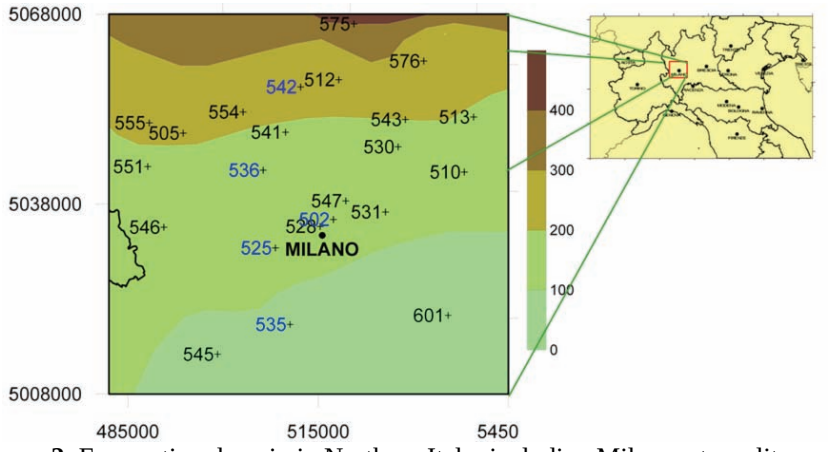


Figure 2: Forecasting domain in Northern Italy including Milan metropolitan area, showing all 23 measurement stations (stations in blue are validation stations).

3.2 Neural Network Validation

LOOCV has been performed for the period of 2000-2006 by leaving each year at a time. Excluded year is the validation data set and rest of the data set is training data set for the NNs. In this study, NNs inputs are the max8h ozone concentrations of previous day (t-1) , the maximum value of first 12 hour ozone for today (t) and the forecasted daily mean temperature for t, t+1 and t+2. The target patterns are the maximum eight hours ozone for today (t), tomorrow (t+1) and day after tomorrow (t+2). To perform each day forecast (t, t+1 and t+2), three individual networks have been identified for each station.

Forecasted max8h ozone has been compared with the corresponding day of measurement. Correlation coefficient and mean error has been computed for each LOOCV year and for all

23 stations. Box plots of correlation coefficient and mean errors for t, t+1 and t+2 are shown in Figure 3 and 4 respectively. Year 2004 has been chosen as a test case year and detailed statistical indexes for selected stations (502, 525, 535, 536 and 542) are shown in table 1. It is clear that networks are able to correctly forecast max8h ozone but performance of the forecast get worse advancing in the day of forecast.

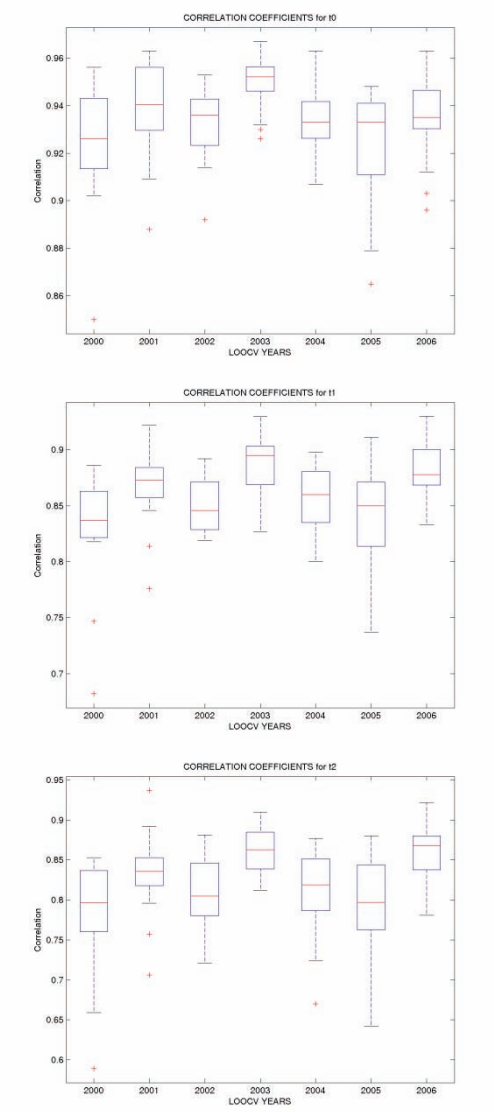


Figure 3: LOOCV correlation coefficients for t, t+1 and t+2

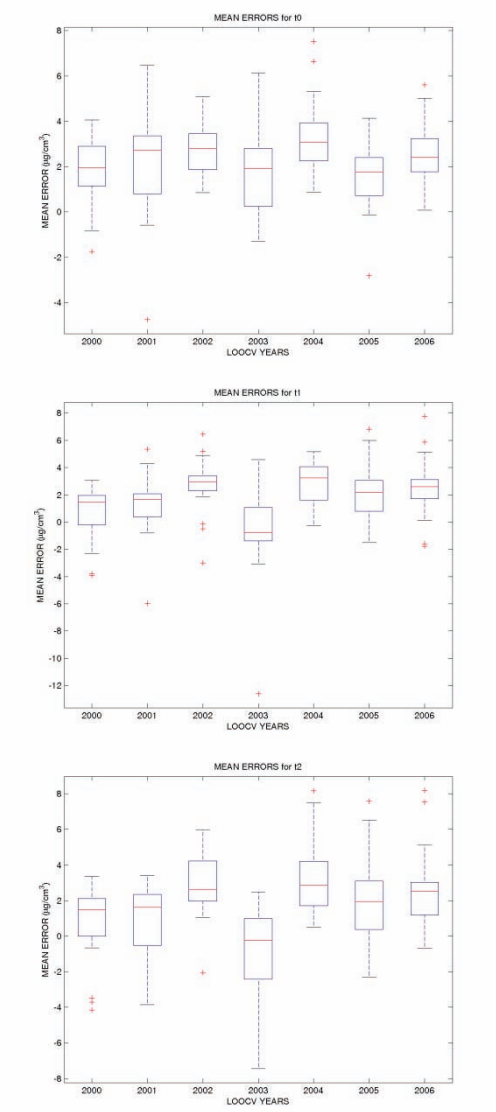


Figure 4: LOOCV mean errors for t, t+1 and t+2

3.3 Forecasting system validation

In the second step, co-kriging has been performed using NNs forecasted max8h ozone values, excluding five stations (502, 525, 535, 536 and 542) out of total 23 stations. These stations have been used as validation stations for forecasting system. Spatially interpolated max8h ozone at these five stations has been compared with the max8h ozone measurement at the respective stations and statistics have been computed to validate the results. Computed statistics for five validation stations, for t, t+1 and t+2, are shown in Table 2.

Table 1: Statistics[1] of the NNs validation for t, t+1 and t+2.

			t		
Stations	**502**	**525**	**535**	**536**	**542**
M. Err	0.94	1.51	1.63	0.86	-0.88
RMSE	13.29	15.73	13.60	14.72	12.52
Corr	0.94	0.92	0.96	0.95	0.97
SR	47%	47%	77%	64%	90%
SP	47%	45%	79%	62%	70%
			t+1		
M. Err	1.33	3.12	4.75	0.96	0.10
RMSE	19.27	22.49	23.12	23.00	23.90
Corr	0.88	0.85	0.88	0.86	0.87
SR	31%	23%	45%	50%	63%
SP	33%	35%	62%	38%	45%
			t+2		
M. Err	1.52	3.64	5.69	1.04	0.60
RMSE	20.68	23.63	24.46	24.77	27.75
Corr	0.86	0.83	0.87	0.83	0.82
SR	20%	16%	47%	40%	50%
SP	20%	30%	65%	28%	36%

Table 2: Statistics[1] of the forecasting system validation for t, t+1 and t+2

			t		
Stations	**502**	**525**	**535**	**536**	**542**
M. Err	2.95	2.94	-3.66	2.25	2.57
RMSE	15.73	19.44	16.76	17.52	14.26
Corr	0.92	0.89	0.93	0.92	0.95
SR	57%	42%	77%	67%	87%
SP	53%	67%	52%	69%	77%
			t+1		
M. Err	3.88	4.28	-3.25	3.86	2.37
RMSE	21.00	24.60	23.45	23.75	24.14
Corr	0.86	0.83	0.87	0.85	0.87
SR	40%	24%	59%	47%	66%
SP	27%	38%	38%	52%	43%
			t+2		
M. Err	4.08	5.12	-3.09	4.93	2.33
RMSE	22.20	25.44	24.80	24.58	27.57
Corr	0.83	0.81	0.85	0.84	0.82
SR	14%	15%	50%	31%	55%
SP	7%	24%	32%	38%	32%

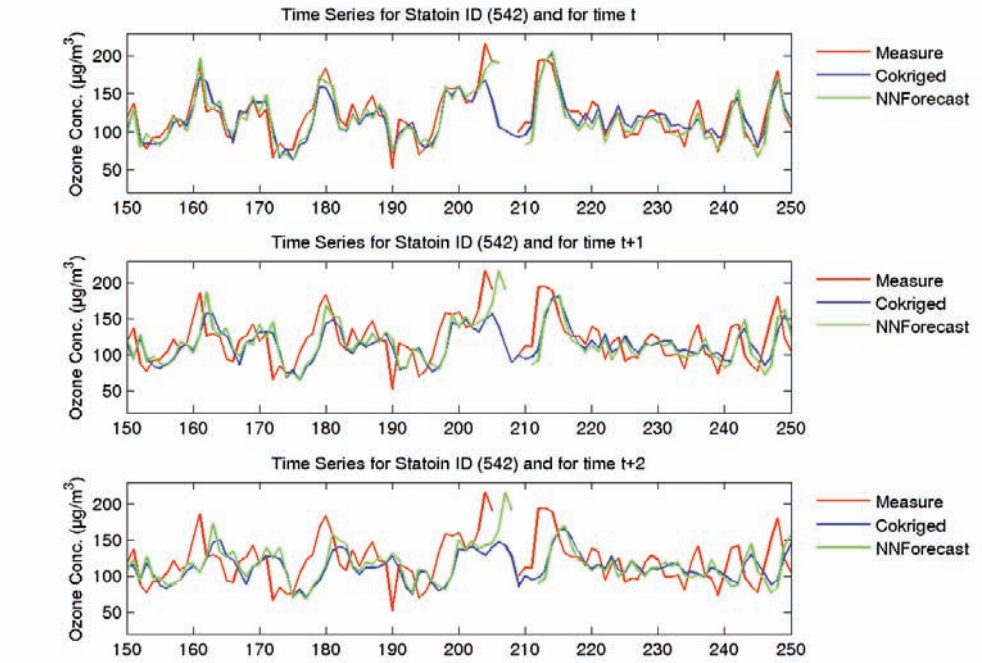


Figure 5: Comparison between measured, NN forecasted and co-kriged max8h ozone for time t, t+1 and t+2.

The performance of the forecast gets worse for the next days of forecast. For example, the correlation, which is close to 0.9 for t, decreases with time, and varies between 0.82-0.86

[1] The indexes in Table 1 and Table 2: **M. Err:** Mean Error; **RMSE:** Root mean square error; **Corr:** Correlation; **SR**: Percentage of forecasted exceedances correctly predicted over measured exceedances (threshold of 120 µg/m3); **SP**: Percentage of forecasted exceedances correctly predicted over forecasted exceedances (threshold of 120 µg/m3).

for t+1 and 0.81-0.85 for t+2. If other statistics are considered, they also worsen with time. The degradation in the performance can also be marked while looking at SP and SR indexes computed by the forecasting system. It is clear that stations 535, 536 and 542 (located inside the Milan) give better performances then the stations 502 and 525, which are close to Milan highway.

Figure 5 shows the time series for 100 days of 2004, showing measured, NN forecasted and Co-kriged max8h ozone concentration ($\mu g/m^3$) for station 542 for t, t+1 and t+2. These figures also show the degradation in the forecasted results for next days forecast. Figures 6-8 show the spatial measured and forecasted images of max8h ozone concentration ($\mu g/m^3$) for 4th, 5th and 6th of August 2004. It can be seen that the developed methodology is able to reproduce spatial patterns of the daily max8h ozone over the domain.

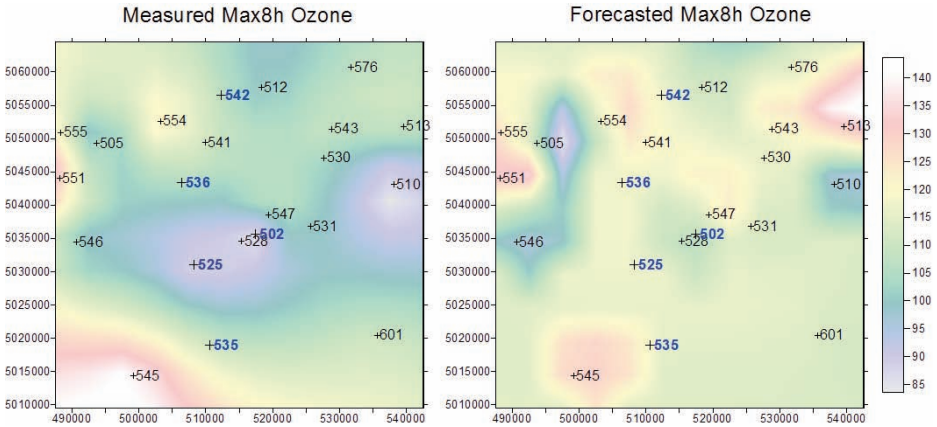


Figure 3: Measured and forecasted maps of max8h ozone over the domain for t

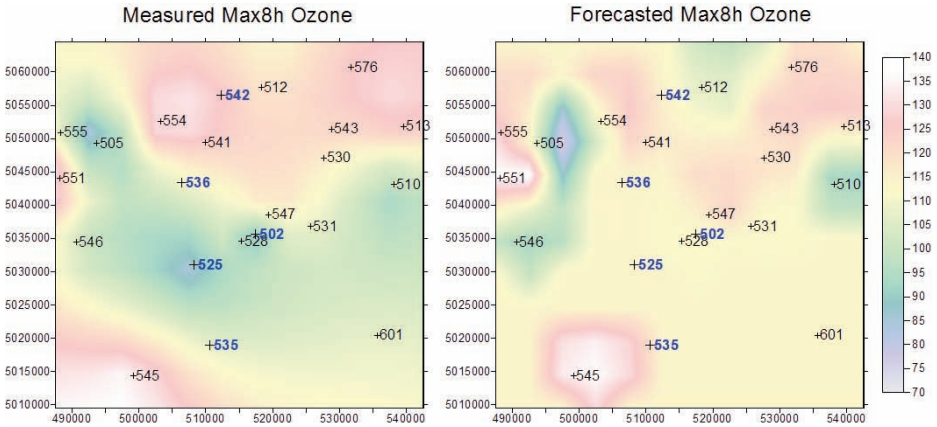


Figure 4: Measured and forecasted maps of max8h ozone over the domain for t+1.

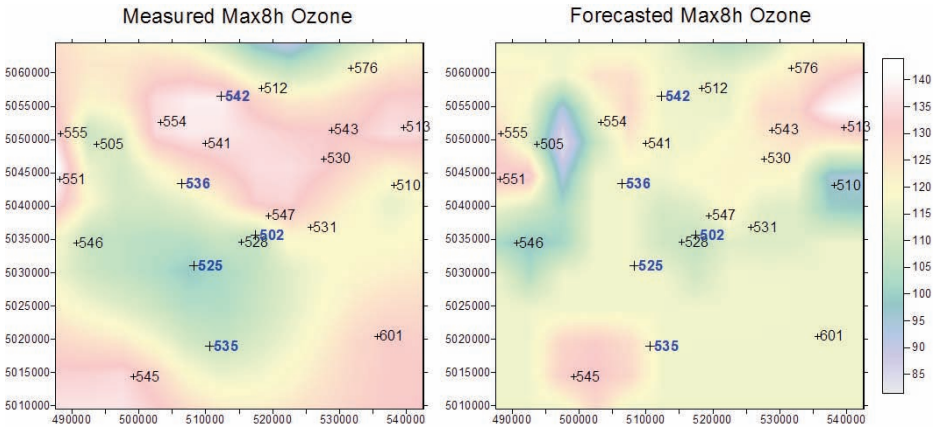


Figure 5: Measured and forecasted maps of max8h ozone over the domain for t+2.

4. CONCLUSIONS

Developed methodology, integrating NNs and co-kriging is able to forecast daily maximum eight hours ozone three days in advance. The results of the methodology application show good agreement in terms of statistical indexes, i.e. showing fair correlation coefficient values. The quality of forecast gets worse for the next days of forecast. This methodology also reproduces the spatial patterns of the daily max8h ozone over the domain. The proposed forecasting methodology represents a fast and reliable way to provide decision makers and general public with ozone forecasting data on an urban area.

ACKNOWLEDGEMENTS

The research has been developed in the framework of the Pilot Project QUITSAT (Qualità dell'aria mediante l'Integrazione di misure da Terra, da Satellite e di modellistica chimica multifase e di Trasporto – contract I/035/06/0 – http://www.quitsat.it), sponsored and funded by the Italian Space Agency (ASI).

REFERENCES

Agirre-Basurko E., Ibarra-Berastegi G. and Madariaga I., Regression and multilayer perceptron-based models to forecast hourly O3 and NO2 levels in the Bilbao area, *Environmental Modelling & Software* 21, 430-446, 2006.

Hooyberghs Jef et al,. Spatial interpolation of ambient ozone concentrations from sparse monitoring points in Belgium, *Journal of Environmental Monitoring*, Vol 8, 1129-1135, 2006.

Carnevale C., Decanini E. and Volta M., Design and validation of a multiphase 3D model to simulate tropospheric pollution, *Science of The Total Environment*, Volume 390, 166-176, 2008.

Clayton V. Deutsch, Andre G. Journel. GSLIB: Geostatistical Software Library and User's Guide, Oxford University Press, 1997.

Corani, G., Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning. *Ecological Modelling* 185, 513 - 529, 2005.

Gabusi V. and Volta M., Seasonal modelling assessment of ozone sensitivity to precursors in Northern Italy, *Atmospheric Environment* 39, 2795–2804, 2005.

Isaaks E.H. and Srivastava R.M., An introduction to applied geostatistics. Oxford University Press, New York, 1990.

Neftel A., Spirig C., Prevot A., Furger M., Stutz J., Vogel B., Hjiort J., Sensitivity of photooxidant production in the Milan Basin: An overview of results from EUROTRAC-2 Limitation of Oxidant Production field experiment. *Journal of Geophysical Research* 107 (D22), LOP1 1–10, 2002.

Pardo-Iguzquiza E., Dowd P.A., FACTOR2D: a computer program for factorial cokriging. *Comput Geosci* 28:857–875, 2002

Schlink U., Herbarth O., Richter M., Dorling S., Nunnari G., Cawley G. and Pelikane E., Statistical models to assess the health effects and to forecast ground-level ozone, *Environmental Modelling & Software*, 21, 547-558, 2006.

Sousa S.I.V., Martins F.G., Alvim-Ferraz M.C.M. and Pereira M.C., Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations, *Environmental Modelling & Software 22*, 97-103, 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

Studying and predicting quality of life atmospheric parameters with the aid of computational intelligence methods

K D. Karatzas and D. Voukantsis
Department of Mechanical Engineering, Informatics Applications and Systems Group, Aristotle University, 54124 Thessaloniki, Greece
(kkara@eng.auth.gr, voukas@isag.eng.auth.gr)

Abstract: Air quality management is among the most challenging problems in terms of analysis and modelling. Air quality modelling and forecasting is directly affected by the highly nonlinear relationships between pollutants and weather, while in many cases there is insufficient domain knowledge due to the influences of local conditions. As atmospheric quality has an impact on the quality of life of millions of people, the ability to reveal interrelationships between parameters that influence environmental decision making is very important. In addition, forecasting of such parameters for the purpose of early warning and health risk prevention is of paramount importance for sensitive parts of the population. In the present paper a number of Computational Intelligence methods are presented with the purpose to the investigate atmospheric quality parameters, towards better understanding of the interrelationships between pollutants, and with the aim to improve forecasting of critical values. For this reason, the use of Fast Fourier Transformation for the construction of Periodograms is firstly presented, followed by the application of Principal Component Analysis. Then, Self Organizing Maps, a method based on the Neural Networks approach, is investigated and applied, for knowledge extraction and atmospheric parameter analysis. Last, an Artificial Neural Network based on the multi-linear perceptron (MLP) model is presented in order to construct prediction models. Results indicate a number of important features within the data investigated, and reveal hidden interrelations, thus providing valuable information for the understanding and the explanation of environmental problems, and for the support of environmental policy and decision making in both long and short terms. It is also demonstrated that the performance of forecasting models justify their selection for early warning information services.

Keywords: Computational intelligence, periodograms, principal component analysis, self organising maps, artificial neural networks, air quality, atmospheric environment, quality of life.

1. INTRODUCTION

A knowledge domain is described on the basis of nominal, categorical or arithmetic values of parameters that serve as the basis for information creation, as they are being processed with the aid of various computational methods, tools or human judgment. In the case of the environmental engineering domain, these data have (in the majority of cases) the form of time stamped records that formulate a multivariate time series within the spatial and temporal scale of the phenomenon of interest.

Although air pollution is “interwoven” to the atmospheric environment, pollutants (i.e. harmful agents) behave differently in various spatial and temporal scales. Nevertheless, it has been agreed upon by the scientific community that certain pollutants should be monitored in locations that are representative of urban spatial forms like city centres, high traffic areas, residential areas, suburban or rural areas. Thus, the quality of the atmospheric environment may be described with the aid of hourly concentration values of various

pollutants like Ozone (O3), Nitrogen Oxides (NO, NO2 etc), Sulphur Dioxide (SO2), Particulate Matter with mean aerodynamic diameter of various scales (coarse, PM10, PM2.5, ultra fine), and other pollutants. In addition, a number of meteorological parameters influence the quality of air and play an important role in our understanding concerning the life cycle of atmospheric pollution. These are parameters like wind speed and wind direction, air temperature, relative humidity, etc.

2. COMPUTATIONAL INTELLIGENCE FOR ATMOSPHERIC PARAMETER INVESTIGATION

The understanding of the relationships, dependencies, profiles and behaviour of parameters describing the quality of the atmospheric environment is of paramount importance for anyone interested to improve quality of life, especially in urban areas. Moreover, the ability to forecast such parameters, and especially those values that affect human health, is critical in all health prevention systems, Although many AQ modelling methods have been applied in this field, Computational Intelligence (CI) provides with an arsenal of scientific approaches that may be applied in order to better understand (and forecast) the behaviour of parameters of interest. Each urban area has its own physical, geographical and meteorological characteristics. Thus knowledge gained in one area may enrich our knowledge for another area and improve our understanding of environmental pressures. On this basis, a number of methods will be described in the next chapters aiming at demonstrating the advantages resulting from the usage of computational intelligence for atmospheric quality study and forecasting.

2.1 Fourier Transformation and Periodogram Construction

The discrete Fourier transform of a stationary discrete time series x(n) is a set of N-discrete harmonics X(k) according to eq. 1.

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j 2\pi k n / N} \qquad (1)$$

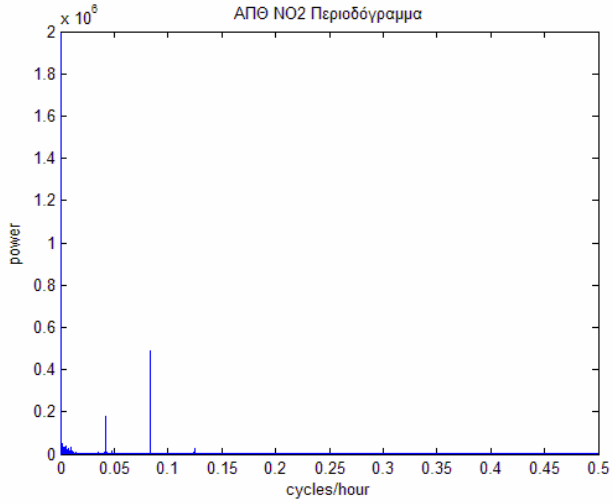


Figure 1. Periodogram of nitrogen dioxide time series (AUTh station, Thessaloniki)

This transformation represents the frequency response of x(n). X(k) is complex. The magnitude of X(k) squared, a real value, is called strength and a diagram of all the strength harmonics is called periodogram. When x(n) is real then the periodogram is symmetric and only half of the harmonics are needed [Karatzas et al., 2007]. Here x(n) is the hourly concentration values of pollutants. The method is demonstrated via its application on data coming from an urban station located in the city of Thessaloniki, Greece, in the campus of

the Aristotle University. These data describe NO2 hourly concentrations for the years 2001-3 (Figure 1). Thessaloniki is a city of 1 million inhabitants and more than 400 000 vehicles, suffering from the asphyxiating pressures of urban development. Thus, it is common than atmospheric pollutants reach high values, and the investigation of their behaviour is of high importance for the authorities responsible for pollution abatement. The specific periodogram reveals that the main periodicity of NO2 is approx 0.08 cycles/hour, i.e. 12 hours. This means that this pollutant has a 12 hour cycle, thus its existence should be attributed to a source having the same periodicity, i.e. urban traffic, a finding indicating the influence of vehicle circulation on air quality in the city centre.

2.2 Principal Component Analysis

Principal component analysis (PCA) is a computational intelligence method originating from multivariate statistical analysis that allows for the identification of the major drives within a certain multidimensional data set and thus may be applied for data compression, identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. Since patterns in data can be hard to find in high dimensional data, where the luxury of graphical representation is not available, PCA is a powerful tool for such an analysis. The other main advantage of PCA is that once these patterns have been identified the data can be compressed, using dimensionality reduction, without significant loss of information [Smith, 2002]. In order to demonstrate the effectiveness of the method, PCA was applied for atmospheric datasets derived from the air quality monitoring station located at Aristotle University, in Thessaloniki, Greece (same with the one used before). The software package employed for this purpose was MATLAB that has a built-in function for applying the method to any given dataset. Data for the period 2001-2003 were used. The data matrix contained measurements for O3, NO2, temperature, humidity, wind speed and wind direction transformed according to eq. 2 into the two new variables, namely sinwd and coswd [Karatzas and Kaltsatos, 2007].

$$v1 = \sin(2\pi(v - \min(v))/(\max(v) - \min(v))) \\ v2 = \cos(2\pi(v - \min(v))/(\max(v) - \min(v))) \quad (2)$$

The method's results as well as the correlation coefficients matrix are presented in Table 1 and Table 2 respectively.

Table 1: PCA results, AUTH station, 2001 – 2003.

	PC1	PC2	PC3	PC4
NO_2	-0.3049	-0.4226	0.0413	-0.7218
O_3	0.5345	0.0421	0.1121	0.1121
Temp	0.3773	-0.2596	0.6095	0.0932
Hum	-0.4233	-0.1367	-0.2845	0.5643
WS	0.3644	0.3986	-0.4673	-0.2001
SinWD	-0.3587	0.3469	0.5232	0.1583
CosWD	-0.1959	0.6741	0.2030	-0.2723
%Var	41.8937	20.7019	13.5555	11.1588
Cum. %Var	41.8937	62.5956	76.1511	87.3099
	PC5	PC6	PC7	
NO_2	0.2097	0.3084	0.2582	
O_3	0.2396	0.0062	0.7937	
Temp	-0.2171	0.5344	-0.2782	
Hum	-0.0627	0.5722	0.2672	
WS	0.3485	0.4993	-0.2814	
SinWD	0.6676	0.0297	-0.0749	
CosWD	-0.5295	0.2042	0.2643	
%Var	5.3721	4.6824	2.6355	
Cum. %Var	92.6820	97.3645	100	

Table 2: Correlation coefficient matrix, AUTH station (2001 – 2003).

	NO_2	O_3	Temp	Hum
NO_2	1.00	-0.51	-0.18	0.20
O_3	-0.51	1.00	0.59	-0.62
Temp	-0.18	0.59	1.00	-0.45
Hum	0.20	-0.62	-0.45	1.00
WS	-0.41	0.52	0.04	-0.42
SinWD	0.09	-0.42	-0.26	0.29
CosWD	-0.08	-0.28	-0.31	-0.00

	WS	SinWD	CosWD
NO_2	-0.41	0.09	-0.08
O_3	0.52	-0.42	-0.28
Temp	0.04	-0.26	-0.31
Hum	-0.42	0.29	-0.00
WS	1.00	-0.34	0.08
SinWD	-0.34	1.00	0.48
CosWD	0.08	0.48	1.00

Table 3: Varimax rotation to the first 2 PCs, AUTH station (2001 – 2003).

	PC1	PC2
NO_2	-0.51	-0.12
O_3	0.43	-0.32
Temp	0.12	-0.44
Hum	-0.41	0.17
WS	0.54	0.06
SinWD	-0.05	0.50
CosWD	0.29	0.64

The cumulative variance percentage of each Principal Component (PC) is considered to be a measure of the representativity of the information contained in the initial data set. Each coefficient (with values from -1 to 1), represents the "weight" of each parameter in the formulation of each PC. The PCs are normalized, so that the squares of each PC coefficients sum to one.

The Humphrey – Ilgen parallel analysis indicates that only 2 PCs, explaining 62.59% of the total variance of the data, should be considered for further analysis, while the rest PCs are more likely to represent random variations of the data. The varimax rotation was applied to the coefficients of the first 2 PCs, and the results are presented in Table 3. It is evident that PC1 reveals the antagonistic relationship between NO_2 with O_3, and thus indicating photochemistry influenced by traffic. Furthermore, the strong contribution of Humidity and Wind Speed to PC1, suggests that O_3 formation is mainly expressed by this PC. In contrast, PC2 expresses the production dispersion of O_3, since it is mostly characterized by wind temperature and direction. These findings are supported by the fact that NO_2 is the result of the chemical degradation of O_3 in the atmosphere, while humidity accumulates under conditions of weak solar radiation, the latter acting as a catalyst (and thus being required) in the production of O_3 from primary pollutants.

2.3 Self Organising Maps

The self-organizing map (SOM) is also referred to as Kohonen Network [Kohonen, 1997], and is a subtype of artificial neural networks. SOM is based on competitive learning, which runs in an unsupervised manner, aiming at selecting the so called winning neuron that best matches a vector of the input space (Figure 2). In this way, "a continuous input space of activation patterns is mapped onto a discrete output space of neurons by a process of competition of the neurons in the network" [Haykin, 1999]. This makes SOM one of best methods for modelling a knowledge domain with the aim to reveal topological interrelations and hidden knowledge, via the visualization of the network's neurons.

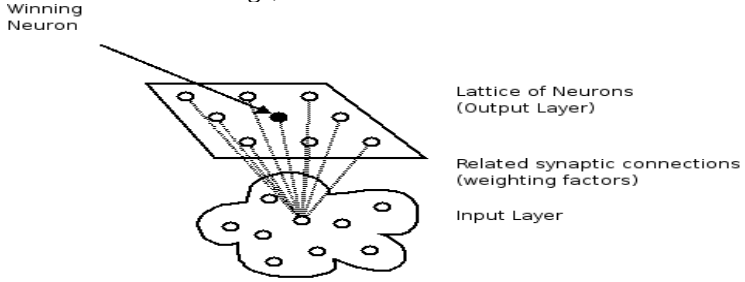


Figure 2. The generalised Kohonen Network. The input space is mapped on the output layer via the winning neurons. The topological relationship of the latter reveal dependencies and relationships between the parameters of the input space, on the basis of the features mapped on the output layer.

SOM is capable of learning from complex, multi-dimensional data without specification of the output. The resulting nonlinear classification consists of clusters that can be interpreted via visual inspection. The methods' unsupervised learning algorithm involves a self-organizing process to identify the weight factors in the network, reflecting the main features of the input data as a whole. In that process the input data is mapped onto a lower dimensional (usually two-dimensional) map of output nodes with little or no knowledge of the data structure being required. The output nodes (neurons) represent groups of entities with similar properties, revealing possible clusters in the input data. It should be noted that, although the method is unsupervised in learning, the number of the output nodes and configuration of the output map (number of nodes included, etc), need to be specified before the learning process [Karatzas and Kaltsatos, 2007b].

The air quality and meteorological data used for the demonstration of SOM capabilities, originate from the following monitoring stations in Thessaloniki, Greece: Aristotle University (city centre, within the University Campus), Kalamaria (urban stations in the east side of the city), Eleftherio-Kordelio (urban stations in the west side of the city) and Sindos (industrial area, located in the west of the city). As a first step to apply the SOM method, input variables were be normalized to unit variance, since Euclidean distance is employed as the error metric, and thus, if the input variables are not normalized, the mapping realized by the SOM may be dictated by some variable which has a much larger variance than the others. Then, the aim of the demonstration was to investigate air quality in one specific location (the Aristotle University- AUTh monitoring station).

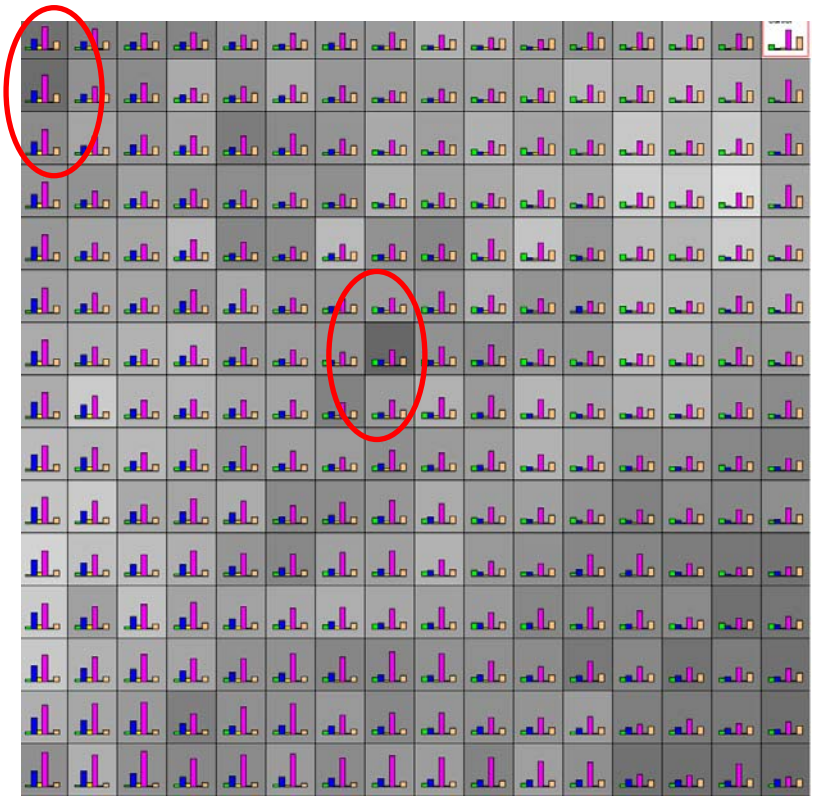


Figure 3. The resulting Self Organising Map for the AUTh station. Each square represents a neuron for which the frequency of occurrence is analogous to the darkness of the grey tone, and the values of the parameters constituting the neuron is analogous to the height of the relevant bars. These parameters (from left to right) are: NO_2, O_3, T, WD, WS, RH.

Figure 3 reveals two dark grey zones, which may be interpreted as clusters of neurons possessing s strong topological relationship. The cluster on the upper left reveals that high NO2 values (the 2nd data bar) are associated with low O_3 (1st data bar), a finding that is expected, due to the photochemical interconnection of these two pollutants. This cluster also reveals that high NO_2 occurs when WD (4th data bar) is high, i.e from 270-360 deg., and WS (5th) is low, indicating that there might be some local transportation from the

western parts of the city. This may be attributed to the sitting of major industrial combustion sources in the western parts of the area. The cluster in the centre reveals that WD (4th data bar) of values around 180 deg (i.e. south) are associated with high RH values, indicating that those are the winds transferring wet air over the city centre. This may be related to the fact that the local sea breeze circulation supports such air mass interactions between land and sea surfaces.

2.4 ANN models for modelling and forecasting ozone concentrations in Thessaloniki, Greece

In addition to the SOM method, Artificial Neural Networks were applied in order to demonstrate forecasting capabilities concerning hourly ozone concentration levels. Ozone is a photochemical pollutant that is being created as a result of the chemical imbalance in the atmosphere caused by pollutants like nitrogen dioxides and hydrocarbons. These pollutants enter the atmosphere mainly as a result of anthropogenic activities, and have a high and direct impact to the quality of life of citizens, especially those belonging to the so called sensitive part of the population.

The network architecture chosen here was the Multi-Layer Perceptron (MLP) with one hidden layer: the network consisted of one input layer, one hidden layer and the output layer. MLP models are widely applied in predicting air pollutant concentrations since they can capture the highly nonlinear relationship between the variables [Gardner and Dorling, 1999; Kolehmainen et al., 2001; Yi and Prybutok, 1996]. The datasets used include hourly concentration values for O_3, NO_2, and hourly records of temperature, humidity, wind speed and the transformed wind direction, as all these parameters were proven to be of importance for describing the investigated datasets via the PCA method previously presented. The choice of one hidden layer was made after several tests with different network structures since it gave us lower error values and smaller convergence times. Table 3 compares the forecasting performance statistical indexes for two different models. Model A uses NO_2, temperature, humidity, wind speed and transformed wind direction as prediction variables, while model B uses the one hour lagged O_3 concentration as an additional prediction variable

Table 3: Evaluation of two ANN models concerning hourly O_3 concentrations at AUTh, Thessaloniki

	Model A	Model B
MAE	27..326	13.532
RMSE	34.19	18.476
IA	0.814	0.948

It is clear that the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) are better for model B. In addition, the Index of Agreement (IA), used for the evaluation of the model's ability to forecast the actual (observed) data, is much higher in the case of model B, indicating the fact that the use of O_3 values of previous hours improves the forecasting performance.

3. CONSLUSIONS AND RECOMENDATIONS

In the present paper a number of Computational Intelligence methods have been applied in order to demonstrate their efficiency for both problem understanding and parameter forecasting concerning air pollution concentration values. The results presented verify the high ability of the methods, and also reveal their unique ability to extract knowledge from the domain of interest and to help us understand and manage environmental problems. On this basis, it would be recommended that such analysis carried out in cases where problems related to the atmospheric environment and quality of life is concerned, in order to better inform policy and decision makers about the constrains of their decisions. In addition, such methods may provide with the necessary scientific competence for services related to public notification, warning and alerts, on the basis of forecasts.

ACKNOWLEDGEMENTS

This paper is related to the work that the authors are conducting in the frame of COST Action ES0602: Towards a European Network on Chemical Weather Forecasting and Information Systems (www.chemicalweather.eu). The data used originate from the freely accessible air quality database (AirBase) of the European Environment Agency (http://air-climate.eionet.europa.eu/databases/airbase) and from the Prefecture of Central Macedonia, Dept. of Environment, Greece, and are related to the project Airthess: Early warning informatics system for air pollution in Thessaloniki (www.airthess.gr).

REFERENCES

Haykin S., *Neural networks a comprehensive foundation*, Pearson Education, 2005

Gardner, M.W., Dorling, S.R., "Neural network modelling and prediction of hourly NOx an NO2 concentrations in urban air in London", *Atmospheric Environment*, vol. 31, 1999, pp. 709–719.

Karatzas K., G. Papadourakis and I. Kyriakidis (2007), Understanding and Forecasting Air Pollution with the Aid of Artificial Intelligence Methods in Athens, Greece. Proceedings of the Int. Workshop on Applications with Artificial Intelligence, 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI) October 29-31, 2007, Patras, Greece (http://ictai07.ceid.upatras.gr/). Springer Series on Studies in Computational Intelligence, in press.

Karatzas K., and Kaltsatos S. (2007), Air pollution modelling with the aid of computational intelligence methods in Thessaloniki, Greece, Simulation Modelling Practice and Theory, Volume 15, Issue 10, November 2007, Pages 1310-1319

Karatzas K. and Kaltsatos S. (2007b), Modelling and forecasting of air quality with the aid of machine learning methods in Thessaloniki, Greece, Proceedings of the 10th International Conference on Engineering Applications of Neural Networks, (Margaritis N. and Iliadis L., eds.), ISBN 978-960-287-093-8, pp. 419-427

Kohonen T., *Self organizing maps*, Springer, 1995

Kolehmainen M., Martikainen H., Ruuskanen J., "Neural networks and periodic components used in air quality forecasting", *Atmospheric Environment*, vol. 35, 2001, pp. 815-825.

Smith L., "A tutorial on Principal Components Analysis", (February 26, 2002, http://kybele.psych.cornell.edu/~edelman/Psych-465-Spring-2003/PCA-tutorial.pdf, last accessed 20.02.2008).

Yi, J., Prybutok, V.R., (1996), A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area, *Environmental Pollution*, vol. 92, pp. 349–357.

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 multiscale Air Quality Forecast System for Europe and selected regions

H. Jakobs, M. Memmesheimer, E. Friese and A. Ebel

Rhenish Institute for Environmental Research at the University of Cologne, Aachener Str. 209, D-50931 Cologne, Germany (Hermann.Jakobs@eurad.uni-koeln.de)

Abstract: A real-time forecast system for atmospheric pollutants is presented. The air quality forecast system is based on the EURAD Model (European Air Pollution Dispersion Model). The daily updated forecast of atmospheric constituents is operational since November 2001. The chemistry transport model EURAD calculates the transport, transformation and deposition of atmospheric constituents for a period of 72 hours on a daily basis. Major products of the AQ forecast system are the atmospheric pollutants O_3, NO_2, SO_2, CO, PM10 and benzene. An intensive effort is done to carry out a verification procedure of the model forecast using observational data from the German Environmental Agency (UBA) monitoring network. Examples of major air pollution events and parts of the verification procedure are shown in section 4. The results of the forecasts on the different regions are published and are updated every day on the EURAD homepage **www.eurad.uni-koeln.**

Keywords: Chemical weather, Modelling, Ozone, Particulate Matter.

1. INTRODUCTION

One major challenge in the last decade was to build up an air quality forecast system compared to numerical weather prediction systems. With increasing computer power it is now possible to run complex chemical weather forecast under near real time conditions. The complex EURAD Model (European Air Pollution Dispersion Model) was established 20 years ago at the Rhenish Institute for Environmental Research (RIU) at the University of Cologne. It calculates the transport, transformation and deposition of atmospheric constituents. The applicability of the model system has been proven through numerous regional simulations of tropospheric composition and boundary layer air quality (e. g. Jakobs et al., 2002; Memmesheimer et al., 1997). In order to apply the EURAD model for operational forecast, a special version of the model was used to speed up the consuming computer time. The features of the EURAD AQ Forecast System and the some results of the daily prediction together with validation results are presented in this paper.

2. THE EURAD AIR QUALITY FORECAST SYSTEM

The air quality forecast service is provided by the multiscale Eulerian chemical transport model system EURAD (European Air Quality Dispersion Model). The EURAD model is well-tested over the course of many case studies (Ebel et al., 1997). The model system was extended for use as a forecast model and has run operationally since 1st November, 2001 with its standard configuration at the RIU in Cologne. Meanwhile the EURAD AQ forecast system was applied for other regions of interest: Turkey, Ireland, the German State of Lower Saxony.

The key features of the model are: High flexibility for selecting forecast domains, an advanced heterogeneous chemistry mechanisms with comprehensive aerosol and photo oxidant chemistry and a focalized and high resolution forecasting by hemispheric/continental to regional scale (optionally 1 km resolution) nesting techniques with an integrated meteorological driver model.

The EURAD Air Quality Forecast System consists of three major components: The PennState/NCAR mesoscale model MM5 (Anthes and Warner, 1978; Grell et al., 1994) which predicts the required meteorological variables, the EURAD Emission Module (EEM) (Memmesheimer et al., 1991) which calculates the temporal and spatial distribution of the emission rates of the major pollutants and the EURAD Chemistry Transport Model (EURAD-CTM) which predicts the concentrations and deposition of the main atmospheric pollutants. Figure 1 gives an overview of all major components, its pre- and post-processors and the relevant input and output data sets. The chemical mechanisms employed in the EURAD system are the so-called RADM2 and its successor RACM. They have been completed by the aerosol mechanism MADE (Modal Aerosol-Dynamics model for EURAD). The RADM2 mechanism contains 63 reactive species treated in 158 chemical reactions. There is an option to run the code with the more sophisticated RACM chemistry as well. Detailed aqueous phase chemistry for the treatment of the air pollutants is incorporated. The horizontal and vertical transport is carried out using the 4th order Bott advection scheme. Vertical mixing of the species is treated by an implicit vertical diffusion scheme. The sink at the lower boundary of the model is treated by wet and dry deposition parameterization. The major driver for wet deposition is the predicted precipitation. The dry deposition is calculated via the deposition velocity for each species, which depends from the particle itself, the atmospheric dynamic and the given land-use type.

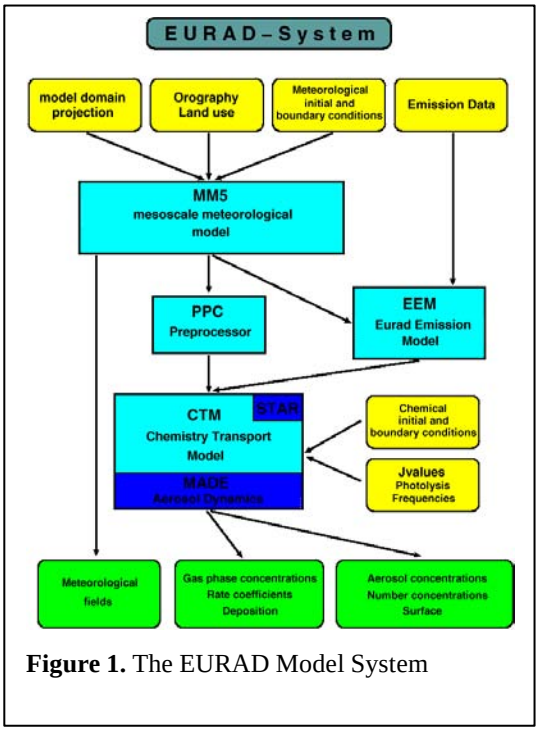


Figure 1. The EURAD Model System

3. THE FORECAST CYCLE

The standard RIU AQ forecast was designed for three different domains: Europe (D01 with 125 km grid resolution), Central Europe (D02, with 25 km grid resolution) and the German state Northrhine-Westfalia (D03, with 5 km grid resolution). The EEM calculates the temporal and spatial distribution of the emission rates of the major pollutants from the available data bases. The EEM was constructed to process different data bases, ranging from continental down to local scale. The biogenic emissions are calculated online with respect to the given atmospheric condition (temperature, radiation, wind) and the given land-use type. The meteorological forecast is obtained using the NCAR/PennState mesoscale model, MM5. For initial and boundary conditions, the model uses the NCEP/GFS global forecast and interpolates the variables on the selected domains.

The forecast model system uses the method of nested simulations. This enables consistent modeling of air quality from small (local) to large (continental) scales. The required geographical information (topography, land-use type) is taken from the USGS data base with respect to the selected horizontal resolution. Applications with coarse resolution usually cover the major part of Europe. They can be zoomed down to regions of the size of central Europe and fractions of it (e.g. Province level). The model uses terrain-following s coordinate in the vertical, with 23 unequally spaced layers, the more dense resolution being used in the lowest part of the model. The model top in the operational mode is 100 hPa.

The regular forecast system starts automatically with the download of the 00 UTC NCEP GFS global meteorological forecast via ftp. Then consecutively the chemical weather forecast was performed from the mother domain (D01) down to the domain covering the German State of Northrhine-Westfalia (D03). Usually the forecast cycle for the standard configuration is finish at around 08:00 MET in the morning, delivering the needed data for three consecutive days.

4. DAILY AQ FORECAST RESULTS AND VALIDATION

Every day, an extensive amount of data is produced by the EURAD forecast system for Europe, Central Europe and the German State of Northrhine-Westfalia. This includes the meteorological prediction variables and the concentrations of the atmospheric constituents at all model levels. In order to be able to subsequently compare the modeled concentrations of air pollutants with the observational data, a considerable effort was made to accurately model the near surface concentrations of the main air pollutants and to produce a combined Air Quality Index (AQI) for the above mentioned domains. For assessment studies, the ranges for the concentration thresholds were selected according to the EU directives. The daily maxima, the daily maximum 8h running mean and the daily maxim 24h running mean of each of the pollutants are calculated. In addition, animations and chemograms of selected regions are produced. These products are graphically processed and published on the EURAD website (http://www.riu.uni-koeln.de). As results, two characteristic episodes are shown to demonstrate the performance of the model.

4.1. Long range transport of PM10, NO_2

One major long range pollution event took place at 15/16 October 2005. With a high pressure system located over Scandinavia, easterly winds were predominant in the central and western part of Europe. Thus pollutants were transported from the main emission centres in Germany, Benelux, UK towards west and reached the relatively clean regions: Ireland and even Iceland at the end of the period. Figure 2 shows the maximum 24h mean of PM10 at the days 15 and 16 October 2005. In addition high concentrations of NO_2 were forecasted during this event over Central Europe. For the validation of the forecast besides the forecasted NO_2 near surface concentrations, the NO_2 column densities, derived from measurements of the OMI satellite are also shown in Figure 3. Since the vertical distribution of NO_2 exhibit an exponential decrease with height above the mixing height, near surface concentrations can be used for comparison with column density values. The comparison of the results shown in Figure 3 indicates that there is a good agreement of the forecast products with observational data.

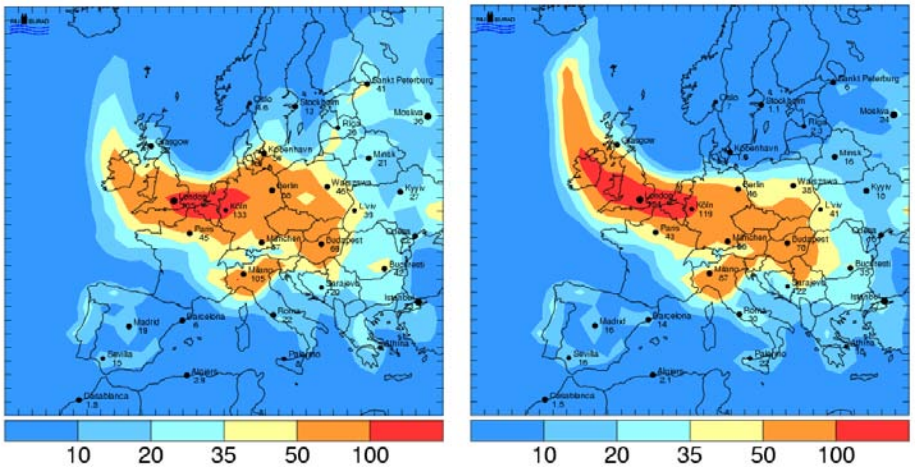


Figure 2. AQ Forecast: near surface maximum 24h mean of PM10 in µg/m³ at 15 October 2005 (left) and 16 October 2005 (right) for domain D01

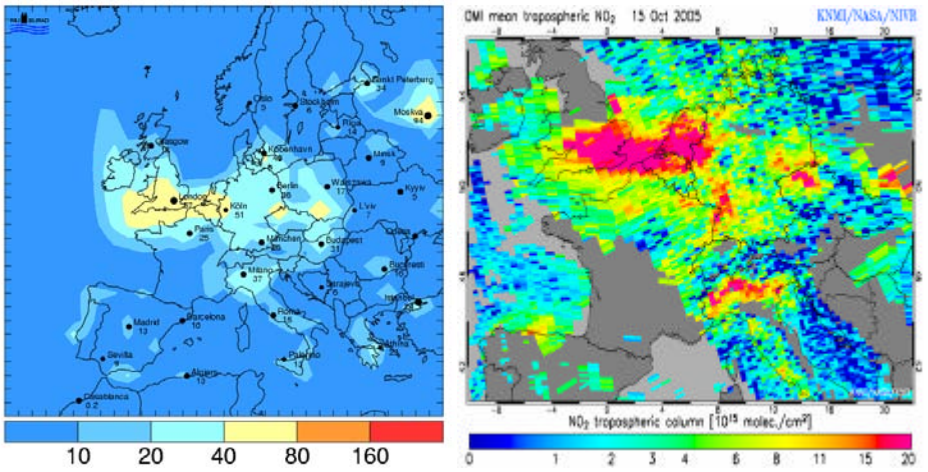


Figure 3. AQ Forecast: near surface maximum 24h mean of NO_2 in µg/m³ for domain D01 (left) and NO_2 column density in 10^{15} molec/cm³ derived from OMI satellite 2005 (right) at 15 October 2005. Note the different size of the domains.

4.2. Ozone production

The summer 2003 in Europe was characterized by extreme temperatures and dry episodes and thus accompanied by high ozone values exceeding critical values. Thus was especially true for Germany, where the information alert (180 µg/m³) and even the alarm threshold (240 µg/m³) were exceeded in many parts of the country. This Ozone values in the first half of August 2003 were well predicted over Germany (see Figure 4). The predicted values are in good agreement with observations (UBA monitoring network).

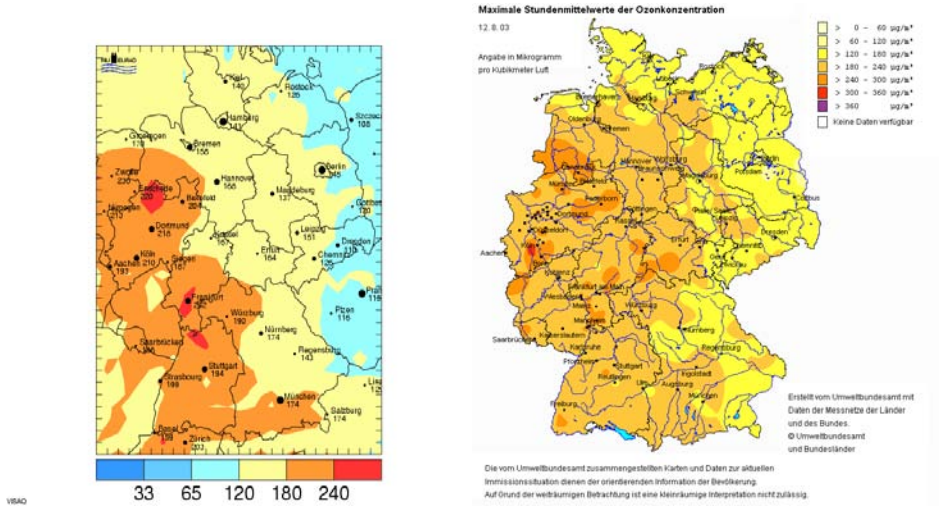


Figure 4. AQ Forecast: near surface daily maximum of Ozone in µg/m³ for Germany (left) and Ozone measured values from UBA monitoring network (right) at 12 August

4.3. Validation

Besides these examples of episodical validation of the predicted atmospheric pollutants a daily full validation of the products is evident and demonstrated on the RIU forecast website. The predictions are compared with over 300 measurement stations of the German UBA network: The observational data are assimilated with the forecast data as first guess and results in a daily analysis of the observed data. In addition scatter diagrams are shown to demonstrate the performance of the model. Figure 5 shows exemplarily the diagrams for an arbitrary selected day in March 2008. The different colours indicate the forecast start (black: start same day, yellow: start 1 day before current day, magenta: start 2 days before current day). The calculated scores are: Bias, Root Mean Square (RMS), Hit rate within a 20% interval (HR-20%) and hit rate within a 50% interval (HR-50%).

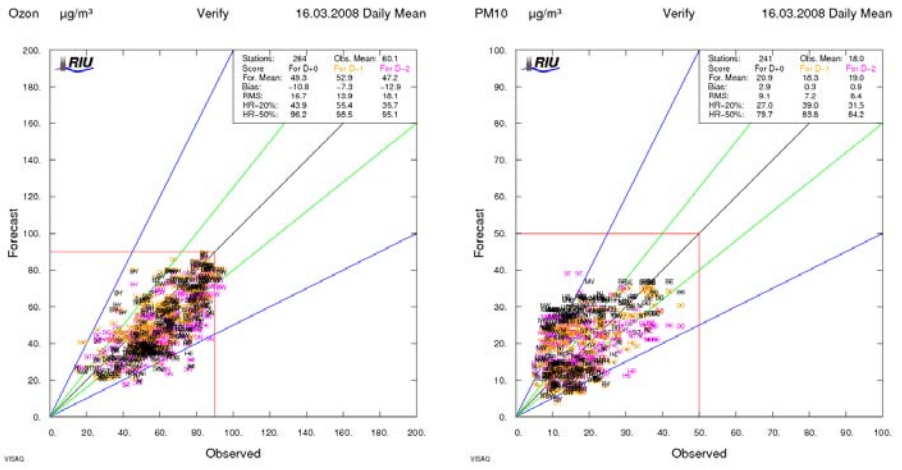


Figure 5. AQ Forecast: Forecasted and observed data of daily mean near surface concentrations of Ozone (left) and PM10 (right) in µg/m³ for 16 March 2008. The measured values are from UBA monitoring network.

Even this verification present only one day scatter data, again these arbitrary chosen results demonstrate the relatively good performance of the AQ forecast system.

5. CONCLUSIONS

The results of the operational EURAD AQ forecasts on different regions are published and are updated every day on the EURAD homepage (www.eurad.uni-koeln.de). Besides ozone, additional atmospheric pollutants as NO_2, SO_2, CO and PM10 were predicted every day for a 72 hour cycle. For assessment studies the ranges for the concentration thresholds were selected according to the EU directives. The verification of the forecast system is shown for selected events. The results of this verification demonstrate that the EURAD air quality forecast system performs reasonably well in predicting the spatial distribution of the main air pollutant concentrations over Europe and Germany and thus it could be used as a tool to support informational strategies for the public. The results of the prediction are also linked to several internet weather portals for distribution to the public. In addition, the predictions are used for assessment and policy making by several regional and local environmental agencies: The forecast maps are linked to the websites of the agencies, information of the public via radio or television, when exceedances of limit values are predicted.

An intensive verification of the forecast data will be performed for the year 2007. Another future task will be the incorporation of observational data within data assimilation techniques to improve the forecast results. In addition a intensive model output statistics will be outlined in order to remove typical prediction errors under certain weather conditions.

ACKNOWLEDGEMENTS

The authors wish to thank the following institutions for their support: European Space Agency (ESA), European commission, Regional Environmental Agency Northrhine-Westfalia (LANUV).

REFERENCES

Anthes, R. A., Warner, T. T., 1978. Development of hydrodynamic models suitable for air pollution and other meso-meteorological studies. Mon. Weather Rev. 106, 1045-1078.

Ebel, A., H. Elbern, H. Feldmann, H.J. Jakobs, C. Kessler, M. Memmesheimer, A. Oberreuter, G. Piekorz, 1997. Air Pollution Studies with the EURAD Model System (3):EURAD - European Air Pollution Dispersion Model System. Mitteilungen aus dem Institut für Geophysik und Meteorologie, Universität zu Köln, Nr. 120.

Grell, G. A., J. Duddhia, and D. Staufer, 1994. A description of the fifth-generation Penn State/Ncar meso-scale model (MM5), Technical note NCAR/TN-398+STR, National Center of Atmospheric Research, Boulder.

Jakobs, H.J., S. Tilmes, A. Heidegger, K. Nester and G. Smiatek, 2002. Short-term ozone forecasting with a network model system during Summer 1999. *Journal of Atmos. Chem.* 42, 23-40.

Memmesheimer, M., J. Tippke, A. Ebel, H. Hass, H. J. Jakobs, and M. Laube, 1991. On the use of EMEP emission inventories for European scale air pollution modelling with the EURAD model. In EMEP Workshop on Photo-oxidant Modelling for Long-Range Transport in Relation to Abatement Strategies, pp. 307 – 324, Berlin, April 1991.

Memmesheimer, M., Roemer, M., and Ebel, A., 1997. Budget calculations for ozone and its precursers: seasonal and episodic features based on model simulations, J. Atmos. Chem. 28:283–317.

Memmesheimer, M., E. Friese, A. Ebel, H. J. Jakobs, H. Feldmann, C. Kessler, and G. Piekorz, 2004. Long-term simulations of particulate matter in Europe on different scales using sequential nesting of a regional model, Int. J. Environ. Poll., 11(1 / 2), 108-132.

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 "Demokritos" web-based air quality forecasting system for the Greater Athens Area

Vlachogiannis D. [(a)], Sfetsos A. [(a)], Andronopoulos S. [(a)], Gounaris N. [(a)], Yiotis A. [(a)], Stubos A.K. [(a)]

(a) Environmental Research Laboratory, Institute of Nuclear Technology - Radiation Protection, NCSR "DEMOKRITOS", 15310 Aghia Paraskevi Attikis, Athens, Greece
(mandy@ipta.demokritos.gr)

Abstract:
The present paper describes the operational air quality forecasting web-based system currently under development by the Environmental Research Laboratory of NCSR "Demokritos". The system integrates three major computational modules and it is configured to apply on the Greater Athens Area, producing meteorological and air quality predictions on a high spatial resolution (1 x 1 km^2) and for a 72-hours time horizon with 1 - hour time step in advance. A case study, as an example of the forecasting ability of the developed system, is provided for February 17 and April 8, 2008. The first date was chosen as it was an extremely cold day for the area with extensive snowfall and varying emission sources. The second date was selected randomly as a business as usual scenario.

Keywords: meteorology, air quality, high resolution forecasting web-based system.

1. INTRODUCTION

Environmental informatics presently is merely the way to link meteorology and pollution sources with personal exposure and determine the impact on human health. More importantly, it is a decision aiding tool for local authorities to predict potential atmospheric pollution problems, optimize actions and policy making activities so as to produce the maximum health benefit. Under current legislation, (EU Directives 90/313/EEC, later amended by 2003/4/EC), the right to access environmental data improved the public's insight into environmental information and obliged policy making organisations to heavily invest in environmental planning and enact a broad spectrum of legislative measures.

In every major European city there is some kind of network system for monitoring and mapping the distribution of air pollution and human exposure. These systems generally encompass in-situ air quality monitoring stations, emission inventories, meteorological and air quality modelling, air quality mapping and air quality impact assessment of various control strategies in support of evaluation of action plans. Examples of decision support systems used in major European cities are the IMPAQT (Integrated Modular Program for Air Quality Tools, UK), the Austrian AirWare, the Norwegian AirQUIS, the Swedish EnviMan. The Finnish AQM system is based on integrated dispersion and exposure modelling system allowing for estimation of the spatial and temporal distributions of concentrations and resulting population exposures. In particular, currently for the Greater Athens Area (GAA) there exists one prognostic system for gaseous air compounds only (http://forecast.uoa.gr). At country level, for the whole domain of Greece, there is a forecasting system for gaseous pollutants on a coarse resolution of 10x10 km^2 (hhtp://lap.physics.auth.gr/forecasting).

As most of the intensely populated European cities hold a decision support tool for forecasting air quality levels and human exposure, it was deemed necessary to develop such

a tool for the prognosis and assessment of air pollution in the GAA, a region of complex topography particularly susceptible to accumulation of air pollutants. The present study is concerned with the operation of a flexible, web-based air quality forecasting system for GAA, currently under development by the Environmental Research Laboratory (EREL) of the National Centre for Scientific Research "Demokritos". The system comprises three major computational modules and is designed to produce meteorological and air quality forecasting on a high spatial resolution (1 x 1 km^2) and for a 72-hours time horizon with 1 hour time step in advance (Figure 1).

The meteorological predictions are produced by the MM5 (Mesoscale Model 5) (Penn State University version 3.7.2), which has been parameterized for application to the particular geographical and climatic characteristics of the Greater Athens Area. Its output is used to produce high resolution daily air emissions inventories for the main anthropogenic and biogenic pollutants with 1-hour time step by an in-house built processor named EMIS-LAB. The meteorological prediction fields in combination with the emissions inventories are used as inputs to the 3-D atmospheric dispersion and photochemical model CMAQ (Community Multiscale Air Quality, v.4.6), which produces prognoses for concentrations of gaseous and particulate pollutants. CMAQ can be employed in particularly when facing problems that have to do with complex emission patterns, source apportionment, as well as atmospheric processes which transport and transform pollutants in a dynamic environment. The analysis of the modelling results is performed via custom made MATLAB© post-processors and GIS (Geographical Information System) environment in an attempt to estimate air quality indices and compare them to the current EC and national regulatory limits.

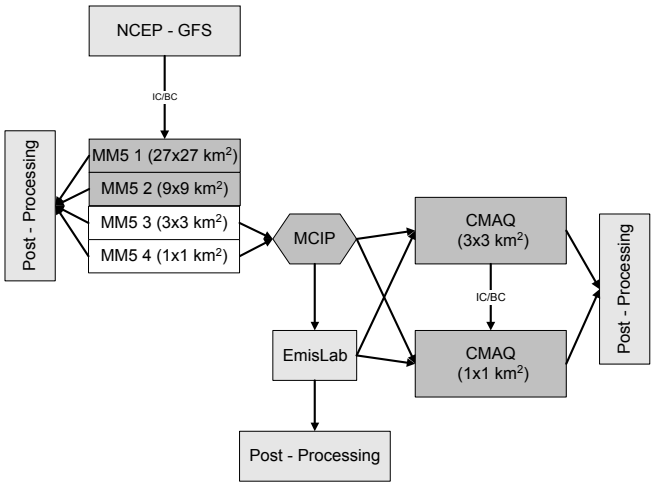


Figure 1. Schematic presentation of the system design.

MM5 and CMAQ models as a forecasting suite have been used extensively over the recent years in the United States, e.g. [Arnold et al., 2003]. However, there have been very few applications of this system for cities outside the USA, for example in the UK and Spain, e.g. [San Jose et al., 2002] and [Sokhi et al., 2006]) and China e.g. [Zhang et al., 2004]). This is the first time this particular set of models (MM5 and CMAQ) is applied for high spatial and temporal resolution air quality prognoses in Greece. In addition, EMISLab provides emission estimates amongst others for fine and coarse particulate matter on a 1 x 1 km^2 domain. Hence, the originality of this work is established by the online prognosis of both categories of air pollutants, gaseous and particulate matter, on a high spatial and temporal resolution.

2. HARDWARE SYSTEM DESIGN

The forecasting model is solved using a distributed memory Beowulf cluster. The cluster consists of an Intel Core 2 Duo@2.66Ghz PC with 2GB RAM, that serves as a frontend,

and 4 Intel Core 2 Duo @3Ghz compute nodes that solve the computational problem. The PC's are interconnected using a private Gigabit Ethernet Switch through the TCP/IP protocol. The theoretical performance of this cluster is 48Gflops and the actual measured Linpack performance is approximately 20Gflops.

3. DOMAIN DEFINITION

The MM5 model has been configured with four nested domains, as shown in Figure 2. A description of the domains with regards to the number of cells and discretisation information is provided on Table 1. The outer domain is sufficiently large to take as input meteorological data from the Global Forecast System (GFS – NCEP) in order to provide more accurate lateral boundary conditions. The area of application of the photochemical model CMAQ coincides with the inner domain of MM5 and covers a rather large section of Attica. In the vertical, there are 23 half-sigma layers.

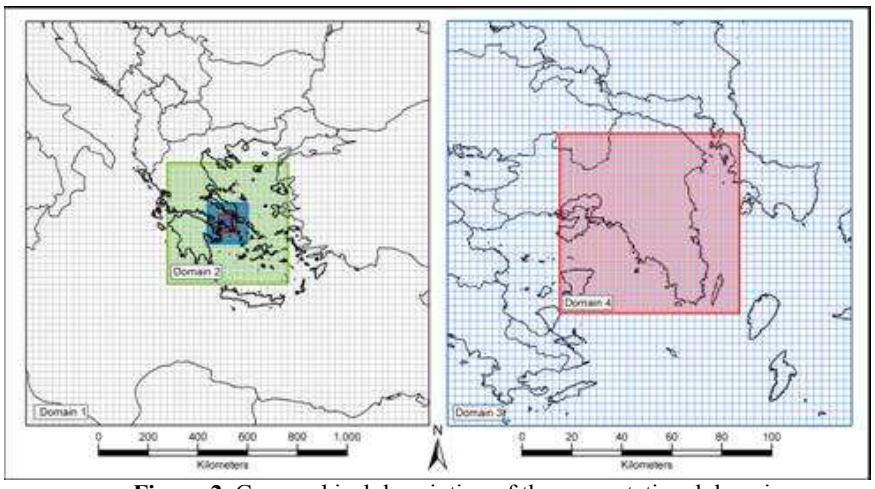


Figure 2. Geographical description of the computational domains.

Table 1. Description of computational domains.

Domains	Number of cells	Horizontal resolution (km)
Domain 1 (outermost)	60 x 60	27 x 27
Domain 2	54 x 54	9 x 9
Domain 3	54 x 54	3 x 3
Domain 4 (innermost)	72 x 72	1 x 1

4. THE MODELLING COMPONENTS

4.1. The Meteorological Forecast

The MM5 mesoscale model has been developed by the Pennsylvania State University (PSU) and the National Centre for Atmospheric Research (NCAR). It is a non-hydrostatic meteorological model and it can be used in cases where high spatial resolution (of a few kilometres) is demanded, as in our case study. MM5 is widely applied for operational forecasting. It has been extensively tested in cases of extreme weather events and over areas of complex topography. As previously mentioned, the model uses input from GFS – NCEP for initial and boundary conditions. In our case study, MM5 (version 3.7.2) has been suitably parameterized for producing the meteorological forecast in the working domain. In Figure 3 the topography and land-cover of the second innermost computational domain is presented. Four two-way nesting domains have been employed to downscale the initial conditions from the NCEP GFS model with cell sizes of 27 – 9 – 3 – 1 km, respectively (Table 1, Figure 2). Thus, a very high resolution weather forecast (1 km^2) is produced in the

innermost grid that covers the GAA. The generated meteorological data are used as input to CMAQ.

4.2. The Air-Emissions Model

The in-house EMIS-LAB emission processing tool is used to calculate emissions of air pollutants with a spatial and temporal resolution as required by the atmospheric chemistry model. EMIS-LAB has been developed to process existing inventory data producing an emission inventory of air pollutants with high resolution. It is designed to convert emission inventories that are of annual-total or daily-average temporal character, to any temporal time scale, typically hourly, required by the atmospheric quality models. Additionally, spatial processing is performed calculating emissions values for each model grid cell and model layer for selected sources (e.g. point and aviation).

EMIS-LAB is an integrated dynamic modelling system based on MATLAB. The calculation of the biogenic emissions is depended on the component BEIS 3.12 (EPA). Emissions from biogenic sources, fugitive dust and anthropogenic activities are treated, incorporating among others the influence of meteorological conditions that are produced from Numerical Weather Prediction models (in this case MM5). Other functionalities of EMIS-LAB include the extensive visualization capabilities, both spatial and in time series mode, allowing the user to obtain a complete overview of the modeling output, facilitating the analysis and assessment of the results. In addition, EMIS-LAB provides correct input format to CMAQ, including the Input/Output Applications Programming Interface (I/O API) and the Network Common Data Form (NetCDF) output format.

EMIS-LAB can process a large number of air pollutants, including toxic gases that are defined by the users. However, its intended application in regional air quality application requires the handling of critical gaseous pollutants such as carbon monoxide (CO), nitrogen oxides (NOx), volatile organic compounds (VOC), ammonia (NH3), sulfur dioxide (SO2) and particulate matter (PM) of different sizes, PM 2.5 and PM10.

4.3. The Air-quality forecast

CMAQ is a complex three-dimensional, (3D), Eulerian numerical air quality model. It is designed as a state-of-the-science ''one atmosphere'' air quality model involving complex atmospheric pollutant interactions and predicts ambient pollutant concentrations at various scales (regional and urban) in the low atmosphere. The CMAQ system can simulate concentrations of tropospheric ozone, acid deposition, visibility, fine particulate and other air pollutants. The CMAQ modelling system requires 4D meteorological data which are generally provided by a mesoscale meteorological model (in this case MM5). The one-way coupling of MM5 to CMAQ is accomplished through a meteorology–chemistry interface processor MCIP (Meteorology Chemistry Interface Processor) [Byun and Ching, 1999]. The adjoint variation of the CMAQ 4.5 model has been parameterized and applied for the two innermost domains (3x3 km^2 and 1x1 km^2, see Table 1) in nested mode. The chemical mechanism incorporated in the system is CB-IV including chemical reactions in the gaseous and liquid phase.

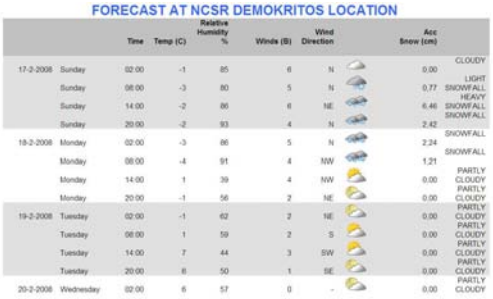
5. CASE STUDY APPLICATION

For the purpose of system testing, a case study for two selected dates February 17 and April 8, 2008 has been simulated. The first date was a particularly cold winter day for Greece, during which the snowfall that occurred was amongst the heaviest snowfalls in the last decade according to the available historic weather reports. The second date was chosen to simulate a business as usual scenario as regards to meteorological and emission conditions.

5.1. Meteorology

The selected parameterization on MM5 was the Grell cumulus scheme, the MRF PBL scheme based on Troen-Mahrt representation, the CCTM2 radiation model, and the simple ice explicit moisture scheme. The number of vertical layers in the modelling study was set to 23, with a telescopic increase with height. The initial data were obtained from the GFS model at the 00:00 run.

The following series of figures yield the various calculated fields of the modelling system (Figure 3 and Figure 4) and present a preliminary comparison of the MM5 output at different domains against experimental observations. Figure 3 shows an MM5 text output referring to the weather forecast of the NCSRD location in Aghia-Paraskevi, for the period from February 17 (Sunday) to February 20, 2008. The model correctly predicted the commencing of the snow fall that started on Sunday midday and lasted until Monday morning. The predicted wind field can be sketched in vector form, an example is shown in Figure 4, from the 3rd modelling domain (3×3 km^2), for the 8th of April on 12:00 UTC.


FORECAST AT NCSR DEMOKRITOS LOCATION

		Time	Temp (C)	Relative Humidity %	Winds (B)	Wind Direction		Acc Snow (cm)	
17-2-2008	Sunday	02:00	-1	85	6	N		0.00	CLOUDY
	Sunday	08:00	-3	80	5	N		0.77	LIGHT SNOWFALL
	Sunday	14:00	-2	86	6	NE		6.46	HEAVY SNOWFALL
	Sunday	20:00	-2	93	4	N		2.42	SNOWFALL
18-2-2008	Monday	02:00	-3	86	5	N		2.24	SNOWFALL
	Monday	08:00	-4	91	4	NW		1.21	SNOWFALL
	Monday	14:00	1	39	4	NW		0.00	PARTLY CLOUDY
	Monday	20:00	-1	56	2	NE		0.00	PARTLY CLOUDY
19-2-2008	Tuesday	02:00	-1	62	2	NE		0.00	PARTLY CLOUDY
	Tuesday	08:00	1	59	2	S		0.00	PARTLY CLOUDY
	Tuesday	14:00	7	44	3	SW		0.00	PARTLY CLOUDY
	Tuesday	20:00	6	50	1	SE		0.00	PARTLY CLOUDY
20-2-2008	Wednesday	02:00	6	57	0	-		0.00	PARTLY CLOUDY

Figure 3. Weather predictions in text format.

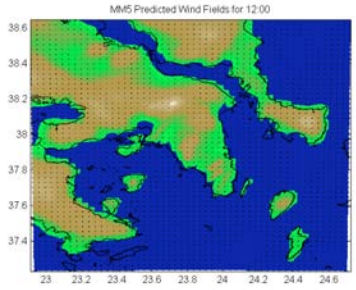


Figure 4. Predicted wind field for April 8.

A preliminary evaluation of the model predictions is presented in the following figures. The model output is able to capture accurately the vertical profiles of temperature and relative humidity as it can be deduced from Figure 5. On the 8th of April the model seems to overpredict relative humidity on the highest layers. Concerning the surface wind speed, the comparison was performed against observational data from the Athens International Airport (code LGAV) recorded every hour for that day. The data used in this validation study were obtained from the monitoring network of the Hellenic National Meteorological Service.

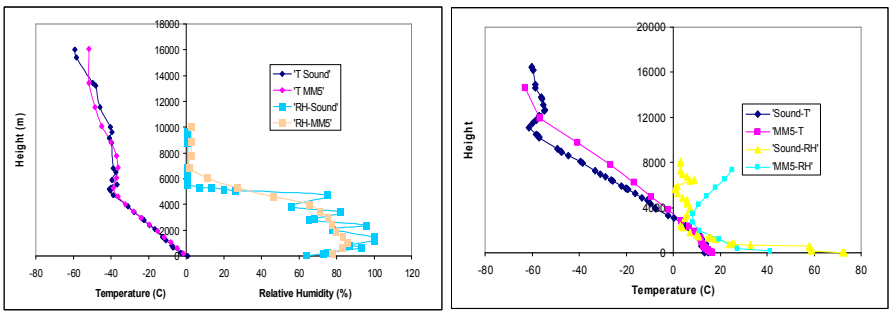


Figure 5. Comparison of vertical profiles of temperature and relative humidity.

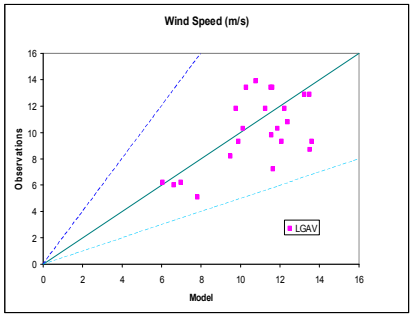


Figure 6. Scatter plot of observed and modelled surface wind at LGAV.

5.2. Emissions

The MM5 output was further processed using MCIP to provide the meteorological fields for the emission processing system EMIS-LAB. EMIS-LAB has the ability to rapidly adapt to the particular conditions of that day. On the 17th and 18th of February the entire Greater Athens Area experienced particularly difficult meteorological conditions, which dramatically affected the economic and transportation activity on these specific days. The Athens International Airport was closed during the entire day and shipping was cancelled due to very strong winds. Therefore, the emissions consisted primarily of: (i) the main industries that were treated as point sources, (ii) road-traffic and rail, as the remaining transportation sources, (iii) domestic heating and (iv) biogenic emissions. Each of the above mentioned sources was processed using a suitable and unique temporal (hourly) and chemical profile based on previous studies (Winiwarter et al., 2003) and existing databases (e.g. EPA SPECIATE, SMOKE) so that finally gridded emissions were generated. Moreover, road-traffic was treated taking into account the highly reduced number of vehicles and vehicle speeds for the particular dates. An example is presented in Figure 7, at 11:00 on the 17th February and at the lowest model layer. Figure 8 presents the PM2.5 emissions from nonroad mobile sources and CO emissions from all sources.

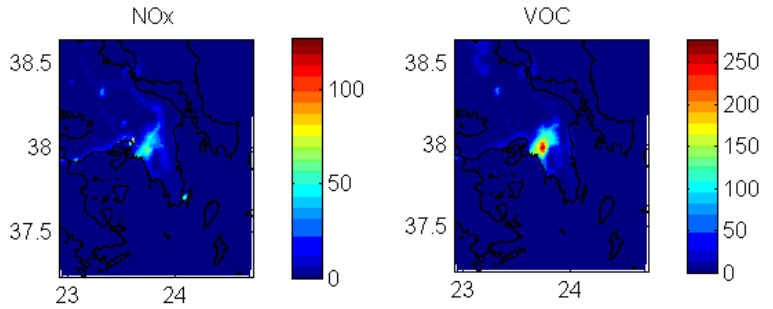


Figure 7. Hourly emissions estimate at 12:00, February 17, 2008.

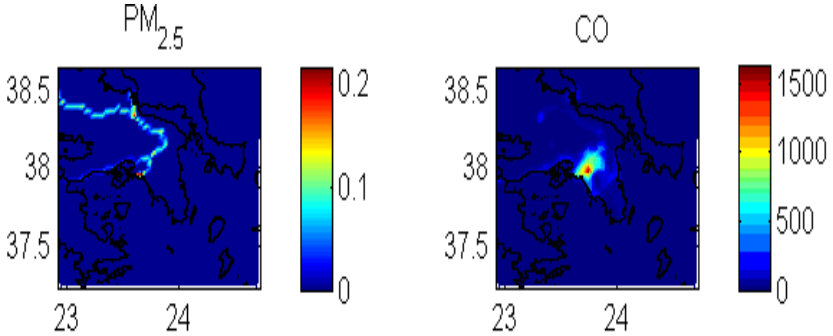


Figure 8. Hourly emissions estimate at 12:00, April 8, 2008.

5.3. Air quality

The MCIP processed output of MM5 on the two finer domains was used as input to the CMAQ in addition to the generated emission file in NETCDF format. As mentioned in Section 2, initially CMAQ was run for DOMAIN 3 and its output was used as initial and boundary conditions to the second run performed in the working DOMAIN 4. Initial and boundary conditions (IC/BC) were also derived from typical profiles of the gaseous species CO, O3, NO2 and SO2 provided with the CMAQ source code, modified to reflect the situation in the area of NCSRD, using background observational data. The CB-IV (Carbon Bond IV) chemical mechanism with aerosols and aqueous chemistry was chosen in addition to the Euler Backward Iterative (EBI) for temporal discretization.

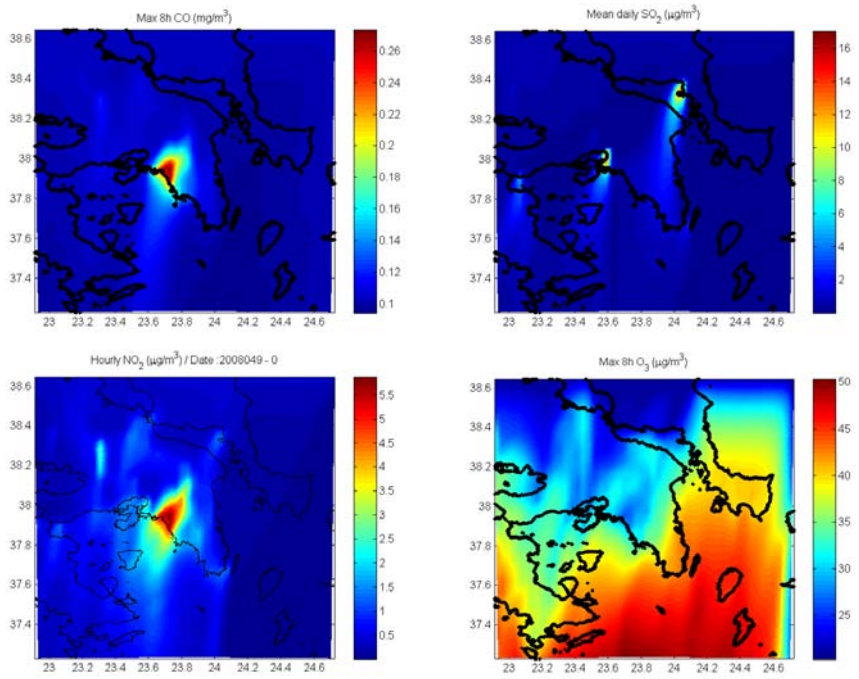


Figure 8. CMAQ output shown for February 17, 2008 in accordance to the national air quality limits set by the Greek Ministry of Environment.

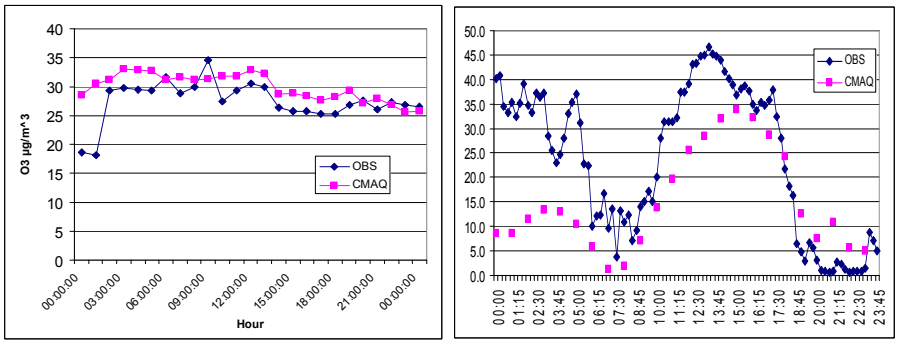


Figure 9. Comparison of model calculated results for O3 with observations at NCSRD station (a) Feb 17, (b) Apr 8.

Figure 8 presents the output of CMAQ for the 17th of February in accordance to the national air quality limits set by the Ministry of Environment. Due to the prevailing meteorological conditions it can be observed that none of the priority pollutants that were examined appear to exceed the national limits. Ozone does not appear to be generated by photochemical reactions due to heavy cloud coverage. Figure 9 depicts the hourly calculated O3 concentrations and the observed values of the pollutants from the NCSRD air quality monitoring station for both days examined. The results show that a relatively good agreement is reached, although further and more thorough model validation is currently under way. Modelling discrepancies in the early hours of the 8th of April are attributed to the initial conditions of the run.

6. Conclusions

The NCSRD air quality forecasting system has been developed by the Environmental Research Laboratory aiming potentially at the assessment of air pollution in the Greater Athens Area. The system is an integrated modelling framework comprising the: (i) MM5 meteorological model, resulting in a very fine resolved domain for Attica (1 x 1 km^2), (ii) EMISLAB, a dynamic anthropogenic and biogenic emission processing model, (iii) CMAQ, which is the state-of-the-art photochemical model and (iv) extensive post processing options for visualization and more user friendly presentation of the results.
As a test case, the 17th of February 2008 was studied and a primary evaluation of the modelling components against observational data from the existing monitoring networks was presented. This work represents the first application of the MM5-CMAQ system in Greece for predicting air pollution levels. Future work will include sensitivity studies of the air quality predictions to the emission inventories.

ACKNOWLEDGEMENTS

This work was partially funded by the Hellenic General Secretariat of Research & Technology (GSRT) and by the European Union - European Regional Development Fund (ERDF).

REFERENCES

Arnold, J.R., R.L Dennis,. G.S Tonnesen,. Diagnostic evaluation of numerical air quality models with specialized ambient observations: testing the Community Multiscale Air Quality modeling system (CMAQ) at selected SOS 95 ground sites, *Atmospheric Environment* 37, 1185–1198, 2003.

Byun, D.W. and J.K.S. Ching, (Eds.). Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. EPA Report N. EPA-600/R-99/030, Office of Research and Development. US Environmental Protection Agency, Washington, DC., 1999.

San Jose, R., J.L. Perez, J.F Blanco, R. Barquın, R.M. Gonzalez. An operational version of MM5–CMAQ modelling system over Madrid City. Forth Symposium on the Urban Environment, 20–24. American Meteorology Society, May 2002.

Sokhi R.S., R. San Jose, N. Kitwiroon, E. Fragkoua, J.L. Pe´rez, D.R. Middleton, Prediction of ozone levels in London using the MM5–CMAQ modelling system, *Environmental Modelling & Software* 21, 566–576, 2006.

Winiwarter W., C. Dore, G. Hayman, D. Vlachogiannis, N. Gounaris, J. G. Bartzis, S. Ekstrand, M. Tamponi and G. Meffeis. Methods for comparing gridded inventories of atmospheric emissions – applications for Milan Province, Italy and the Greater Athens Area, Greece, *The Science of the Total Environment,* 303, 231-243, 2003.

Zhang, M., I. Uno, Y. Yoshida, Y. Xu, Z. Wang, H. Akimoto, T. Bates, T. Quinn, A. Bandy, B. Blomquist. Transport and transformation of sulfur compounds over East Asia during the TRACE-P and ACE-Asia campaigns. *Atmospheric Environment* 38, 6947–6959, 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

The Use of Third Generation Air Quality Modeling Systems for Web Operational Real-time Forecasting Decision Support Systems: Spain Case

R. San Jose[a], J.L. Pérez[a], J.L. Morant[a], F. Prieto[c] and R.M. González[b]

[a]*Environmental Software and Modelling Group, Computer Science School, Technical University of Madrid (UPM), Campus de Montegancedo, Boadilla del Monte 28660 Madrid (Spain), (roberto@fi.upm.es).*

[b]*Department of Meteorology and Geophysics, Faculty of Physics, Complutense University of Madrid (UCM), 29040 Madrid (Spain), (barras@fis.ucm.es).*

[c]*Spanish Observatory for Sustainability, Universidad de Alcalá, plaza de San Diego, s/n 28801 Alcalá de Henares (Madrid), (fprieto21@gmail.com).*

Abstract: New generation of electric power plants and industrial plants are important emission sources generally surrounding large, medium and small cities. Future tendencies on energy production envision that the growing demand of electricity will push forward the need to construct more electric power plants with different combustion material. In Spain, the growing electric energy demand has conducted to a growing market on natural gas combined cycle electric power plants. The air quality impact of such a plant is considerable smaller than the old ones but the modern EU Air Quality legislation is starting to obligate to control the air quality impact in real-time and forecasting mode. This tool presents the first implementation in Spain over a 4 group combined cycle electric power plant, located in the surrounding area of Madrid Community (Spain) by using the MM5-CMAQ-EMIMO modeling tool. The MM5 model is widely used al over the world developed by PSU/NCAR (USA) and the CMAQ model is the so-called Community Multiscale Air Quality Modelling System developed by EPA (USA) and finally EMIMO is an anthropogenic and biogenic emission model to produce hourly emissions per pollutant per squared kilometer. The system is accessed over the Internet for the environmental authorities and company managers under daily basis. The tool produces alerts every day according to the results of the model. Final decision related to possible shut-down for a limited period of time of the different power plant groups – in case of exceeds of EU Directive limits due to the emissions of the different power groups – is taken by environmental authorities in real-time. The system prototype was part of the EUREKA project TEAP (A tool to evaluate the air quality impact on industrial plants) (2001-2003). The same approach can be used for any other industrial plant and also for any emission source apportionment such as traffic over specific sections of the model domain or even specific pollutants over determined areas in the model domain.

Keywords: Air Quality Modelling, Industrial impact, real-time control.

1. INTRODUCTION

The air quality impact of industrial plants is an essential issue in air quality assessment and modelling which is becoming more important due to the more strict EU legislation in air quality standards. The 2002/3/EC Directive of the European Parliament and of the Council

of 12 February 2002 related to ozone in ambient air provides information related to shirt-term action plans at the appropriate administrative levels. In accordance with this legislation, industrial plants are being requested by environmental authorities to have appropriate control systems to provide in real-time and forecasting mode information related to the relative impact of their industrial plants in the forecasted pollution levels – in particular on the O3, SO2, NOx, CO and PM10 levels which are limited by different EU Directives -. The capability to reduce specific emissions in real-time according to a forecast for a specific area and period of time is actually a challenging issue. In the past years, this objective was limited due to the limited computer power and the cost of vector parallel computers. Nowadays, cluster system composed by PC processors (3,4 Ghz or 3,6 Ghz) provide an acceptable capability – once we have designed a proper architecture of the air quality modelling systems – to run this complex systems in real-time and forecasting mode.

The concept of real-time in our case is related to the fact of taking appropriate decisions in advance to avoid specific exceeds of the EU Directive limits. Following above Directive the responsibility to design of short-term action plans, including trigger levels for specific actions, is the responsibility of Member States. Depending of the individual case, the plans may provide for graduated, cost-effective measures to control and, where necessary, reduce or suspend certain activities, including motor vehicle traffic, which contribute to emissions which result in the alert threshold being exceeded. These may also include effective measures in relation to the use of industrial plants or products. In this application we focus on the possible reduction of industrial activities – in our case, a combined cycle power plant -.

The complete tool designed for this application is called TEAP (a Tool to Evaluate the Air quality impact of industrial Plants) [San José et al., 1994, 1996]. This tool is designed to be used by the environmental impact department at the industrial site. The tool provides a response to air quality impact to industrial emissions in the form of surface patterns and lineal time series for specific geographical locations into the model domain. The model domain is designed in a way that the industrial source point is located approximately in the centre of the model domain. The model domain can be as large as wished but a specific nesting architecture should be designed for each case together with balanced computer architecture.

The TEAP tool (an EUREKA-EU project) has the capability to incorporate different modelling systems. In a preliminary stage we have tested thee system with the so-called OPANA model (ETC/ACC03). OPANA model [San José et al., 1996] stands for Operational Atmospheric Numerical pollution model for urban and regional areas and was developed at the middle of the 90's by the Environmental Software and modelling Group at the Computer Science School of the Technical University of Madrid (UPM) based on the MEMO model (ETC/ACC03) developed in the University of Karlsruhe (Germany) in 1989 and updated on 1995, for non-hydrostatic three dimensional mesoscale meteorological modelling and SMVGEAR model for chemistry transformations based on the CBM-IV mechanism and the GEAR implicit numerical technique developed at University of Los Angeles (USA) in 1994. The OPANA model has been used (different versions) for simulating the atmospheric flow – and the pollutant concentrations – over cities and regions in different EU funded projects such as EMMA (1996-1998), EQUAL (1998 – 2001), APNEE (2000-2001). In these cases and others the model has become an operational tool for several cities such as Leicester (United Kingdom), Bilbao (Spain), Madrid (Spain), Asturias region (North of Spain) and Quito (Ecuador, BID, 2000). In all these cases the model continue to operate under daily basis and simulates the atmospheric flow in a three dimensional framework. The OPANA model, however, is a limited area model – which means that the model domain is limited by the earth curvature – and the cloud chemistry and particulate matter is not included (aerosol and aqueous chemistry).

Examples of "state-of-the-art" meteorological models are: MM5 (PSU/NCAR, USA), RSM (NOAA, USA), ECMWF (Redding, U.K.), HIRLAM (Finnish Meteorological Institute, Finland), etc. Examples of "state-of-the-art" of transport/chemistry models – also called "third generation of air quality modelling systems" – are: EURAD (University of Cologne, Germany), [Stockwell et al., 1977], EUROS (RIVM, The Netherlands), [Lagner et al., 1998], EMEP Eulerian (DNMI, Oslo, Norway), MATCH (SMHI, Norrkoping, Sweden),

[Dervent R. and Jenkin M. 1991], REM3 (Free University of Berlin, Germany), (2000), CHIMERE (ISPL, Paris, France), [Schmidt et al., 2001], NILU-CTM (NILU, Kjeller, Norway), [Gardner et al., 1997], LOTOS (TNO, Apeldoorm, The Netherlands), [Roemer et al., 1996], DEM (NERI, Roskilde, Denmark), [Gery et al., 1989], STOCHEM (UK Met. Office, Bracknell, U.K.), [Collins et al., 1997]. In USA, CAMx Environ Inc., STEM-III (University of Iowa) and CMAQ (EPA, US) are the most up-to-date air quality dispersion chemical models. In this application we have used the CMAQ model (EPA, U.S.) which is one of the most complete models and includes aerosol, cloud and aerosol chemistry.

2. THE MM5-CMAQ MODELING SYSTEM

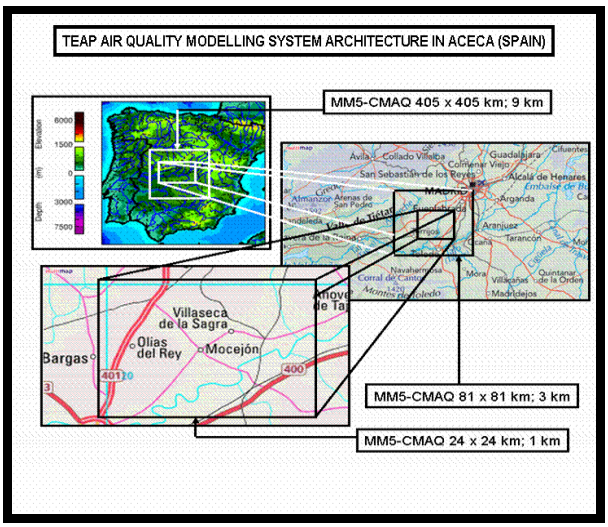


Figure 1. Model domain architecture.

The CMAQ model (Community Multi-scale Air Quality Modeling System, EPA, US) is implemented in a consistent and balanced way with the MM5 model. The CMAQ model is fixed "into" the MM5 model with the same grid resolution (6 MM5 grid cells are used at the boundaries for CMAQ boundary conditions). As an example a domain architecture is showed in Figure 1 for an application in a combined cycle power plant in the south area of Madrid Community . MM5 is linked to CMAQ by using the MCIP module which is providing the physical variables for running the dispersion/chemical module (CMAQ) such as boundary layer height, turbulent fluxes (momentum, latent and sensible heat), boundary layer turbulent stratification (Monin-Obukhov length), friction velocity, scale temperature, etc. We have run the modeling system (MM5-CMAQ) with USGS 1 km landuse data and GTOPO 30'' for the Digital Elevation Model (DEM) which can be substituted for any more accurate high spatial resolution landuse information in the implementation of the input data.

The system uses EMIMO model to produce every hour and every 1 km grid cell the emissions of total VOC's (including biogenic), SO2, NOx and CO. EMIMO is a emission model developed at our laboratory in 2001. This model uses global emission data from EMEP/CORINAIR European emission inventory (50 km spatial resolution) and EDGAR global emission inventory (RIVM, The Netherlands). In addition the EMIMO (EMIssion Model) model uses data from DCW (Digital Chart of the World) and USGS land-use data from AVHRR/NOAA 1 km satellite information. The EMIMO model includes a biogenic module (BIOEMI) developed also in our laboratory based on the algorithms for natural NOx, monoterpene and isoprene emissions in function of LAI (leaf Area Index) and PAR (photosynthetic active radiation). The emission inventory is a model which provides in time and space the amount of a pollutant emitted to the atmosphere. In our case we should quantify the emissions due to traffic, domestic sources, industrial and tertiary sector and also the biogenic emissions in the three model domains with 9 km, 3 km and 1 km spatial resolution mentioned above.

The mathematical procedures to create an emission inventory are essentially two: a) Top-down and b) Bottom-up. In reality a nice combination of both approaches offers the best results. Because of the high non-linearity of the atmospheric system, due to the characteristics of the turbulent atmospheric flow, the only possibility to establish the impact of the part of the emissions (due to traffic or one specific industrial plant, for example) in

air concentrations, is to run the system several times, each time with a different emission scenario. The general scheme of application of a set of combined cycle power plants can be seen in Figure 2.

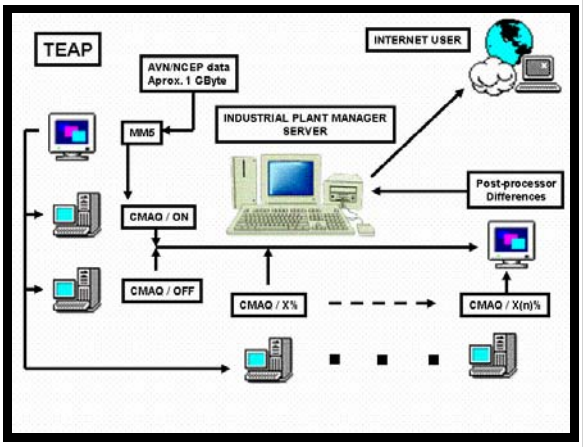


Figure 2. General scheme for full application with N different combined cycle power plants.

In this contribution we have applied the MM5-CMAQ modelling system over a power plat with 4 400 MW combined cycle power groups which are expected to operate simultaneously. The simultaneous simulations of the so-called ON run (the four groups running) and OFF1, OFF2, OFF3 and OFF4 runs – which are representing the air concentrations when switching off the first combined cycle power group, the second one and so on successively – is analyzed by generating the corresponding differences between ON-OFF1, ON-OFF2 and so on. These differences represent the respective impact of each of the successive switching off process until the OFF4 scenario which represent the complete disconnection of the 4 400 MW combined cycle power plant groups and the subsequent zero emissions from the power plant. The system was calibrated for 60 day selected periods by using 5 day periods per month along the 2004 year. Figure 3 shows the comparison between NO air concentration in Leganes monitoring station (located in the surrounding area of Madrid city) and the simulated NO concentrations.

Figure 4 shows the comparison between O3 hour observations and modeling results for several monitoring stations located in the surrounding area of Madrid. Note that the total amount of data is 8760 hours (100 % of the year's hours) and the correlation coefficient is 0,797.

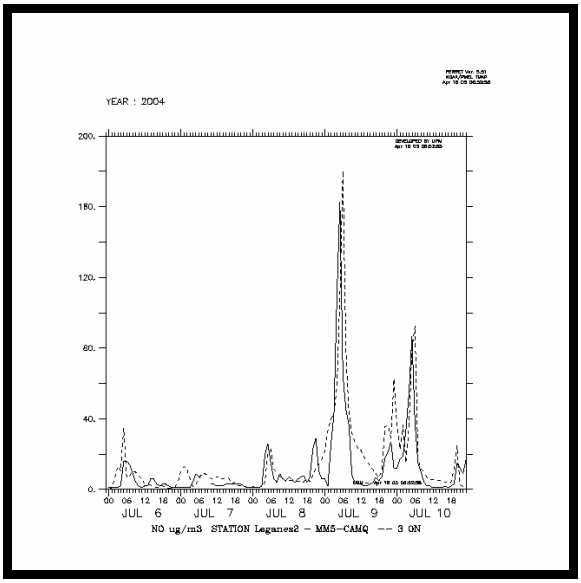


Figure 3. Comparison between NO observations and simulated data by using TEAP tool for analyzing the impact of emissions from combined cycle plants.

The impacts – due to the high chemical non linearity involved – are analyzed respecting on the absolute concentration pollution values and these values respecting the EU Directive limits. The post-processing is done automatically and presented in the specifically designed Web site.

The system is operating since July, 1, 2005 with full success. The developed system provides 72 hours forecasts for the impact of the 4 independent 400 MW combined cycle power groups. The access to the web site is restricted to environmental authorities and company authorized personnel. The system is operating from the Computer Center at Computer Science School at Technical University of Madrid (UPM) managed by the Environmental

Software and Modelling Group (ESMG). The system has been mounted over an 8-node 3,4 Ghz cluster platform. Figure 5 shows an example of the TEAP prototype mounted over a Petrochemical plant also in the South of Madrid Area (150 km away). In figure 6 we show the impact on O3 concentrations at a different location in the south of Madrid Community on September, 4, 2002, 15h00 GMT as obtained during the tests.

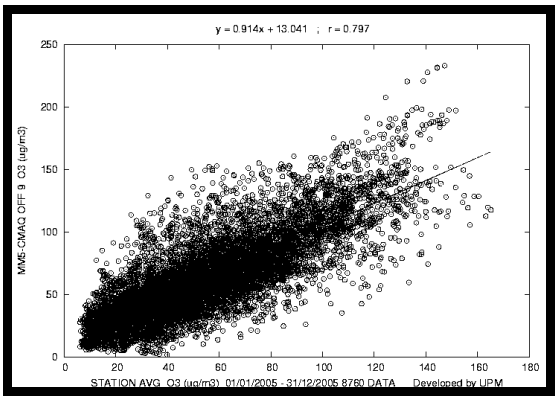


Figure 4. Comparison between O3 hour observations and modeling data for a group of monitoring stations located in the surrounding area of Madrid city. The presented vales are averaged between all stations and corresponding grid cells during the full 2005 year. The correlation coefficient shows an excellent value near 0,8.

The O3 concentrations are increased up to 11 % over the levels (without the emissions from the power plant) and with decreases up to 14 %. The increased levels are located at distances between 10 – 20 km to the East of the power plant location (centre of the domain) and the decreased levels are located in the immediate surrounding areas of the power plant (center of the domain). Figure 7 shows the O3 percentage change during 120 hours simulation and when activating the power plant after 72 hours. We observe that in that specific location (4336,1770 m, UTM) we expect increases and decreases on O3 concentration in a range up/down 6%. The TEAP tool provides a full time and spatial information related to the absolute and relative impact of different emissions (different combined cycle power groups) in real-time and forecasting mode under daily operations with an 8-node cluster platform. The system is the first operating in Spain by using such a sophisticated 3rd Generation Air Quality Modelling System and it is expected to be installed in several other combined cycle power plants and in general in different industrial plants to help the local and regional authorities to identify the relative impact of the different industrial plants located in the surrounding area. The system can be adapted to identify the impact of traffic sources and also different scenarios.

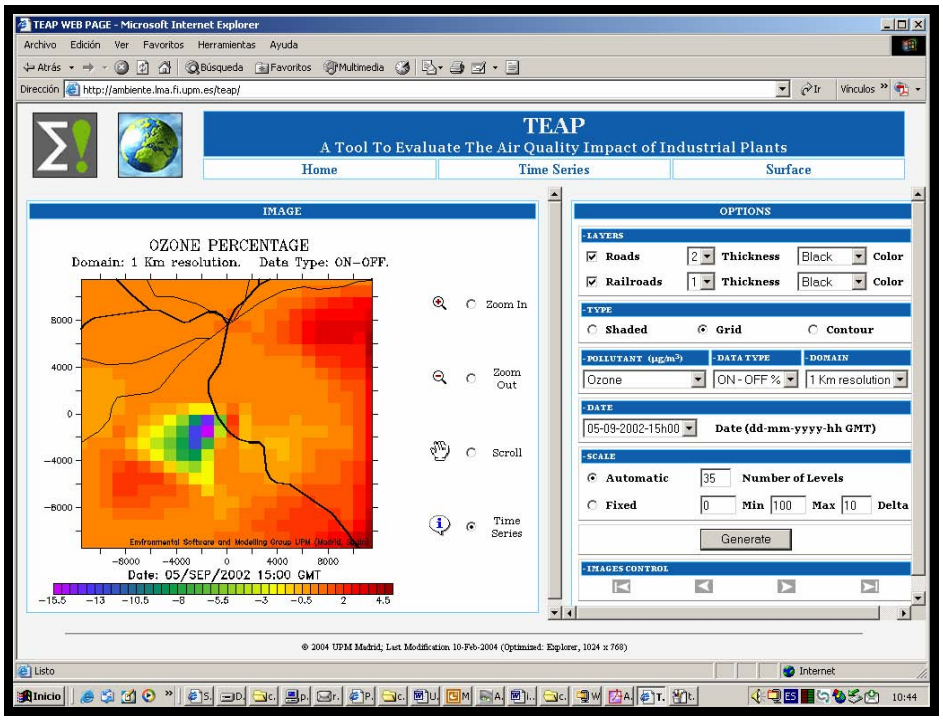


Figure 5. TEAP web site to analyze the impact of a petrochemical plant located 150 km at the south of Madrid area.

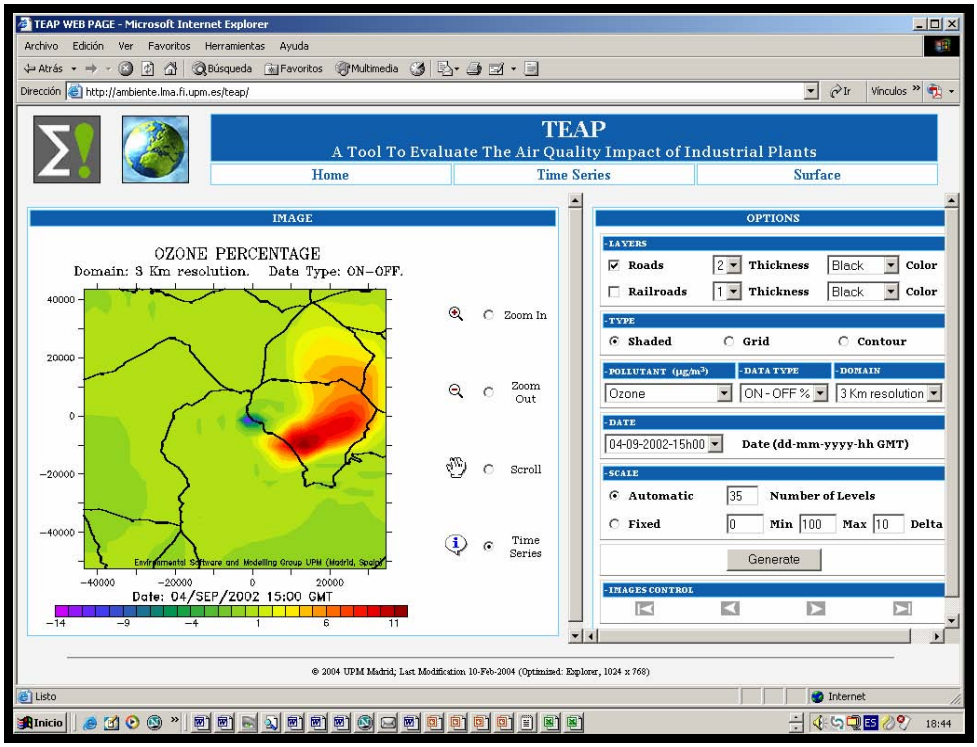


Figure 6. O3 power plant concentrations impact at 15h00 on September, 4, 2002 for the 3 km spatial resolution model domain showing the areas with increase and decrease on O3 concentrations due to the emissions from the power plant.

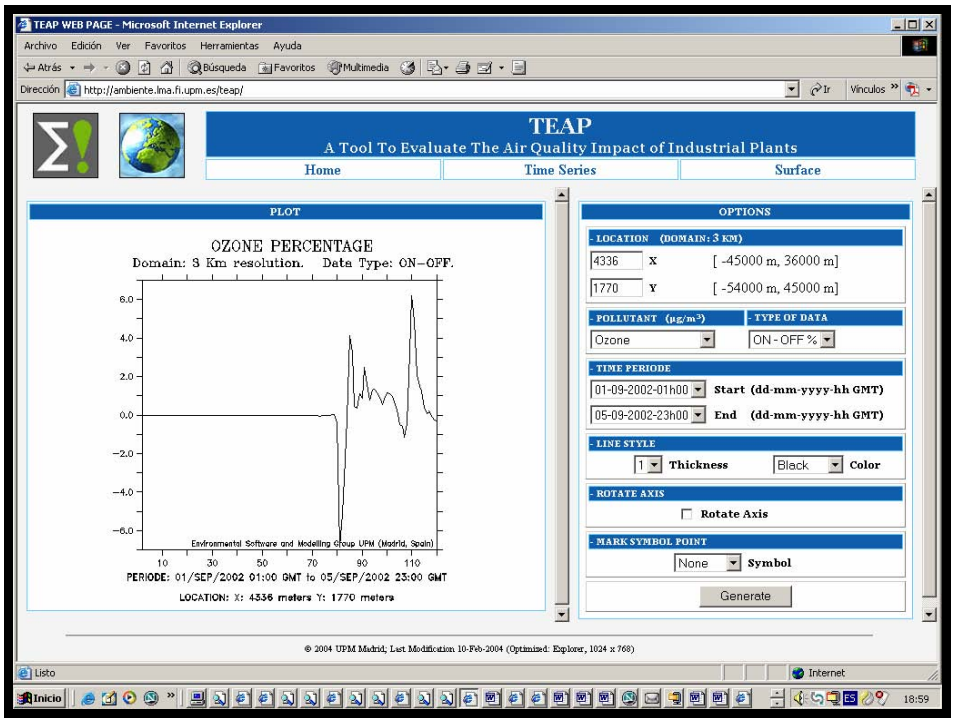


Figure 7. O3 concentration percentage impact due to the emissions of the power plant located in the Southern area of Madrid Community for the test period (1-5, September, 2002) at location (4336,1770, UTM).

3. CONCLUSIONS

The use of state-of-the-art (third generation of air quality models) air quality modelling systems such as MM5-CMAQ-EMIMO provides a robust capability to predict the air quality impact of industrial emissions in real-time mode. These models are usually complex codes which simulate all the atmospheric process and are capable to simulate the nonlinear chemical process in the atmosphere. Due to this non-linearity feature, the impact must be calculated by parallel runs in cluster parallel computer platforms. Each run simulates all the process in the atmosphere including all the emission sources but the different between one process and another is just the emissions of the industrial plant to be analyzed. The computer runs should be scheduled under daily basis (faster computer platforms can try several runs every day) with an horizon of 48-72 hours or more (depending on the computer capabilities). The process of switching on and off the industrial emission should reproduce the approved protocol in case of exceedances – according to EU limits - in the air quality concentrations predicted by the daily simulations. The impacts are calculated at the end of the runs by showing maps, statistics, time series, etc. of the absolute values and differences and percentages between different scenarios (ON and OFF). The TEAP system is a prototype developed during the TEAP Eureka Project and operational versions are applied in two different industrial complexes located in the surrounding areas of Madrid: ACECA power plant and Portland Valderrivas Cement company. Similar systems can be used and developed for the impact of roads and areas in cities and take actions in advance to avoid pollution concentration exceedances by applying specific emission reduction strategies according to the forecasts produced by a system TEAP-like.

ACKNOWLEDGEMENTS

We would like to thank PSU/NCAR and EPA (US) for the MM5 and CMAQ codes. The Spanish Industrial Funding Agency (CDTI) and the SOCOIN S.A. company.

REFERENCES

Collins W.J., Stevenson D.S., Johnson C.E. and Derwent R.G. (1997). Tropospheric ozone in a global scale3D Lagrangian model and its response to NOx emission controls. J. Atmos. Chem., 86, 223-274.

Derwent, R. and Jenkin, M. (1991). Hydrocarbons and the long range transport...Techniques for pollution monitoring in remote sites, II: Ozone. Water, Air, and Soil Pollution, 68, 391-397.

Gardner R.K., Adams K., Cook T., Deidewig F., Ernedal S., Falk R., Fleuti E., Herms E., Johnson C., Lecht M., Lee D., Leech M., Lister D., Masse B., Metcalfe M., Newton P., Schmidt A., Vandenberg C. and van Drimmelen R. (1997). The ANCAT/EC global inventory of NOx emissions from aircraft. Atmospheric Environment 31, 1751-1766.

Gery M.W., Whitten G.Z., Killus J.P. and Dodge M.C (1989), A photochemical kinetics mechanism for urban and regional scale computer modelling, Journal of Geophysical Research, 94, D10, pp. 12925-12956.

Langner J., Bergstrom R. and Pleijel K. (1998). European scale modeling of sulfur, oxidized nitrogen and photochemical oxidants. Model development and evaluation for the 1994 growing season. SMHI report RMK No. 82. Swedish Met. And Hydrol. Inst., SE-601 76 Norrkoping, Sweden.

Roemer M., Boersen G., Builtjes P. and Esser P. (1996). The budget of ozone and precursors over Europe calculated with the LOTOS model. TNO publication P96/004, Apeldoorn, The Netherlands.

San José R., Rodriguez L., Moreno J., Palacios M., Sanz M.A. and Delgado M. (1994) Eulerian and photochemical modelling over Madrid area in a mesoscale context, Air Pollution II, Vol. 1Computer Simulation, Computational Mechanics Publications, Ed. Baldasano, Brebbia, Power and Zannetti., pp. 209-217.

San José R., Cortés J., Moreno J., Prieto J.F. and González R.M. (1996) Ozone modelling over a large city by using a mesoscale Eulerian model: Madrid case study, Development and Application of Computer Techniques to Environmental Studies, Computational Mechanics Publications, Ed. Zannetti and Brebbia, pp. 309-319.

Schmidt H., Derognat C., Vautard R. and Beekmann M. (2001). A comparison of simulated and observed ozone mixing ratios for the summer 1998 in Western Europe. Atmospheric Environment, 35, 6277-6297.

Stockwell W., Kirchner F., Kuhn M. and Seefeld S. (1997). A new mechanism for regional atmospheric chemistry modeling. J. Geophys. Res., 102, 25847-25879.

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

Web-system for air quality assessment of the city area based on the mathematical modeling data

A. Titov[a,b], **D. Belikov**[c], **E. Gordov**[a,b] **and A. Starchenko** [c]

[a]*Siberian Center for Environmental Research and Training, Akademicheski ave., 10/3, Tomsk, Russia (titov@scert.ru)*
[b]*Institute of Monitoring of Climatic and Ecological Systems SB RAS, Akademicheski ave., 10/3, 634055, Tomsk, Russia*
[c] *Tomsk State University, 36 Lenin Prospekt, Tomsk, 634050, Russia (dimon@math.tsu.ru, starch@math.tsu.ru)*

Abstract: The limitations of the widespread methods of air pollution level estimation allow producing quite rough air quality estimations for the areas not covered by observation stations. Application of the computational models to this purpose requires usage of the meteorological observations, atmospheric emissions of industrial enterprises, measurements of the concentration of the atmospheric pollutants as well as the numerical modelling of the gaseous substance transformation and transport processes. High accuracy of estimation and forecast of air quality for such models are achieved with the use of information on pollutant emission from sources within the city and distribution of the meteorological parameters obtained with the use of model of atmospheric boundary layer. In this report the results of the work devoted to creation of the web-system (http://air.risks.scert.ru/tomsk-mkg/) for effective air chemical composition assessment and forecast in the conditions of industrial city area and its suburbia are presented. Tomsk city air pollution estimation data obtained as an output of the pollution transport mathematical model has been employed as a research basis for selected periods from 2000 to 2006. The air quality model is based on prognostic transport equations taking into account chemical reactions of air pollutants on the basis of reduced model of chemical reactions. Dynamic and thermal characteristics of the atmospheric boundary layer (ABL) are calculated on the basis of simplified model of urban ABL. Working version of the dedicated web-application has been created based on the web-portal ATMOS engine. At present it allows to estimate such characteristics as average monthly and seasonal pollutions for chosen time interval of the day along with their yearly dynamics as well as daily dynamics for various pollutants such as sulphur dioxide, nitric oxide, carbon monoxide, dust, etc. The web-system being considered could be used for long duration ecological risks assessment as well as for calculation of pollution scenarios resulted from different technological breakdown events for various meteorological conditions.

Keywords: Air pollution; Ecological monitoring; Web applications; Meteorology.

1. INTRODUCTION

Atmospheric boundary layer processes have an essential impact on the biosphere and human activity [Oke, 1987]. Even minor changes in moisture exchange and radiation balance between the Earth surface and atmosphere, chemical composition of the air and other characteristics may have serious consequences for environment. For this reason one of the key tasks in the area of environment preservation at present is modelling of

atmospheric processes and creation of applied Internet-accessible information-computational systems aiming at monitoring and forecast of ecological and meteorological conditions of atmospheric boundary layer. In particular, this task becomes very important in the current period of rapid development of industry, power engineering and road network that causes permanent atmospheric air deterioration due to increase of the number of factors affecting its chemical and aerosol composition. During the past decades simulation models have been widely used for scenario computations designed for clarifying specific features of pollution propagation over chosen area under various meteorological conditions. Scenario analysis is performed to study the contribution of separate emission sources into general air pollution, as well as to estimate effects of possible emergency situations at extra-hazardous objects, such as nuclear power-stations, chemicals plants, etc. Simulation models are also included in on-line information systems for air quality monitoring and forecast. Such systems provide real-time detailed information on distribution of air pollution concentration over the urban territory, with the stationary observation data being used as initial and boundary conditions as well as for validation of calculation results.

It should be noted that analysis of the air pollution state can not be completed without taking into account the contribution of secondary emission products, i.e., resultants of chemical and photochemical reactions between constituents of anthropogenic emissions and atmosphere gases [Moussiopoulos et al., 1995]. Many of such compounds are highly toxic; they form so-called urban photochemical smog, which lowers the visibility and affects detrimentally human beings, animals and plants. At present many models are designed to estimate concentrations of secondary pollutants with a precision depending in many respects on the number of constituents being considered and connecting equations varying from tens to few hundreds for different procedures.

2. MODEL DESCRIPTION

Pollution transport mathematical model represents the model of turbulent diffusion, including the transport equations with the description of advection, turbulent diffusion and chemical reactions, which is applied to calculate the air pollutant concentrations taking into account the chemical changes [Belikov et al., 2005]:

$$\frac{\partial C_i}{\partial t}+\frac{\partial UC_i}{\partial x}+\frac{\partial VC_i}{\partial y}+\frac{\partial W_cC_i}{\partial z}=-\frac{\partial}{\partial x}\langle c_iu\rangle-\frac{\partial}{\partial y}\langle c_iv\rangle-\frac{\partial}{\partial z}\langle c_iw_c\rangle+S_i+R_i;\quad i=1,..,n_s \tag{1}$$

Here $C_i(t,x,y,z)$, $c_i(t,x,y,z)$ are pulsating and averaged constituents of i-components concentration of impurity respectively; u,v and U,V are pulsating and averaged horizontal components of wind speed respectively; w_c, W_c are pulsating and averaged vertical components of impurity speed; S_i represents emissions of i-component of impurity in the atmosphere; R_i describes generation and transformation of substance during chemical reactions; n_s is a number of chemical impurity components considering in the chosen scheme of chemical reactions.

Original expressions for turbulent fluxes $\langle c_iu\rangle\ \langle c_iv\rangle\ \langle c_iw_c\rangle$ are used in the given work [Belikov, 2006]:

$$-\langle c_iu\rangle=\frac{\tau}{C_{1\theta}}\left(\left(1-C_{2\theta}\right)\langle c_iw\rangle\frac{\partial U}{\partial z}+\langle u_ju\rangle\frac{\partial C_i}{\partial x_j}\right); \tag{2}$$

$$-\langle c_iv\rangle=\frac{\tau}{C_{1\theta}}\left(\left(1-C_{2\theta}\right)\langle c_iw\rangle\frac{\partial V}{\partial z}+\langle u_jv\rangle\frac{\partial C_i}{\partial x_j}\right); \tag{3}$$

$$-\langle c_i w_c \rangle = \frac{\tau}{C_{1\theta} + D_{1C}F}\left(-(1-C_{3\theta})\frac{g}{\Theta}\langle c_i \theta \rangle + \langle u_j w_c \rangle \frac{\partial C_i}{\partial x_j} \right). \quad (4)$$

Here $u_j = (u, v, w_c)$, F is the function defining the surface influence on the turbulent flow structure; $c_x = 5.0$, $C_{1\theta} = 3.0$, $C_{2\theta} = 0.346$, $C_{3\theta} = 0.333$, $D_{1C} = 0.806$ are empirical constants, $\tau = l/C_D\sqrt{k}$ is the time scale of turbulent pulsations, g is the gravity acceleration, θ, Θ are pulsating and averaged components of potential temperature. Repetitive index j means summation.

To calculate the unknown correlation of pulsations of temperature and concentration $\langle c\theta \rangle$ included in equation (4), the following differential equation is used:

$$\frac{\partial \langle c_i \theta \rangle}{\partial t} + U_j \frac{\partial}{\partial x_j}\langle c_i \theta \rangle = -\langle c_i w_c \rangle \frac{\partial \Theta}{\partial z} - \langle \theta u_j \rangle \frac{\partial C_i}{\partial x_j} + \frac{\partial}{\partial x_j}\left(\alpha \frac{k}{\varepsilon} \langle u_k u_j \rangle \frac{\partial \langle c_i \theta \rangle}{\partial x_j} \right) - \frac{2}{\tau \cdot c_x}\langle c_i \theta \rangle; \quad (5)$$

$$i = 1,..,n_s; \;\; j = 1,..,3; \;\; k = 1,..,3.$$

Expressions (2)-(4) have the form of the Boussines gradient ratio: $\langle cu_i \rangle = -K_{ij}\frac{\partial C}{\partial x_j}$.

Turbulent flows of heat and impulse, included into the derived expressions (2)-(4) are defined using the algebraic ratios presented in [Starchenko, 2000].

3. PROBLEM FORMULATION

Mainly this work is devoted to creation of the web-system (http://air.risks.scert.ru/tomsk-mkg/) for effective air chemical composition assessment and analysis in the conditions of industrial city area and its suburbia. At present detailed pollutant concentration fields were obtained as an output of the pollution transport mathematical model described earlier for selected periods from 2000 to 2006. Dynamic and thermal characteristics of the atmospheric boundary layer are calculated on the basis of simplified mesoscale model of urban ABL developed at the Tomsk State University and Institute of Atmospheric Optics RAS.

The problem was solved numerically in the investigated domain that covers an urbanized territory with numerous high-rise point sources, as well as linear, point and area emission sources situated on the underlying surface. Predictions were performed with the use of finite volume method and new explicit-implicit calculation scheme of solving three-dimensional prognostic transport equations at Skif-Cyberia multiprocessor systems of Tomsk State University.

Here the dataset obtained as an output of pollution transport and transformation mathematical model was put onto the powerful server and provided with basic software toolset accessible via unified web-interface for data processing and visualization and subsequent analysis. Web-portal ATMOS software [Gordov et al., 2006], specially designed for rapid development of scientific applications, has been used as a base for web-system development. Graphic user interface of the system is planned to realize using DHTML technology for it allows to provide more friendly user interface than standard HTML [Okladnikov and Titov, 2006]. PHP scripting language is used for implementation of relevant program modules within the framework of ATMOS web-portal. The work is devoted to the practical realization of the working version of the dedicated web-application for effective air quality assessment of the city area [Gordov et al., 2007].

4. SYSTEM FUNCTIONAL DESCRIPTION

The system being developed comprises three following parts:

- Generated by the model [Belikov et al., 2005] data archives containing detailed fields of the pollutant concentrations for selected for each season periods in 2000-2006 interval
- Graphic user interface
- A set of PHP-scenarios to perform data processing and conversion with the following visualization

It should be noted that raw data archives obtained as a result of numerical modelling are incompatible with the data used by the system for a number of technical reasons, so that first they were structured and converted into the standard binary format used by GrADS software package for the following online visualization.

Graphic user interface is developed based on the web-portal ATMOS engine and represents dynamic HTML form to enter calculation and visualization parameters (Figure 1).

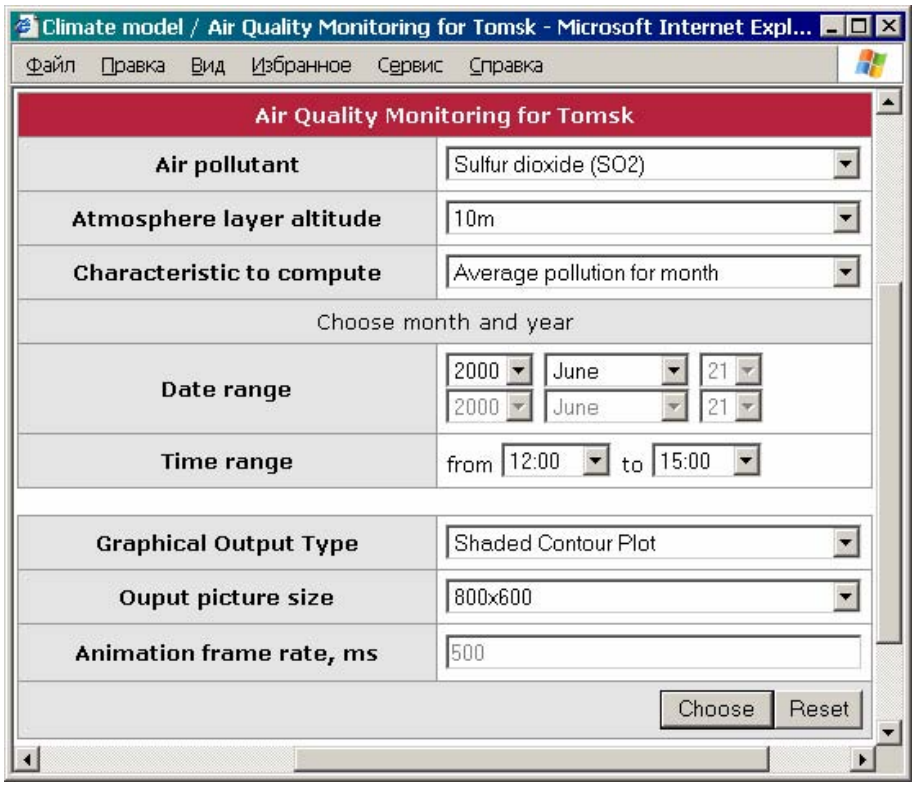


Figure 1. Dynamic HTML form for entering calculation parameters

The form allows setting the following parameters:

- "Air pollutant" with such values as airborne particulate matter, sulfur dioxide, nitrogen dioxide, carbon oxide and ozone)
- "Atmospheric layer altitude" ranging from 10 to180 meters
- "Characteristic to compute". At present it is possible to calculate such characteristics as average pollution for month and season for chosen time of the day and their dynamics within the selected time interval as well as hourly dynamics for the selected day

- Date range, time range, graphical output type and picture size. There's also a possibility to choose animation frame rate to see dynamics of the pollutant concentration.

"Date range" fields allow to set the time interval of interest. It should be noted that contents of the fields is generated dynamically according to the model data currently available for processing. It is an important issue from the point of view of effective access to relevant information because model computational process is very heavyweight so that output data can be received in relatively small blocks. It is also possible to set "Graphical output type" and animation frame rate. After the form is filled in and calculations are carried out web-application will render the results on the screen.

5. CURRENT RESULTS

Below several examples demonstrating present capabilities of the web-system are to be found. Figure 2 shows average nitrogen dioxide pollution of the Tomsk city area for September, 2006 from 6.00 pm till 9.00 pm.

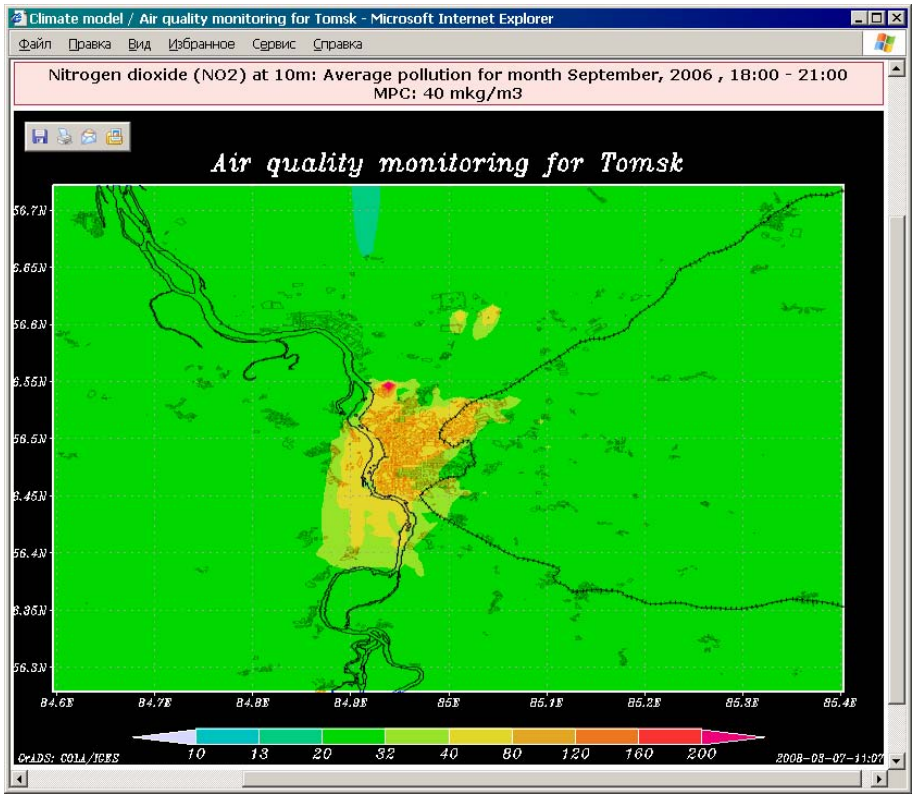


Figure 2. Average NO_2 pollution at 10m for September, 2006, 18.00 – 19.00.

It should be noted that the scale displayed on the graphics generated by the system is normalized according to the corresponding maximum permissible concentration value residing in the middle of the scale that makes the picture easy for interpretation.

The functional capability of the system of calculating monthly and seasonal averages for a number of years and representing them in animation mode would be useful for analysing pollutant concentration dynamics. Table 1 that follows below shows an example of visualization of average seasonal SO_2 pollution for 6 years.

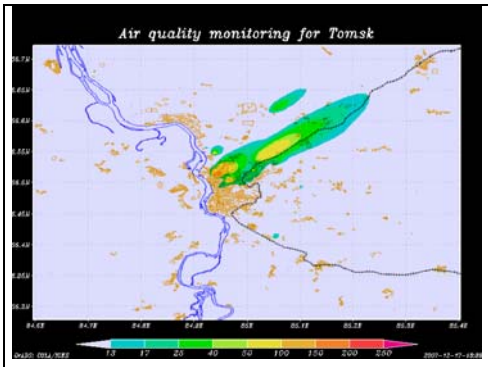


Figure 3. Winter, 2001

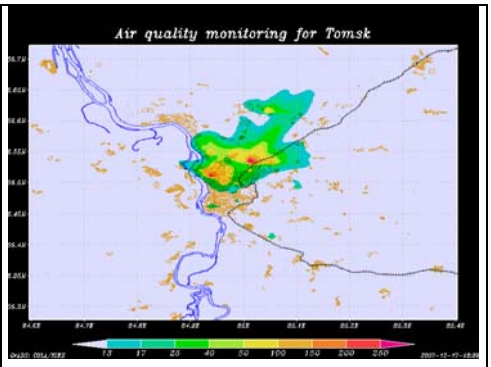


Figure 4. Winter, 2002

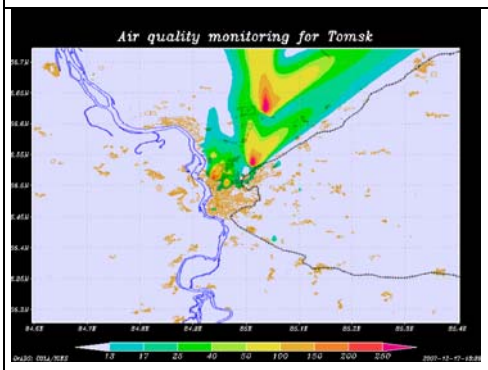


Figure 5. Winter, 2003

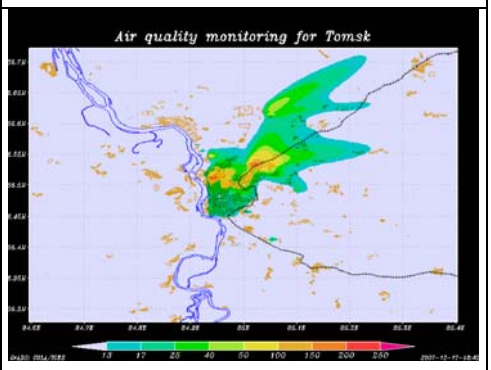


Figure 6. Winter, 2004

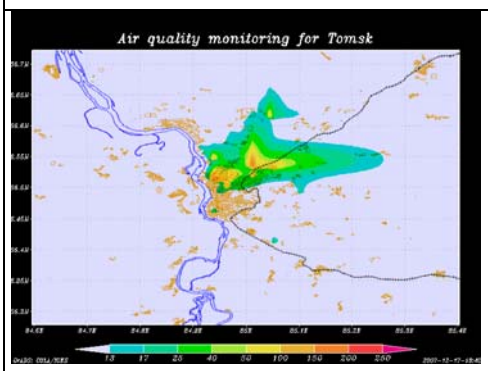


Figure 7. Winter, 2005

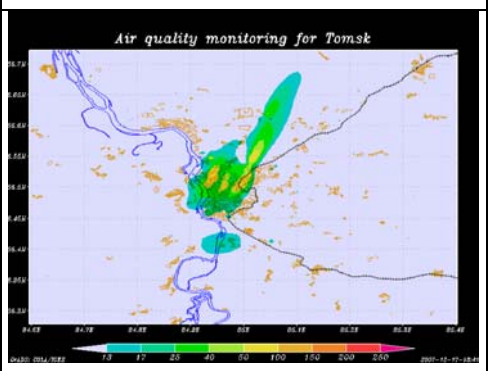


Figure 8. Winter, 2006

Table 1. Dynamics of the average SO2 pollution for winter season at 10m, 12.00 – 15.00.

6. CONCLUSION

The system is aimed at regional ecologists and decision makers and provides them with images of pollutant concentrations fields at different altitudes for the industrial city area. It might be used to determine characteristics of pollution distribution above the territory and their dynamics under different weather conditions, to estimate input of selected pollution sources (industry enterprises, transport, etc.) into pollution fields, as well as to estimate consequences of possible accidents leading to additional pollutants blowouts. Also it might be used to understand degree of anthropogenic influence on regional environment and climate. It should be added that the system has generic character and being provided with characteristics of industrial and transport pollution sources, local meteorology data, surface properties and generated by the pollution transport and transformation mathematical model datasets it can be easily adjusted for conditions of an arbitrary city.

ACKNOWLEDGEMENTS

This work is partially supported by SB RAS Integrative Project 34, SB RAS Basic Research Program Project 4.5.2.2, RFBR grant #05-05-98010 and FP6 Enviro-RISKS project (INCO-CT-2005- 013427).

REFERENCES

Belikov, D.A., A.V. Starchenko, An Investigation of Secondary Pollutants Formation (Ozone) in the Atmosphere of Tomsk, *Journal of Atmospheric and Oceanic Optics*, 18(05-06), 391-398, 2005.

Belikov D.A., Parallelization method achievement of numerical model of atmospheric diffusion for investigation of primary and secondary pollutants transport over urban territory, *Thesis for the degree of Candidate of Physics-and-Mathematics Science*, Tomsk State University, 2006.

Gordov, E.P., V.N. Lykosov, and A.Z. Fazliev, Web portal on environmental sciences "ATMOS", *Advances in Geosciences*, 8, 33–38, 2006.

Gordov, E.P., I.G. Okladnikov, and A.G. Titov, Development of elements of web-based information-computational system supporting regional environment processes investigations, *Computational Technologies*, 12(3), 20-28, 2007.

Moussiopoulos, N., P. Sahm, and C. Kessler, Numerical simulation of photochemical smog formation in Athens, Greece - a case study, *Atmospheric Environment*, 29(24), 3619-3632, 1995.

Oke, T.R., *Boundary layer climates*, Methuen, 435 pp., London and New York, 1987.

Okladnikov, I.G., A.G. Titov, Web-system for processing and visualization of meteorological data, *Environmental Observations, Modeling and Information Systems*, 42-45, Tomsk, 2006.

Starchenko A.V., Modeling of impurity transport in a homogeneous atmospheric boundary layer, *Proceedings of International conference ENVIROMIS-2000*, 77-82, Tomsk, 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

Runoff projection sensitivity to rainfall scenario methodology

Freddie S. Mpelasoka and Francis H. S. Chiew

CSIRO Land and Water, Black Mountain, Canberra, Australia

(Freddie.Mpelasoka@csiro.au)

Abstract: Runoff characteristics are inextricably linked with climate, particularly the spatial and temporal patterns of precipitation and evapotranspiration. The need for demonstrably objective climate change scenarios consistent with what is realistic under global warming predicted conditions is increasingly growing. Global climate models (GCMs) are the best tools available for simulating global and regional climate for predicting future climate. However, GCMs provide information at a resolution that is too coarse to give results that can be directly used in hydrological studies.

This paper quantifies three simple methods of rainfall scenarios construction informed by GCMs for their potential use in providing desirable future rainfall scenarios for modelling runoff projections. Runoff simulations from daily rainfall time series obtained using three simple methods (*constant scaling*, *daily scaling* and *daily translation*) to transform GCM outputs to catchment-scale rainfall over South-East Australia for 1981-2000 and 2046-2065 periods are used. In the *constant scaling* and *daily scaling* methods, the historical observed daily rainfall is scaled by the changes indicated by the GCM. In the *constant scaling* method, the entire daily rainfall series (in seasons) is scaled by the same factor. In *daily scaling*, the different daily rainfall amounts are scaled differently. In *daily translation*, a relationship between GCM simulation of the present rainfall and the observed catchment-scale rainfall is established, and used to convert the future GCM daily rainfall time series to catchment-scale rainfall series.

In summary, the *constant scaling* method can be used in most applications to transform climate outputs from GCMs to drive hydrological models. However, where more detailed analyses of runoff distribution is required, the *daily scaling* and *daily translation* methods are potentially better, particularly if the GCMs used have skill in modelling extreme rainfall and daily rainfall series.

Keywords: Runoff projections; climate change; scenario construction; rainfall

1. Introduction

Water shortage is a recurrent phenomenon of climate variability in Australia and has significant environmental and socio-economic impacts. For example, the drought of 2002-2003 cost Australia A$10 (US$7.6) billion (Adams *et al*. 2002), about 70,000 jobs were lost and had significant impacts on tourism which currently contributes to 4.5% of Australian Gross Domestic Product (Allen 2005). Water restrictions are becoming almost permanent features in many regions and cities of Australia in response to water shortages. This emphasises Australia's vulnerability to climate variability and limitations of adaptive capacity (Mpelasoka et *al*. 2007).

Most of the current projections from GCMs show a general decrease of mean rainfall over Australia coupled with increase in temperature (Suppiah et *al*. 2007). Changes in spatial and temporal patterns of climate variables will undoubtedly impact on regional hydrological processes. In particular, changes in rainfall will be amplified as an impact on runoff (Chiew, 2006). This will have significant implications on water resources, due to changes in both mean rainfall and increases in variability and extreme events. GCMs are the best tools available for simulating global and regional climate for future climate. However, GCMs provide information at a resolution that is too coarse to give results that can be used directly in hydrological modelling (Mpelasoka *et al* 2001; Mearns and Hulme, 2001).

Various approaches have been employed to develop climate change scenarios at different scales. For example, dynamic and statistical downscaling techniques, ranging from the simple to the complex, are frequently used to transform large-scale GCM outputs to catchment-scale climate variables. Statistical downscaling techniques relate large synoptic-scale atmospheric predictors to catchment scale rainfall (Charles *et al*., 2004; Mpelasoka *et al*., 2001). However, the choice of appropriate atmospheric predictors and the calibration of statistical downscaling models are fairly laborious and subject to expert judgement. Dynamic downscaling techniques, by nesting high resolution limited area models within a GCM, account for processes that are not resolved by the GCM (Nunez and McGregor, 2007; Lorenz and Jacob, 2007). While this approach is conceptually consistent with the GCM representation of the climate system dynamics, they suffer from several limitations, including high computing costs. Also, the success of a dynamic model strongly depends on its horizontal resolution and the computational expenses often limit the sample sizes.

Simpler empirical approaches such as pattern scaling and translation methods offer a more immediate solution (Kidson and Thompson, 1998). Because they are simple to use, they can be applied to outputs from a number of runs of different GCMs, therefore taking into account the large uncertainties associated with global warming and local climate change projections. However, there is need for these methods to also account for changes in extremes and sequence of events, as it is evident that changes will not be restricted to the mean state of the climate but also in the higher order moments (Mearns *et al*., 1996; IPCC-TGCIA, 1999).

This paper compares runoff results from a rainfall-runoff model driven with future climate inputs obtained using three simple methods for transforming rainfall from several GCM transient experiments: *constant-scaling*, *daily-scaling* and *daily-translation*. We demonstrate that the *constant scaling* method can be used in most applications to transform climate outputs from GCMs to drive hydrological models. However, where more detailed analyses of runoff distribution is required, the *daily scaling* and *daily translation* methods are potentially better, particularly if the GCMs used have skill in modelling extreme rainfall and daily rainfall series.

2. Data and Methods

2.1 Historical climate data and modelling of historical runoff

Daily runoff is modelled using a lumped conceptual daily rainfall-runoff model, SIMHYD, for 0.25^o x 0.25^o grid cells across South Eastern Australia (SEA). The SIMHYD model has been used successfully for various applications across Australia. The structure of SIMHYD and the algorithms describing the processes modelled by SIMHYD are shown in Figure 1. The seven parameters in SIMHYD have been calibrated against observed streamflow data for 331 catchments across Australia by Chiew *et al.* (2002). For this study, runoff for each 0.25^o x 0.25^o grid cell is modelled using optimised parameter values from the geographically closest 'calibration catchment'.

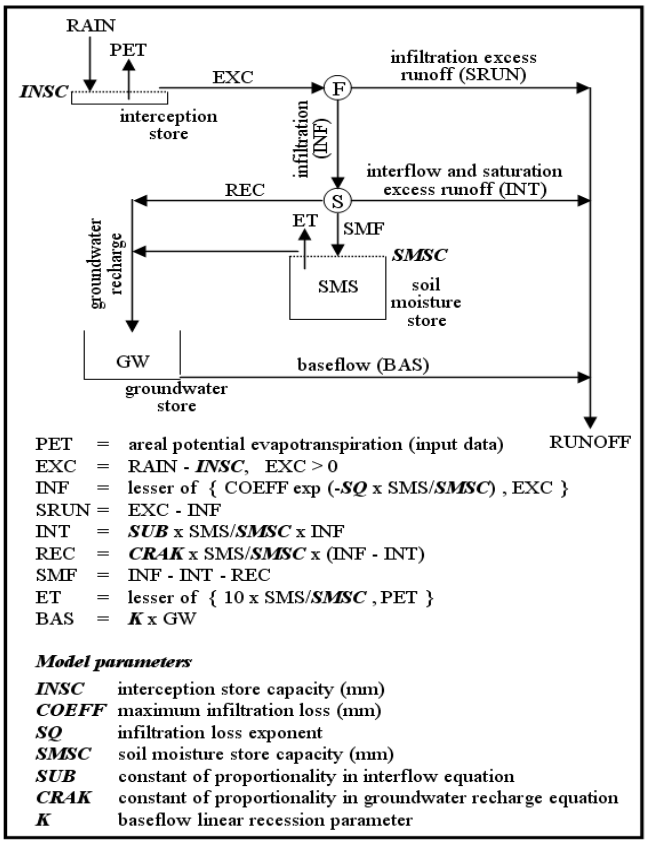


Figure 1: Structure of the SIMHYD model.

Daily rainfall and areal potential evapotranspiration (APET) are required to run SIMHYD. The observed climate series for 1981-2000 are obtained from the 'SILO Data Drill' of the Queensland Department of Natural Resources and Water (Jeffrey *et al.*, 2001). Areal PET (APET) is calculated from the SILO climate surface using Morton's wet environment evapotranspiration algorithms (Morton, 1983). The 1981-2000 daily rainfall and APET series are used to drive SIMHYD to estimate daily runoff for 1981-2000.

2.2 Future climate and modelling of future runoff

To obtain the future climate series, daily simulations from three GCMs for 1981-2000 and for 2046-2065 for the mid-range IPCC A1B global warming scenario are used. The three GCMs used are *cccma* (Climate Modelling and Analysis Centre, Canada, 3.8^o x 3.8^o horizontal resolution, 31 levels), *mk3* (CSIRO, Australia, 1.9^o x 1.9^o horizontal resolution, 18 levels) and *gfdl* (Geophysical Fluid Dynamics Lab, USA, 2.5^o x 2.0^o horizontal resolution, 24 levels). The simulations were obtained from the World Climate Research Program (WCRP) Coupled Model Intercomparison Project (CMIP3) (https://esg.llnl.gov:8443/index.jsp).

Since GCM simulations of current climate generally validate well against observations at seasonal to annual temporal scales, we applied the three methods of scenario-construction to seasonal changes. In the *constant scaling* method, the mean rainfall and APET simulated by the GCM for 2046-2065 and for 1981-2000 are compared to determine the change in the mean rainfall and APET. The ratios of mean rainfall and APET between 2046-2065 and 1981-2000 are then used to scale the historical 1981-2000 daily rainfall and APET series to obtain the 2046-2065 rainfall and APET series to drive SIMHYD. Each of the four seasons are considered separately (i.e., a different constant scaling factor is used for each season).

In the *daily scaling* method, the daily rainfalls from the GCM for 2046-2065 and for 1981-2000 are ranked from highest to lowest, and changes at the different ranks/percentiles are determined. The percentage change to each rainfall rank/percentile is then used to scale the observed 1981-2000 daily rainfall (i.e., different rainfall amounts are scaled differently) to obtain the 2046-2065 rainfall series to drive SIMHYD. Again, each of the four seasons is considered separately. The 2046-2065 daily rainfall series obtained is then rescaled such that the mean rainfall for each season is the same as that in the *constant scaling* method. Common to 'scaling' approaches is the assumption that the relative pattern of change is scale invariant, that is, a relative pattern of change calculated at the GCM grid-scale can be applied at any point within a GCM grid-box.

The *daily translation* approach is based on the comparison of the GCM cell simulations for the 20^{th} century to observed data over the same period on finer grid (at different rainfall percentiles), to establish a model 'biases-correction scheme' against the catchment scale climate variable of interest. The 1981-2000 daily rainfall distribution from the GCM is compared with the 1981-2000 observed rainfall at each 0.25^o grid cell to relate the GCM daily rainfall to 0.25^o grid cell daily rainfall at the different rainfall ranks/percentiles. This relationship is then used to translate the 2046-2065 daily rainfall series simulated by the GCM, to 0.25^o grid cell daily rainfall series. The 2046-2065 daily rainfall series are also rescaled such that the mean rainfall for each season is the same as that in the *constant scaling* method.

The same SIMHYD parameter values used for the 1981-2000 modelling are used for the 2046-2065 modelling with future rainfall obtained using the *constant scaling*, *daily scaling* and *daily translation* methods.

3. Results and discussion

3.1 Changes in rainfall and potential evapotranspiration

The three scenario-construction methods produce exactly the same changes in mean rainfall and APET for each of the four seasons due to the common rescaling (2046-2065 relative to 1981-2000). Figure 2 shows changes in annual rainfall derived from the three GCMs by applying *constant scaling*, *daily scaling* and *daily translation* methods.

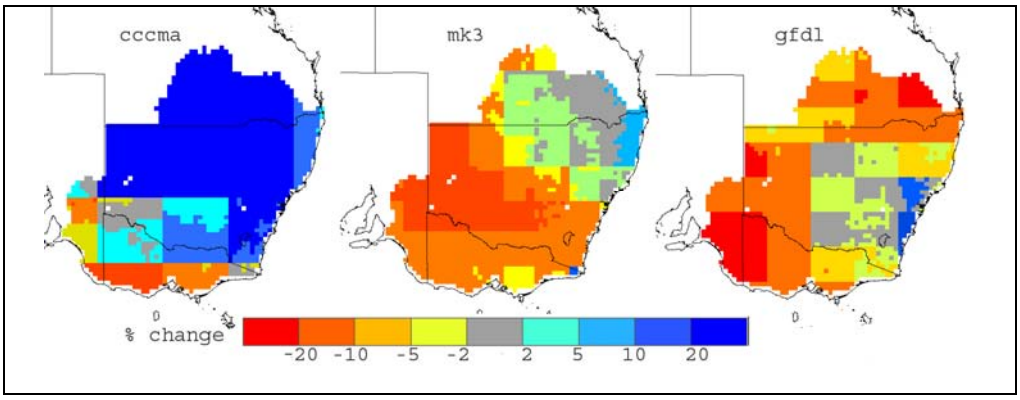


Figure 2: Changes in annual rainfall for 2046-2065 with respect to 1981-2000 derived from *cccma*, *mk3* and *gfdl*.

To interpret the differences in the runoff characteristics between the scenario-construction methods of rainfall, which is the main driver of runoff, only *constant scaling* was used in the construction of APET scenarios.

3.2 **Changes in runoff characteristics**

Figure 3 shows changes in mean annual runoff for 2046-2065 relative to 1981-2000 modelled by the rainfall-runoff model using the *constant scaling*, *daily scaling* and *daily translation* methods to provide future rainfall series as informed by the three GCMs. A comparison of percent of study area with change in mean annual rainfall less than a specified percentage change shows systematic differences between the *constant scaling* and *daily scaling* methods.

The *daily scaling* method shows higher runoff than the *constant scaling* method. This is because the three GCMs generally indicate that the increase in extreme rainfall in a future climate is more than the increase in mean rainfall (or the decrease in extreme rainfall is less than the decrease in mean rainfall). The *daily translation* method considers changes in the future rainfall time series, and runoff results from the *daily translation* method can be significantly different to the *constant scaling* and *daily scaling* methods.

Figure 4 shows an example of the potential use of the proposed simple scenario construction methods in providing objective information on the impacts of climate change in water accounting. A scenario of water availability at Torrumarry Weir, in the Murray-Darling Basin derived from *gfdl* GCM for 19 years (2047-2065) relative to 1982-2000 period show a general decline in water flows. However, occasional relatively high/extreme flows are also exhibited by flows derived using *daily scaling* and *daily translation* rainfall scenario construction methods. The Torrumarry Weir facilitates the division of flows for the main irrigation agriculture area in the Murray-Darling basin. The amount of water supplied for irrigation is a function of demand, but water restrictions are applied when river low flows reach pre-established thresholds. Table 1 shows a summary of the decreases in the projected flows relative to respective monthly means. The *constant scaling* method shows relatively low projections of the number of months under decreased flows compared to *daily scaling* and *daily translation* methods. This can be attributed to the tendency of the *daily scaling* method to suppress the variability of the flow. Potentially such information in the table is useful in decision making on levels of water restrictions over various periods.

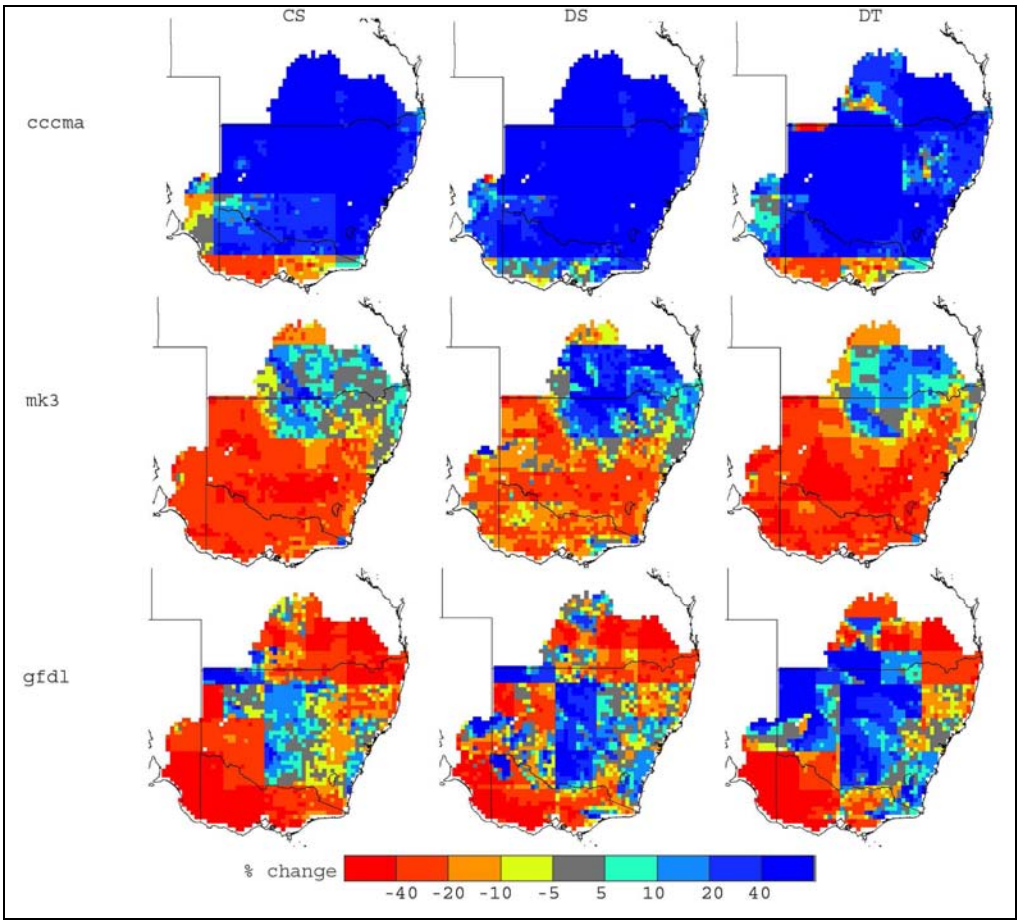


Figure 3: Change in mean annual runoff for 2046-2065, with respect to 1981-2000, derived from *cccma* (row 1), *mk3* (row 2) and *gfdl* (row 3) by using *constant-scaling* (CS), *daily scaling* (DS) and *daily translation* (DT) methods.

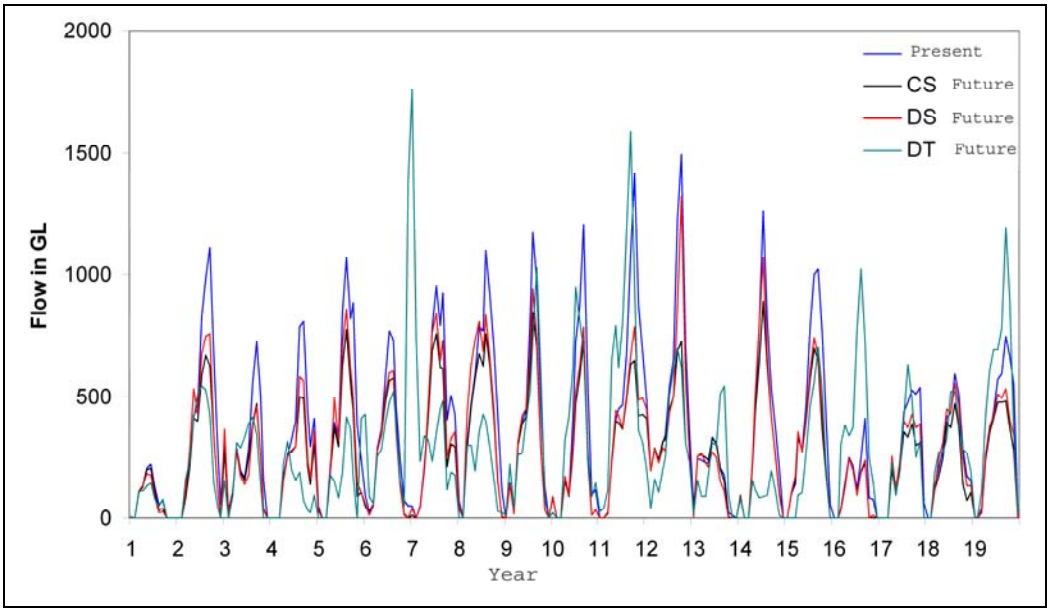


Figure 4: Monthly flows of the Murray-Darling River at Torrumbary Weir for 1982-2000 (present) and projected flows for 2047-2065 (future) derived from *gfdl* GCM rainfall scenarios using *constant-scaling* (CS), *daily scaling* (DS) and *daily translation* (DT) methods.

Table 1: Total number of months under projected changes in monthly flows at Torrambarry Weir in the Murray-Darling Bain for the 2047-2065 period relative to 1982-2000.

Change (%)	Number of months showing change		
	CS	DS	DT
-50	18	26	24
-25	30	38	38
-10	41	51	44

4. Conclusions

The *constant-scaling* method is currently used in almost all hydrological impact modelling studies due to its simplicity. However, the *daily scaling* method is likely to be a better method because it considers different changes to different rainfall amounts. This can be particularly important because many GCMs indicate that the extreme rainfall in a future climate is likely to increase by more than the increase in mean rainfall or decrease by less than the decrease in mean rainfall. As high rainfall events generate significant runoff, the runoff estimated from the *daily scaling* method will be higher than that estimated by the *constant-scaling* method. However, the difference between the *constant-scaling* and *daily scaling* modelling results is small compared to the differences between the three GCMs.

The *daily translation* method considers changes in the future rainfall time series, and runoff results from the *daily translation* method can be significantly different to the *constant-scaling* and *daily scaling* methods. As the *daily translation* method uses the future rainfall series modelled by the GCM and translates it to catchment scale rainfall, the *daily translation* method potentially exhibits the dynamics of the climate system associated with global climate change as simulated by the GCM.

The *constant-scaling* method is useful because it is simple and can be used with results from many GCMs and global warming scenarios to take into account the large uncertainties associated with climate change projections. The GCM simulation of mean climate is also considerably better than its ability to simulate extreme climate and daily series. However, the *daily scaling* method is conceptually better than the *constant-scaling* method if GCMs have skill in modelling extreme rainfall. Likewise, the *daily translation* method is conceptually better than the *constant-scaling* and *daily scaling* methods if GCMs have skill in modelling daily rainfall series and if the rainfall at GCM grid cell can be meaningfully translated to catchment scale rainfall.

In summary, the constant scaling method can be used in most applications to transform climate outputs from GCMs to drive hydrological models. However, where more detailed analyses of runoff distribution is required, the daily scaling and daily translation methods are potentially better, particularly if the GCM used have skill in modelling extreme rainfall and daily rainfall series.

ACKNOWLEDGEMENTS

This work was funded by the South Eastern Australian Climate Initiative (SEACI). The authors would like to thank Harvey Davies for his support in developing Tcl/NAP software tools used extensively in this study, Mainuddin Mohammed for water accounting analysis and Steve Charles for reviewing the manuscript.

REFERENCES

Adams, P.D., Horridge M., Masden J.R. and Wittwer G., Drought, regions and the Australian economy between 2001-02 and 2004-05, *Australian Bulletin of Labour*, 28, 233-249, 2002.

Allen Consulting Group, Climate Change Risk and Vulnerability: Promoting an Efficient Adaptation Response in Australia, *The Allen Consulting Group*, 159 pp, 2005.

Charles SP, Bates BC, Smith, IN & Hughes JP, Statistical downscaling of daily precipitation from observed and modelled atmospheric fields, *Hydrological Processes*, 18, 1373-1394, 2004.

Chiew FHS, Estimation of rainfall elasticity of streamflow in Australia, *Hydrological Sciences Journal*, 51, 613-625, 2006.

Chiew, F.H.S., Peel, M.C. and Western, A.W.., Application and testing of the simple rainfall-runoff model SIMHYD, In: Mathematical Models of Small Watershed Hydrology and Applications Editors: V.P. Singh and D.K. Frevert, *Water Resources Publication, Littleton*, Colorado, USA, pp. 335-367, 2002.

IPCC-TGCIA, Guidelines on the use of scenario data for climate impact and adaptation assessment. Version 1. Prepared by Carter, T.R., Hulme, M. and Lal, M., Intergovernmental Panel on Climate Change, Task Group on Scenarios for Climate Impact Assessment, 69pp, 1999.

https://esg.llnl.gov:8443/index.jsp, World Climate Research Program (WCRP) Data. Last cited, 19 February,2008.

Jeffrey, S.J., Carter, J.O., Moodie, K.M and Beswick, A.R, Using spatial interpolation to construct a comprehensive archive of Australian climate data", *Environmental Modelling and Software*, Vol **16/4**, pp 309-330, 2001.

Kidson J W, Thompson CS., A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate* **11**: 735–753, 1998.

Lorenz P. and Jacob D., The link between regional and global climate -applications of the two-way nested GCM-RCM system. *Second International Conference on Earth System Modelling (ICESM) Abstracts*, Vol. 1, ICESM2007-A-00139, 2007.

Mearns, L.O., Rosenzweig, C. and Goldberg, R., The effect of changes in daily and interannual climatic variability on CERES-wheat: A sensitivity study. *Climatic Change* 32, pp. 257–292, 1996.

Morton, F. I., Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. .*Journal of Hydrology* **66**: 1-76, 1983.

Mpelasoka F.S., Hennessy K., Jones R. and Bates B., Comparison of suitable drought indices for climate change impacts assessment over Australia towards resource management, *International Journal of Climatology*, (www.interscience.wiley.com) DOI: 10.1002/joc.1649, 2007.

Mpelasoka F., Mullan A.B. and Heerdegen R.G., New Zealand climate change information derived by multivariate statistical and artificial neural network approaches. *International Journal of Climatology*, 21, 1415–1433, 2001.

Nunez M. and McGregor J. L., Modelling future water environments of Tasmania, Australia. *Climate Research*, 34, 1-14, 2007.

Suppiah, R., K. J. Hennessy, P. H. Whetton, K. McInnes, I Macadam, J. Bathols, J. Ricketts and C. M. Page, Australian climate change projections derived from simulations performed for the IPCC 4th Assessment Report. *Australian Meteorological Magazine* (In press), 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

Assessing Activity-Related Vehicle Emissions through an Integrated Activity-Based Modelling Framework

C. Beckx[a,b], T. Arentze[c], L. Int Panis[a], D. Janssens[b] and G. Wets[b]
[a] *Flemish Institute of Technological Research. Boeretang 200, 2400 Mol, Belgium (carolien.beckx/luc.intpanis@vito.be)*
[b] *Transportation Research Institute, Hasselt University. Wetenschapspark 5 bus 6, 3590 Diepenbeek, Belgium* (carolien.beckx/davy.janssens/geert.wets@uhasselt.be)
[c] *Eindhoven University of Technology. PO Box 513, 5600 MB Eindhoven, The Netherlands (T.A.Arentze@bwk.tue.nl)*

Abstract: Policymakers are increasingly recognizing the importance of source-related measures, instead of technological actions, to tackle the problem of traffic air pollution. Unfortunately, traditional trip-based models fail to make accurate predictions for activity-related policy questions. Due to the richer set of concepts which are involved in activity-based transportation models, the use of these models should be encouraged to contribute to this part. In this research the activity-based model ALBATROSS was used to assess trips and emissions produced by passenger cars in the Netherlands. The results were segregated according to trip motive to gain more insights into the contribution of different trips towards the total amount kilometres and air pollution. The predicted values correspond well with the reported values from the Dutch Scientific Statistical Agency. Predictions for total travelled distance, carbon dioxide, carbon monoxide, nitrogen oxide, sulphur dioxide and particulate matter (PM) differed not more than 8% from the officially reported values. Concerning the classification into trip motive, the commuter trips produced almost half of the PM emissions. Further, trips with a social purpose caused 17% of the PM emissions, and shopping and leisure trips each accounted for 10% of the total PM emissions. This paper is novel in the sense that it reports on the applied methodology and presents the practical results from a case study of the activity-based modelling approach as well.

Keywords: Activity-based modelling; ALBATROSS; Emissions; trip motive.

1. INTRODUCTION

The last few years, there is an increasing concern over the environmental impacts of traffic as traffic volumes have continued to rise. Technological innovations on motor vehicles (e.g. the European directives 91/441/EEC, 94/12/EC and 98/69/EC) are able to diminish the environmental consequences of this increase. However, due to the continuing increase of vehicle kilometres on the road, the results of these improvements will be partially offset, forcing the implementation of other traffic measures. National and local policymakers are therefore recognizing the importance of source-related measures [EEA, 1999]. Unlike technological measures, focusing on the end of the problem chain, source-related measures intervene in an earlier stage of the problem and focus on the demand for travel. Concerning the problem of traffic air pollution, focusing on the source of the problem means concentrating on people's travel behaviour: why are people travelling? Good quantitative information about people's travel behaviour should provide policymakers with the answers to this kind of question.

Unfortunately, questions that involve the linkages between activities and travel decisions can not be examined through a traditional trip-based four-step transportation model. This kind of model focuses on individual trips where the spatial and temporal interrelationships between trips and activities are ignored. Together with other limitations [see e.g. Beckx et al. 2005], this inability has prompted the development of a new approach to travel analysis, the activity-based approach.

This paper presents the first results of a research project that applies an activity-based travel demand model to the evaluation of activity-related exhaust emissions. The activity-based model ALBATROSS, a fully operational computational process model developed by Arentze and Timmermans [2005], was used to predict trip distances and emission estimates for personal vehicle travel in the Netherlands. Trips were classified according to trip motive to gain an insight into the contribution of different trips.

The remainder of this paper is organized as follows. In the next section, the activity-based modelling approach is briefly explained and the activity-based model ALBATROSS is described. Next, the methodology to apply such an activity-based model for air quality purposes is explained followed by a presentation of the results from the trip analysis and emission assessment procedure. Finally, the results are discussed and the paper concludes with some important aspects of future research applications.

2. ACTIVITY-BASED MODELLING OF TRAVEL BEHAVIOUR

2. 1 Introduction

Modelling traffic patterns has always been a major area of concern in transportation and environmental research. Since 1950, due to the rapid increase in car ownership and car use in the US and in Western Europe, several models of transport mode, route choice and destination have been used by transportation planners, often referred to as “four-step models” [Ruiter and Ben-Akiva, 1978]. These models were necessary to predict travel demand on the long run and to support investment decisions in new road infrastructure. A lot of these aggregate four-step models failed to make accurate predictions. The major drawback is the focus on individual trips, where the spatial and temporal interrelationships between trips are ignored. Furthermore, the overall behaviour is represented as a range of constraints which define transport choice, while it is in fact both an outcome of real human decision making and of a complex choice process. The last drawback clearly is the complete negation of travel as a demand derived from activity participation decisions. This is where activity-based transportation models came into play.

The major idea behind activity-based transportation models is that travel demand is derived from the activities that individuals and households need or wish to perform [Ettema and Timmermans, 1997]. Activity-based approaches aim at predicting which activities are conducted, where, when, for how long, with whom and the transport mode involved. An activity-based model allows for spatio-temporal linkages between the collection of activities that individuals and households perform as part of their daily schedule. Besides providing detailed activity-travel information for individuals within a population, the advantages of activity-based modelling lie in its ability to give a better prediction of travellers’ responses to transportation control measures. Over the last years, several research teams have focused on building activity-based models of transport demand [e.g. Bhat et al. 2004; Arentze and Timmermans, 2005; Pendyala et al. 2005]. But, although the potential advantage of an activity-based approach for air-quality purposes has been recognized from the beginning [Spear, 1994] and has been re-iterated more recently [Shiftan, 2000; Beckx et al. 2005; 2007a], to the best of our knowledge- models that have been developed along these lines are still scarce.

2.2 The activity-based model ALBATROSS

The activity-based model 'A Learning-Based Transportation Oriented Simulation System' (ALBATROSS) was developed for the Dutch Ministry of Transportation, Public Works and Water Management as a transport demand model for policy impact analysis. ALBATROSS is a computational process model that relies on a set of decision rules, typically extracted from activity diary data, to predict activity-travel patterns [Arentze and Timmermans, 2000; Arentze et al. 2003]. The model is able to predict which activities are conducted, when, where, for how long, with whom, and the transport mode involved.

The activity scheduling agent of ALBATROSS is the core of the system which controls the scheduling processes. The scheduling model of ALBATROSS, which generates a schedule for each individual and each day, consists of three components. The first component generates an activity skeleton consisting of fixed activities and their exact start time and duration. Given the skeleton, the second component then determines the part of the schedule related to flexible activities to be conducted that day, their duration, time of day and travel characteristics. Both components use the same location model component determining the location of activities. All three components assume a sequential decision process in which key choices are made and predefined rules delineate choice sets and implement choices made in the current schedule. Interactions between individuals within households are to some extent taken into account by developing the scheduling processes simultaneously and alternating decisions between the persons involved.

Figure 1 schematically presents the structure of the first model component: the process model for generating fixed activities.

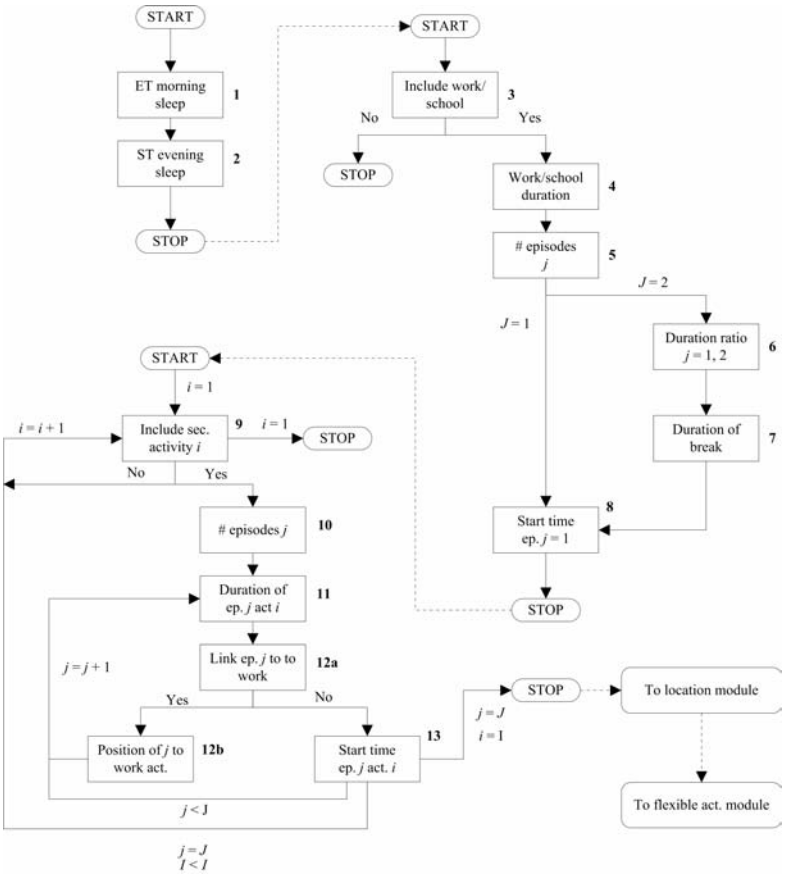


Figure 1. Process model for generating fixed activity patterns (ET = end time; ST = start time, sec. = secondary, ep. = episode, act. = activity)

Each numbered rectangle corresponds to a decision tree. The indices used in the figures are defined as follows: *i* is the index of activity in order of priority, $i = 1,\ldots, I$ and *j* is the index of episode of activity i in order of start time, $j = 1, \ldots, J$. The skeleton component comprises decisions 1 tot 13 and consists of several subprocesses, including determining the pattern of sleep activities (decisions 1 and 2); determining the pattern of the primary work/school activity (decisions 3 to 8) and determining the pattern of secondary, fixed activities (decisions 9 to 13). In a next step, the location of each fixed activity episode will be predicted in the location module and finally the flexible activities, and their locations, will be scheduled in the flexible activities model component. The model chooses the end time of the morning sleep episode and the start time of the evening sleep episode. The primary work/school activity has maximally two episodes and a minimum duration of 1 hour per episode. The pattern is defined by decisions about the number of episodes, start time, duration(s) and interepisode time. Work/school activities with shorter duration are treated as a separate category of secondary fixed activities in the next step.

Compared with other activity-based models, ALBATROSS is unique in that rules as opposed to principles of utility maximization underlie the scheduling decisions. Furthermore, the rather detailed classification of activities and inclusion of a full set of space-time and scheduling constraints are distinctive features of the model compared with most other models.

Detailed information regarding the working of this model and the applied decision rules can be found in Arentze et al. [2003] or in Arentze and Timmermans [2005].

3. METHODOLOGY

The methodology to use an activity-based approach for estimating trip distances and vehicle emissions per trip motive consists of three successive steps. Focusing on the application in the Netherlands, this section will describe each of these steps briefly.

3.1 Activity-based modelling

In this study ALBATROSS was applied to predict activity travel schedules for the total Dutch population. The model is estimated on approximately 10,000 person-day activity-diaries collected in the period of 1997-2001 in a selection of regions and neighbourhoods in the Netherlands. First, a synthetic population, representing 30% of the households in the Netherlands, was created with iteratively proportional fitting (ITF) methods, using demographic and socio-economic geographical data from the Dutch population in base year conditions (2000) and attribute data of a sample of households from a National survey including approximately 67,000 households. Next, activity schedules were assigned to each individual within this synthetic population using the scheduling process in ALBATROSS as described before. The activity-based approach hereby offers information on different facets of the individual trip like the trip purpose, the trip duration and the characteristics of the trip performer. Of course, when concentrating on the impact of personal travel on the environment, only the trips that were made as a car driver need to be selected from the trip dataset, other transport modes can be ignored. After the prediction of these activity-travel schedules, detailed Origin-Destination (O/D) matrices, presenting the trips per trip motive, were extracted from the simulated activity-travel patterns.

3.2 Traffic assignment

The O/D-matrices per trip motive, predicted by the activity-based model Albatross, represent the trip behaviour for 30% of the Dutch population. In a second step, these O/D matrices were multiplied with the reversed sample fraction and assigned to a transportation network using a standard traffic assignment algorithm embedded in the software package TransCAD. TransCAD is a GIS platform designed for use by transportation professionals to store, display, manage, and analyze transportation data [Caliper, 2004]. Approximately 120,000 traffic links were present in the Dutch road network and the region was divided into 1,308 traffic zones.

After the traffic assignment procedure, detailed information was present about the activity-related traffic flows on the road links and total distances were calculated for each trip motive. The predicted results were compared with reported values from the Dutch National Travel Study (NTS) to validate the model results. The travel results from the 2000 survey were used for comparison. This survey includes trip diaries of more than 100,000 persons and results were reweighed to compensate for the under –and over-representation of certain groups, e.g. degree of urbanization, age, and journeys. Although the NTS uses trip diaries (as opposed to activity diaries), this travel survey has analogue survey characteristics as the activity-based survey (no holiday trips, no freight trips, no vehicle kilometers in other countries, no vehicle kilometers of foreigners), and is executed yearly to gain more insights into the travel behaviour of the Dutch population. More information about the NTS in the Netherlands can be found on the Dutch Road Safety Website [SWOV, 2000].

3.3 Emission calculation

Finally, in the last step of the methodology, the activity-related traffic flows on the different traffic links were converted into vehicle exhaust emissions based on pollutant-specific emission factors (g/km).

The most common emission factors result from the COPERT/MEET methodology [MEET, 1999]. According to this approach, vehicle type-specific emission functions combine the average speed data with specific emission parameters to estimate vehicle emissions for a certain time period and a certain region. This is a standard procedure for transport emission inventories in Europe although the uncertainty of the result is often underestimated [Int Panis et al. 2001; 2004]. Instead of working with separate emission factors per vehicle type, some regions also apply vehicle fleet emission factors per pollutant, taking into account the composition of the vehicle fleet and the characteristics of the road network. Of vehicle fleet emission factors are available, only information about the travelled distance is needed to calculate the resulting vehicle emissions.

In this study, we used the vehicle fleet emission factors for the Netherlands in base year conditions. The emission factors for carbon dioxide (CO_2), carbon monoxide (CO), nitrogen oxide (NO_x), sulphur dioxide (SO_2) and particulate matter (PM) are presented in Table 1. To validate the emission results, emission values were compared with emission estimates from the Dutch Scientific Statistical Agency (CBS).

Table 1. Vehicle fleet emission factors for the Netherlands in the year 2000 (CBS, 2000)

	CO_2	CO	NO_x	SO_2	PM
Emission factor (g/km)	190	2.9	0.7	0.014	0.032

4. RESULTS

In this section, trip distances and emission results per trip motive are presented. When available, predicted results are compared with reported values to validate the model results.

4.1 Travelled distances

By aggregating the results from the traffic assignment procedure and extrapolating the values for a whole year, the total travelled distance during a whole year was calculated. In Table 2 the calculated value is presented next to the reported travelled distance value from the NTS, representing the total number of travelled kilometres by passenger car travel during the base year. The relative difference between both values is less than 10%.

Table 2. Total travelled distance by non-commercial vehicles in the Netherlands in the year 2000 [SWOV, 2000]

	Travelled distance (x 10^9 km)
Modelled travelled distance	93.3
Reported travelled distance	85.9
Relative difference (%)	8.6

After segregating the trips and the corresponding travelled distances according to trip motive, the annual number of vehicle kilometres per trip motive could be determined. To compare the predicted results with reported travel values, trips were classified according to the following seven NTS trip motives: work, school, shopping, leisure, social, service and other purposes (personal care, bring/get).

Figure 2 presents the travelled distances by trip motive for the activity-based method and the NTS method. Almost half of the total distance travelled is due to commuter purposes (more than 40 billion kilometres), followed by trips for social purposes (16 billion) and other purposes (12 billion). The activity-based approach predicts approximately the same amount of kilometres travelled for shopping and leisure trips (10 billion), and only a small amount of kilometres covered for school trips and services (1.5 billion).

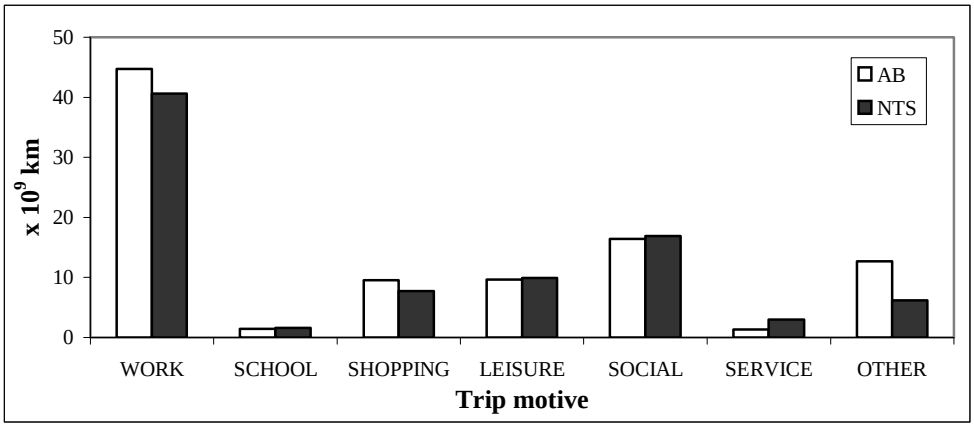


Figure 2. Total travelled distances per trip motive for passenger car trips in the year 2000: predicted activity-based (AB) values vs. reported NTS values [SWOV, 2000]

The results for the activity-based approach and the NTS approach seem to correspond well. The travelled distances for commuter trips, shopping trips and other trips are only slightly overestimated by the activity-based approach. On the other hand, school trips, leisure trips and trips with social purposes are slightly underestimated by the activity-based approach.

4.2 Vehicle emissions

Table 3 presents the total emission estimates per pollutant when the calculated travelled distance is multiplied with the corresponding fleet emission factor (see Table 1). These results were compared to the reported emission values for the year 2000 [CBS, 2000]. Model results for most pollutants seem to correspond well to the reported values. Relative differences between predicted and reported values fluctuate between approximately 2 and 9%.

Table 3. Calculation of total vehicle emissions for the year 2000, using vehicle fleet emission factors per pollutant [CBS, 2000]

	CO_2	CO	NO_x	SO_2	PM
Modelled emissions in 2000 ($x10^6$ kg)	17,729.06	270.60	65.32	1.31	2.99
Reported emissions in 2000 ($x10^6$ kg)	17,346.00	263.60	60.10	1.26	2.88
Relative difference (%)	2.21	2.66	8.69	3.97	3.82

As with the travelled distances, vehicle emissions can also be classified according to trip motive. By means of example, Figure 3 presents the contribution of different trips to the total amount of PM emissions. Calculations for other emissions yield similar results. As can be expected based on the travelled distance results, the largest amount of PM emissions is due to commuter trips: 47% of the total amount of PM emissions is attributed to work-related trips. Further, 17% of the emissions is caused by trips with social purposes, 20% is due to shopping and leisure trips and only a small percentage is caused by school trips and services-related trips. Trips with 'other' purposes still have a quite large contribution in the total PM emissions of approximately 13%.

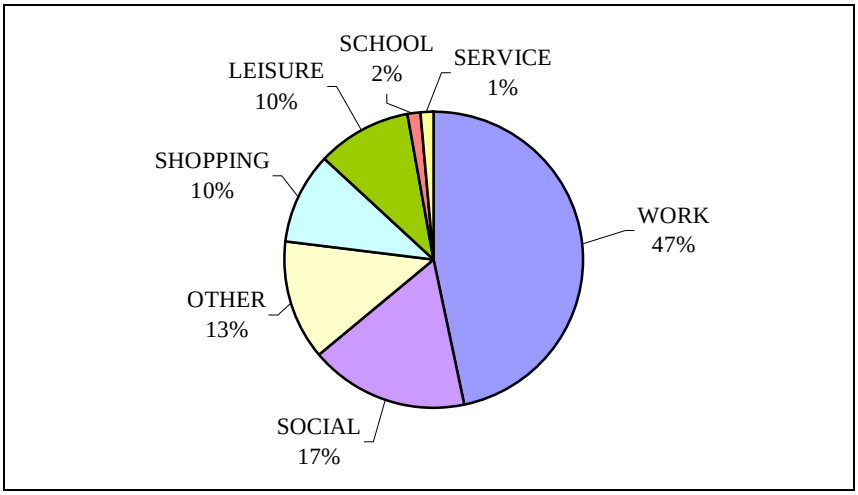


Figure 3. Activity-related PM emissions. Presentation of the relative contribution of different trip motives to the total amount of PM emissions

5. DISCUSSION AND CONCLUSION

In this paper we report on the use of an activity-based model to assess the activity-related distances and emission values for passenger car travel in the Netherlands. The activity-based model ALBATROSS was used to evaluate the travelled distance and the emissions of passenger car trips. First total travelled distance and total emissions were calculated and compared with officially reported values. Predictions for total travelled distances and total emissions of CO_2, CO, NO_x, SO_2 and PM differed not more than 8% from the reported values. Next, the trips (and the ensuing emissions) were classified according to trip motive to provide information about the contribution of different trips to traffic air pollution. By means of example, the results for the emissions of PM were presented in this paper. The commuter trips produced almost half of the total amount of PM emissions. Further, shopping and leisure trips together produced approximately 20% of the PM emissions, trips with a social purpose caused 17% of these emissions and school trips and service-related trips only caused a small percentage of the PM emissions. Calculations for other emissions yield similar results.

Of course, there are some qualifications to our research that need to be discussed. First, one can argue the validation method in itself. The predicted results from the activity-based emission modelling approach were compared with national travel and emission values from the Dutch Scientific Statistical Agency whose data originates from other surveys and model simulations. A good agreement between both values does not automatically indicate a good representation of the real situation, and only states the similarity between both models. Ideally, a validation method should comprise the use of measurements instead of simulation values, but the procedure of comparison with other models has been adopted by several authors for lack of a better alternative [e.g. Int Panis et al. 2006; Schrooten et al. 2008].
Next, the use of vehicle fleet emission factors can be a subject for discussion. The vehicle fleet emission factor takes into account the characteristics of the total vehicle fleet and general driving conditions, but takes no notice of local conditions. Since local road conditions and speed limits can significantly alter vehicle emissions [Int Panis et al. 2006; Beckx et al. 2007b], one should be careful when applying this approach for local policy questions. On the other hand the choice for a simplified emission calculation procedure with a constant emission factor also has some advantages. It is fast and provides us with a general evaluation of the impact of a policy measure on emissions. Nevertheless, future research should examine the use of more realistic emission factors to this extent. Leisure trips conducted in off-peak conditions will involve different emission factors than work trips occurring in the rush hour. Future research should therefore take into account trip departure times and locations to differentiate the use of emission factors for different trip motives.

The validation test in this research was an essential first step: if a model is unable to replicate its base year behaviour, it has little hope of forecasting the future adequately. Based on the results of this research we can conclude that the activity-based modelling approach is able to reproduce base year conditions with sufficient accuracy. Since the activity-based approach allows for impact analysis of other, new, policy measures (see e.g. Beckx et al. 2005), the use of an activity-based travel model will certainly put a new perspective on the research of source-related policy measures. Moreover, due to the fact that this approach provides information about the activity-related emissions, policymakers will gain more insights into the impact of certain policy measures on different trips. Future studies will certainly comprise the use of an activity-based model for the evaluation of different policy measures on travel behaviour and vehicle emissions and will also take into account the use of other transport models, e.g. public transport, for different trip motives.

ACKNOWLEDGEMENTS

In this research data have been used from NTS (2000), which are available through the Dutch Scientific Statistical Agency.

REFERENCES

Arentze, T., Hofman, F., Timmermans, H., Reinduction of ALBATROSS Decision Rules with Pooled Activity-Travel Diary Data and an Extended Set of Land Use and Cost-Related Condition States, *Transportation Research Record*, 1831, 230-239, 2003.

Arentze, T., Timmermans, H., ALBATROSS: A Learning-Based Transportation Oriented Simulation System, *Transportation Research Part B*, 38, 613-633, 2000b.

Arentze, T., Timmermans, H., *ALBATROSS 2.0: A Learning-Based Transportation Oriented Simulation System*, European Institute of Retailing and Services Studies, Eindhoven, The Netherlands, 2005.

Beckx, C., Broekx, S., Janssens, D., Activity-based policies to reduce human exposure to traffic air pollution, *Proceedings of the 32nd International "Transportation Research Colloquium*, Antwerp, Belgium, Volume 7, 1955-1972, 2005.

Beckx, C., Int Panis, L., Vanhulsel, M., Wets, G. and Torfs, R., Gender-linked disparity in vehicle exhaust emissions? Results from an activity-based survey, *Highway and Urban Environment*, Morrison, G.M, and Rauch, S. (Springer, The Netherlands) 63-70, 2007a.

Beckx, C., Int Panis, L., De Vlieger, I., Wets, G., Influence of gear changing behaviour on fuel-use and vehicular exhaust emissions, *Highway and Urban Environment*, Morrison, G.M, and Rauch, S. (Springer, The Netherlands) 45-51, 2007b.

Bhat, C.R., Guo, J.Y., Srinivasan, S., Sivakumar, A., A comprehensive micro-simulator for daily activity-travel patterns, *Proceedings of the Conference on Progress in Activity-Based Models*, Maastricht, 2004.

Caliper Corporation, Travel Demand Modelling with TransCAD 4.7, Caliper Corporation, Newton, Massachusetts, 2004.

CBS, Emissies en Parkemissiefactoren (Emissions and Fleet Emission Factors), Analyses from the Dutch Scientific Statistical Agency, http://www.cbs.nl, 2000.

EEA, Environmental Indicators: Typology and Overview, Technological Report No 25, Copenhagen, 1999.

Ettema, D., Timmermans, H., Theories and models of activity patterns, *Activity-based approaches to travel analysis*, Ettema, D., and Timmermans, H. (Pergamon, Oxford), 1-36, 1997.

Int Panis L., De Nocker L., De Vlieger I., Torfs R., Trends and uncertainty in air pollution impacts and external costs of Belgian passenger car traffic, *International Journal Of Vehicle Design*, 27(1-4) 183-94, 2001.

Int Panis L., De Nocker L., Cornelis E., Torfs R., An uncertainty analysis of air pollution externalities from road transport in Belgium in 2010, *Science Of The Total Environment*, 334-35 287-98, 2004.

Int Panis, L., Broekx, S., Liu, R.H., Modelling instantaneous traffic emission and the influence of traffic speed limits, *Science of the Total Environment*, 371 (1-3) 270-285, 2006.

MEET, Methodology for calculating transport emissions and energy consumption, Transport Research, Fourth Framework Programme, Strategic Research, DG VII, ISBN 92-828-6785-4, 1999.

Pendyala, R.M., Kitamura, R., Kikuchi, A., Yamamoto, T. and Fujji, S., FAMOS: Florida activity Mobility simulator, *Proceedings of the 84th Annual Meeting of the Transportation Research Board*, Washington D.C., January 9-13, 2005.

Ruiter, E.R., Ben-Akiva, M., Disaggregate travel demand models for the San Francisco bay area, *Transportation Research Record*, 673, 121-128, 1978.

Schrooten, L., De Vlieger, I., Int Panis, L., Styns, K., Torfs, R., Inventory and forecasting of maritime emissions in the Belgian sea territory, an activity based emission model, *Atmospheric Environment*, 42(4), 667-676, 2008

Shiftan, Y., The Advantage of Activity-based Modelling for Air-quality Purposes: Theory vs Practice and Future Needs, *Innovation*, 13(1), 95-110, 2000.

Spear, B.D., New Approaches to Transport Forecasting Models, *Transportation*, 23, 215-240, 1996.

SWOV, Results from The National Travel Survey, The Netherlands, http://www.swov.nl/cognos/cgibin/ppdscgi.exe?toc=%2FEnglish%2FMobility, 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

A storm surge model implementation and identification of coastal areas in risk of inundation, in the Mediterranean Sea.

Y. Androulidakis, Y. N. Krestenitis, Y. Kontos and G. Georgakopoulos

Laboratory of Maritime Engineering & Maritime Works, Department of Civil Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece (iandroul@civil.auth.gr)

Abstract: Low-elevation coastal areas and their populations are at risk during and after the appearance of a storm surge event. Coastal flooding as a result of storm surge events is examined in this paper for a number of areas around the Mediterranean Sea. A 2-dimensional hydrodynamic model has been implemented in the region, calculating the sea level altimetry in a $1/10^{o}$x$1/10^{o}$ grid. The sea level rise due to storm surge events is examined for the period 2000-2004. In situ measurements were collected from stations around the Mediterranean coasts, which were compared with the simulation results and several parameters of the model have been tested and calibrated. Potential inundation zones were then identified using a 90-m resolution digital elevation model (DEM). At these zones, the sea-level alterations were calculated for the study period, where the area affected by sea level rise of 1 m for various regions was estimated. Moreover, the implication of storm tracks in major sea level rise incidents in the Mediterranean was investigated for the studied period. The computation and plotting of storm tracks, was accomplished using an algorithm that tracked down sea level pressure (SLP) minima, recording their development.

Keywords: storm surge, modelling, coastal inundation, Mediterranean Sea

1. INTRODUCTION

Low-elevation land areas and their populations are at risk during and after the appearance of a storm surge event above or near coastal areas.

High sea levels and the strong forces exerted by accompanying waves, impact directly or by over-topping of sea defenses on humans, property and habitats. They may even cause loss of life, damage (through inundation and waves), loss of habitat and useful land, property, infrastructure, services and so forth. Traveling mid-latitude low pressure systems act to raise the sea level directly below them, but this effect alone is quite weak in semi enclosed basins such as the Mediterranean Sea [Pirazzoli, 2000]. The most important meteorological factors are the associated winds, turning anticlockwise in the northern hemisphere. These winds tend to drag the water in the same direction, with a deflection to the right due to the Coriolis force. Although, the Mediterranean Sea is not on the main storm track of the European and North Atlantic area [Rogers, 1997], storm track events originated mainly from Africa, with a direction from south to north, affect significantly local sea-level alteration.

This paper describes the major storm surge events that appeared in the Mediterranean region in additions with the inundation areas in danger for the period 2000-2004. Numerical modeling, storm tracks computation and plotting, in situ sea level measurements, satellite images and a digital elevation model were used for this purpose.

2. STORM SURGE MODEL AND SIMULATION OF SEA LEVEL ALTERATION

2. 1 Introduction

A 2-dimensional hydrodynamic model has been implemented in the region, calculating the sea level altimetry in a $1/10^{o}$x$1/10^{o}$ grid (Figure 1). The sea level rise due to storm surge events is examined for the period 2000-2004. Wind data, atmospheric pressure (SLP) and wave data [Soukissian et al., 2007] are the model's forcing. Several parameters of the model have been tested and calibrated, such as the surface friction coefficient and sea bed roughness coefficient.

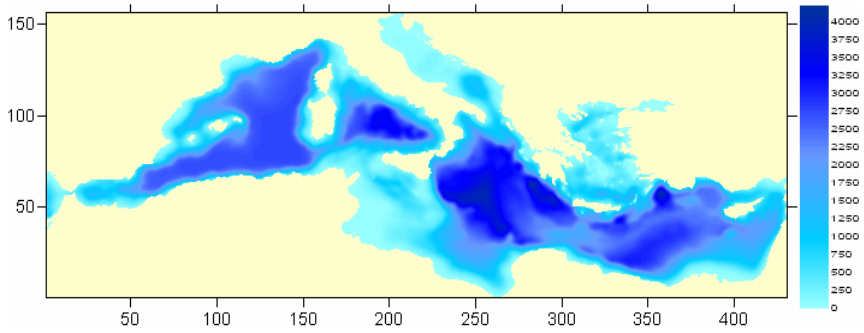


Figure 1. Storm surge model domain and bathymetry

2.2 Model Description

The basic equations of the model are the depth-averaged shallow water equations:

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}-fu+g\frac{\partial z}{\partial x}=-\frac{1}{\rho_o}\frac{\partial P}{\partial x}+\frac{1}{\rho_o}\frac{\tau_\chi}{(h+z)}-k\frac{u\sqrt{u^2+v^2}}{\rho_o(h+z)} \quad (1)$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+fu+g\frac{\partial z}{\partial y}=-\frac{1}{\rho_o}\frac{\partial P}{\partial y}+\frac{1}{\rho_o}\frac{\tau_y}{(h+z)}-k\frac{u\sqrt{u^2+v^2}}{\rho_o(h+z)} \quad (2)$$

$$\frac{\partial z}{\partial t}+\frac{\partial (h+z)u}{\partial x}+\frac{\partial (h+z)v}{\partial y}=0 \quad (3)$$

where t is time, (x,y) are spatial coordinates, z is the water level elevation, u, v are the x and y components of the depth-mean current, h is the undisturbed water depth, g is the acceleration of gravity, f is the Coriolis parameter, k is the bottom friction coefficient, τ_x and τ_y are the x and y components of the wind stress, ρ_o is the density of the water and p is the atmospheric pressure at sea level.

The calculation of the climatological wind stress fields is based on the transformation wind velocity data at 10m, to the zonal/meridional components of the wind stress exerted on the sea surface, according to the formula

$$\tau = \rho_A C_D |W| W \quad (4)$$

where ρA is the air density, W is the wind velocity and CD is the surface friction coefficient. Experiments with various values of the surface friction coefficient were executed, including the calculation, according the formula [Smith & Banke, 1975]:

$$10^3 CD = 0.63 + 0.066 W \quad (5)$$

Tidal boundary conditions where imposed in the Gibraltar open boundary by using specific tidal components measured in area's stations and the tidal harmonic analysis:

$$H = 0.01 Ho_j \cos (F_j dt 2\pi - B_j \pi / 180 \quad (6)$$

where F_j, Ho_j, B_j are area's tidal measured components and dt is the model's time-step.

The contribution of the wind factor, the wave factor, the atmospheric pressure factor and the tidal factor at the sea elevation field was examined both separately and simultaneously. Additional characteristics of the model are given in Table 1 [De Vries et al., 1995].

Table 1. Storm surge model characteristics

Grid	Depths	Bottom Friction	Coastal Boundary Condition	Advection	Integration Scheme	***Time Step (sec)***
C	h at ζ	Quadratic Manning $N=E_b^2 g/H^{1/3}$	No normal flow	Smagorinsky	Explicit leap frog	30

2.3 Additional data

A thorough search of sea level data from gauge stations around the Mediterranean was done for the period 2000 – 2004. Data were collected from the websites of the Med-GLOSS program and the European Sea-level Service. Moreover, sea level measurements have been gathered for 5 stations in the Greek region, regarding the years 1992 until 2004 showing the mean sea level alteration and the extreme values and dates that these values have been appeared. Additional sea-level data were collected by using satellite data concerned the study period and area. These altimetry data were collected from the French space agency (Aviso/Altimetry project) and regard the composition of various scans of different satellite projects (Topex/Poseidon, Jason-1, ERS-1 and ERS-2, EnviSat, and Doris) offering continuous recording of the study area

3. STORM SURGE DETECTION AND STORM-TRACK CALCULATION

Literally, storm tracks can be identified by following low pressure centers on synoptic charts and plotting their trajectories on maps, thereby producing ‘‘cyclone tracks’’ in the pure sense. A method for automated detection and tracking of storms or cyclones is also presented by the ‘National Aeronautics and Space Administration - Goddard Institute for Space Studies’ (NASA-GISS) (see reference) and a similar simplified methodology is used in this study.

3.1 Basic characteristics of the method used

The computation and plotting of storm tracks, is accomplished by using an algorithm that identifies and tracks sea level pressure (SLP) minima. The algorithm searches for and identifies absolute minima from the gridded field of every 12-hour period of the desired year. The minimum for the ensuing 12-hour SLP grid is searched, and its position is again located. Any SLP minima within a critical radius of 1440 km are joined by a segment, representing the path of the low pressure center during that 12-hour period (a cyclone's center can travel at a mean speed of no more than 120 km/hr or 1440 km / 12 hr). Any two associated minima identify one storm track segment, as long as the storm lasts at least 36 hours. If at any time two segments on the same track are found to define an acute angle of less than 85°, the low pressure centers are considered to represent separate storms (extra tropical cyclones are not found to "double back" on themselves over 12 hours). Finally, throughout the full duration of a storm track, SLP in its center is required to be less than 1015 hPa (low pressure means SLP<1015 mbars).

In this way, storm tracks are computed and plotted on a Mediterranean map. The use of absolute minimum SLPs instead of local minima (like NASA’s method) produces the risk of not identifying some storm tracks generated by the passage of secondary low pressure centers. For that, but also for presentation reasons and visual confirmation, a contour map for every 12-hour period is plotted, creating a slideshow of 730 maps (732 for leap years) for every year, presenting low pressure regions (SLP<=1015mBars) and their development.

3.2 Correlation procedure

The full understanding of the magnitude of the impact a storm track may have to the sea level height (SLH) of adjacent areas, also requires studying the relationship between SLP and SLH in the Mediterranean basin in general and in specific areas of interest in particular. The *Pearson product-moment correlation coefficient* (r) [Pearson, 1929 and Pearson, 1931], first introduced by Francis Galton [1988], is used to document the relationship

between 12hourly mean SLP and 12hourly max SLH measured in several gauge stations around the Mediterranean basin. The correlation coefficient indicates the strength and direction of a linear relationship between two random variables and usually refers to the departure of two variables from independence. The correlation coefficient between two properties x and y ($r_{x,y}$), which is commonly used in similar studies [Moron, Ullmann, 2005], is given below:

$$r_{x,y} = Cov(X,Y) / \sigma_x \cdot \sigma_y \quad (7)$$

where $-1 \le r_{x,y} \le 1$

and $$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_x)\cdot(y_i - \mu_y) \quad (8)$$

where μ_x, μ_y are the averages of x, y respectively.

The correlation is 1 in the case of an increasing linear relationship, −1 in the case of a decreasing linear relationship. The closer the coefficient is to either −1 or 1, the stronger the correlation between the variables. For every gauge station in the eastern Mediterranean , for each year between 2000 and 2004, the correlation coefficient between the mean 12-hourly SLP and the max 12-hourly of observed SLH is computed. On Table 2 the SLP-SLH correlations are presented for every gauge stations, for the years that SLH timeseries exist. Then, the local SLH maxima are searched and every one of them is considered as day 0. A six-day window is considered preceding the day of local max SLH appearance (time lags 5, 4, 3, 2, 1 and 0).

For each "6-day window" the SLP-SLH correlation coefficient is again computed and it is compared to the annual correlation.

Table 2. Mean 12hr SLP–max 12hr SLH corr. coeff. at Mediterranean gauge stations (x(-1))

Station	'00	'01	'02	'03	'04
Ancona(IT)	0.63	0.43	0.61	0.58	0.55
Antalya(TR)	-	0.44	0.50	0.55	0.62
Alexandroupoli (GR)	-	-	0.62	-	0.62
Barcelona(ES)	0.77	0.55	0.70	0.64	-
Catania(IT)	0.59	0.26	0.56	0.52	0.56
Chios(GR)	-	-	0.51	0.37	0.62
Dubrovnik(CR)	0.59	0.36	0.52	0.53	0.56
Genova(IT)	0.73	0.59	-	-	-
Lefkas(GR)	-	-	0.43	0.27	0.58
Naples(IT)	0.61	0.41	0.57	0.58	0.57
Otranto(IT)	-	0.25	-	-	0.57
Rovinj(HR)	0.57	0.41	0.55	0.52	-
Split(HR)	0.59	0.38	0.52	0.47	0.54
Trieste(IT)	-	0.43	-	-	0.45
Zadar(HR)	0.62	0.40	0.58	0.56	0.55

4. RESULTS

Sea surface elevation timeseries data from several stations around the Mediterranean are compared with model output showing good performance of the model in a qualitative level. Especially in areas characterized as "in danger" and in areas that in situ measurements exist, model output of SLH was produced, in order to check the possibility of flooding and to test the performance of the model, respectively. Two characteristic cases are presented below.

A storm track (15th of 2001) appeared on 1st of March, 2001 (Figure 2) affecting the sea surface elevation in western Mediterranean. Though, the storm's center does not move directly above the affected coastal areas (523 km NW of Barcelona gauge station at the time of max SLH, afternoon of 2nd of March, 2001), the low pressure system's extend (Figure 3) and its strength (SLP=981 hPa in the center and 994 hPa above Barcelona station, lowest 12hr mean value for 2001), affect the SLH of the adjacent areas.

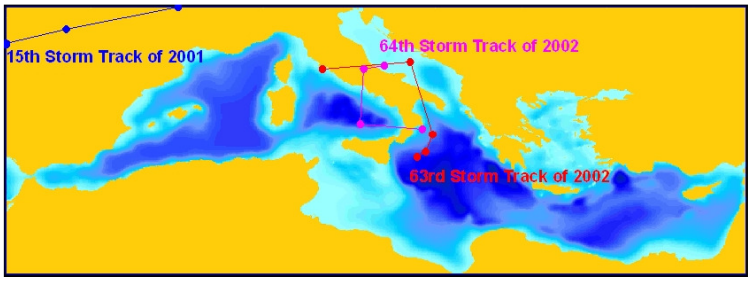


Figure 2. Storm tracks 15, 63 and 64 related to SLH- rise in west and central Mediterranean respectively

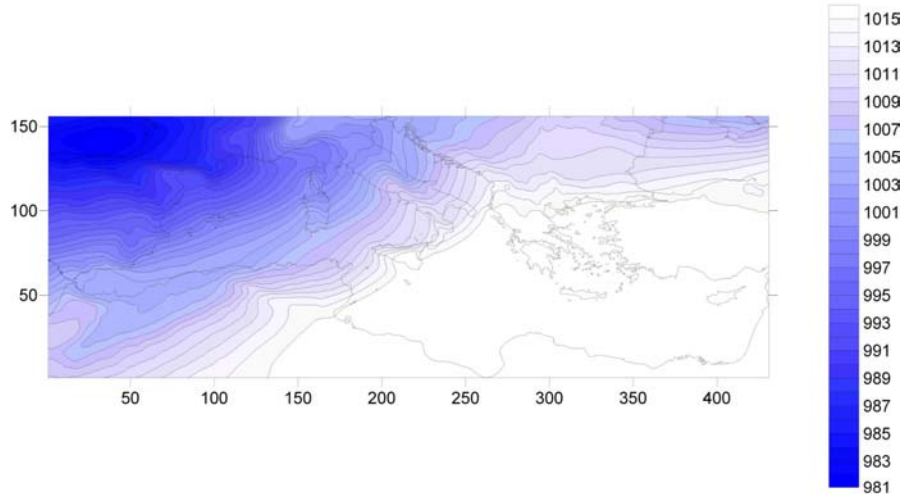


Figure 3. Contour map presenting an atmospheric low pressure system above western Mediterranean on the afternoon of the 2nd of March , 2001.

In Figure 4, the modeling and in-situ 2001 sea elevation timeseries are presented for Barcelona. High SLH values appeared (>0.4m) at the beginning of March (15th storm track of 2001) and in mid November (highlighted in red circles).

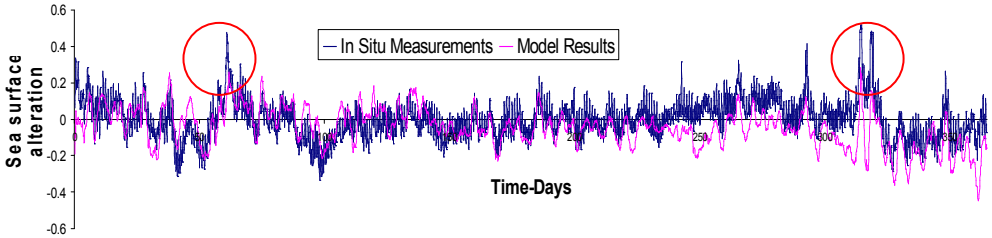


Figure 4. Barcelona model and in-situ surface elevation during year 2001

Table 3 represents Barcelona year 2001 max SLH analysis, from which useful conclusions can be made about the storm tracks' relationship with big sea level heights. The SLH-SLP Barcelona correlation for the year 2001 is r=-0.55 regarding the in situ measurements (Figure 5a) and r=-0.51 regarding the SLH model results (Figure 5b). As expected, the correlation for the in situ values is higher, but the small variance between the two values confirms the good model performance. Likewise, the correlation for the '6-day-window' of the SLH-rise incident described above (relevant to the 15th storm track 2001) r=-0.84, significantly greater than the annual one.

The case studied here is 2nd on the list of the highest observed SLH values in Barcelona station, with a sea level height of 40.7cm, displaying the lowest SLP of all. Most of the incidents are related with the appearance of a storm track nearby, while one could argue that the more the implication of a storm track in the rise of SLH (smaller distance, lower sea level pressures, greater extend of low pressure system), the closer the correlation to -1.

Table 3. Barcelona 2001 max SLH and storm track implication analysis.

Gauge station		**year**	**r**	**SLP**		**SLH**	**SLH annual average**	
BARCELONA		**2001**	**-0.55**	**12hr mean**		**12hr max (observed)**	**0 cm**	
Incident	**SLH (cm)**	**SLP (hPa)**	**r**	**day**	**12hr**	**Storm track**	**storm distance (km)**	**storm direction**
1	45.7	1006.923	-0.45	314	628	65 (627-631)	708	S-N
2	40.7	994.088	-0.84	61	122	15 (120-123)	523	S-N
3	39.7	1022.09	+0.14	319	638	-	-	W-E
4	35.7	1006.493	-0.87	292.5	585	62 (584-587)	776	SW-NE
5	26.7	1002.987	-0.88	265.5	531	55 (528-531)	348	-
6	25.7	1014.609	-0.14	242.5	485	-	-	W-E
7	25.7	1010.475	-	1.5	3	1 (1-4)	802	S-N
8	24.7	1007.632	-0.90	66	132	16 (128-134)	802	-

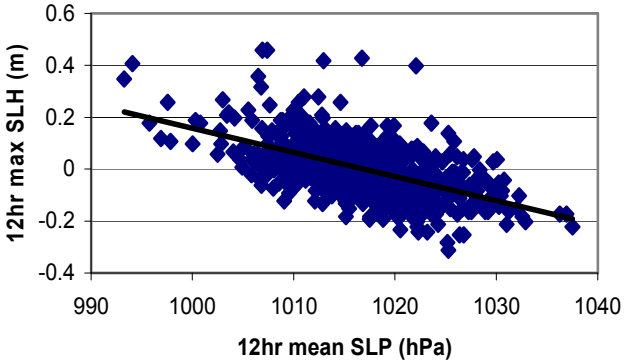


Figure 5a. SLH – SLP correlation for Barcelona gauge station, year 2001

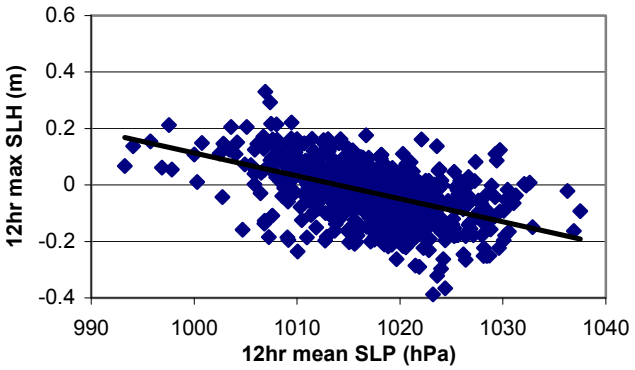


Figure 5b. SLH – SLP correlation for Barcelona model output, year 2001

Another interesting case ascertained at the beginning of December 2002, where the successive storm tracks 63 and 64 (Figure 2) affected the sea level elevation in several areas of central and eastern Mediterranean (Figure 6). One could support that it is, in fact, a single, integral storm track that projects an anti-clockwise rotation of the low pressure center around Italy. The existence of two separate successive storm tracks is suggested by the method because during the studied 12-hour period the storm seems to 'double-back' on itself. This criterium applies mostly to extratropical cyclones, while this one is obviously restrained within the Mediterranean basin, hence, it is safe to assume that storm tracks 63 and 64 are, in fact, two phases of the same integral storm surge. The events presented in Figure 6 represent the transition from the first phase (storm track 63) to the second (64) of the aforementioned unified storm surge event.

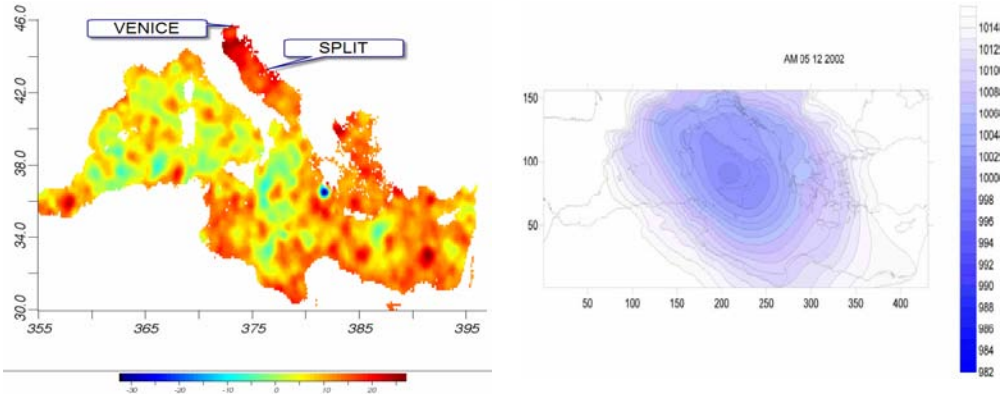


Figure 6. Satellite sea-level altimetry of the Mediterranean on the 4th of December 2002 and 2 specific affected areas (left). An SLP contour map showing the spatial extension of the low pressure system (right).

In Figure 7, the alteration of sea level is presented for 2 areas in Adratiatic Sea, as calculated for the period of the 63 storm track. The highest SLH values were detected in the Adriatic Sea, where the low pressure centre (min SLP=998 hPa) clouds over for days.

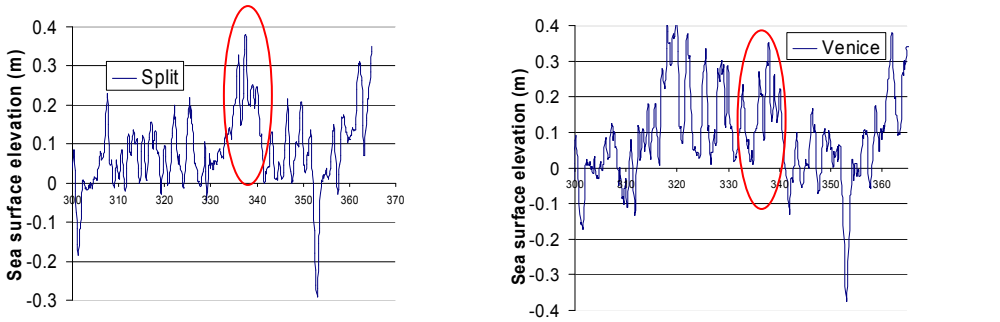


Figure 7. Model's SLH – Time output for Split and Venice on 4th of December, 2002

Potential inundation zones were calculated with the 90-m resolution shuttle radar topography mission (SRTM) digital elevation model (DEM). SRTMDEM has spatial resolution of 90 m, with horizontal and vertical accuracies of 45 and 15 m, respectively [Sun et al. 2003]. The absolute vertical accuracy of the SRTM DEM is about 15 m, and the relative accuracy to the coastline is less than 1 m. The resolution of the SRTM DEM is, therefore, satisfactory to determine vulnerable coastal areas.

Areas like the Venice Lagoon and its city Venice, are very sensitive to water level variations, given that the city of Venice is only about 1 meter above mean sea level and the medium depth of the lagoon is about 1 meter. The area where the land elevation is lower than 1m, produced by the 90-m elevation model is presented in red in Figure 8. The city of Venice often experiences flooding events induced by abnormally high water levels at the three entrances of the lagoon. Water levels significantly higher than the expected astronomical tide level are reported several times a year, especially between October and January [Bargagli et al, 2001]. So, events like the one of the 4th of December could affecte significantly all the coastal zone around the lagoon.

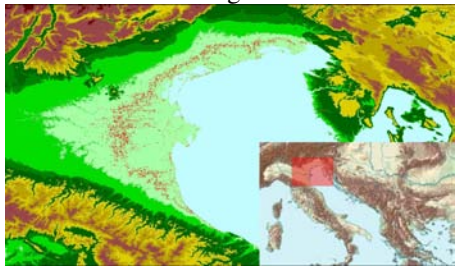


Figure 8. Map of Venice Lagoon of coastal vulnerability to sea-level rise up to 1m.

CONCLUSIONS

The atmospheric pressure influence is an important factor of simulation performance and hence at modelling results. Additionally, the influence of the wind stress is also significant. As far as the storm track identification method is concerned, the fact that the algorithm tracks absolute SLP minima (instead of local minima NASA proposes) produces the risk of not identifying some storm tracks generated by the passage of secondary low pressure centres. On the other hand, the visualization of low pressure systems and the correlation coefficient calculation strengthen the efficiency of the above method. Visualization of secondary lows, covers the gaps that the storm track identification method produced, while the correlation coefficient calculation and the relation between the '6-day-window' and the annual SLP-SLH correlation coefficient indicates possible errors in relating high sea level alterations to storm surge events. Modelling the sea-level alteration due to storm surge events can be employed as a common prediction method of SLH-rise incidents all over the Mediterranean. So, the combination of atmospheric forecasting modelling with a storm surge hydrodynamic model can estimate with good accuracy the near future storm surge events and their accompanied sea level alterations. Combined with a method of storm tracks identification and observation of their dominant characteristics together with the detection of coastal areas in risk of inundation, a useful tool, in terms of generating reasonably based scenarios of fore coming flooding events in the coastal areas of the Mediterranean, can arise. Warning systems of coastal flooding improve authorities' preparedness and help coastal human society to obtain proper measurements in cases of near future-coming extreme meteorological incidents. In sustainable and integrated development of the coastal zone, a storm surge event is a crucial factor, affecting significantly the quality of life. Additionally, hind casting storm surge events and detecting previous inundation cases can offer to stakeholders a better view of possible changes and corrections that need to be done in already developed risky coastal areas.

REFERENCES

Archiving, Validation and Interpretation of Satellite Oceanographic data - Aviso/Altimetry project, http://www.aviso.oceanobs.com

Bargagli, A., A. Carillo, G. Pisacane, P.M. Ruti, M.V. Struglia, and N. Tartaglione, An Integrated Forecast System Over The Mediterranean Basin: Extreme Surge Prediction In The Northern Adriatic Sea, Volume Monthly Weather Review 130,2002.

De Vries, H., M. Breton, T. Mulder, Y. Krestenitis, J. Ozer, R. Proctor, K. Ruddick, J.C. Salomon and A. Voorrips, A Comparison Of 2D Storm Surge Models Applied To Three Shallow European Seas, Environmental Software, Vol. 10, No. 1, Pp 23-42, 1995.

Moron, V., A. Ullmann, Relation between sea-level pressure and sea-level height in the Camargue (French Mediterranean coast), International Journal of Climatology 25, 1531–1540, 2005.

National Aeronautics and Space Administration - Goddard Institute for Space Studies (GISS), http://data.giss.nasa.gov/stormtracks.

Smith, S.D. and E.G. Banke, Variation Of The Sea Surface Drag Coefficient With Wind Speed. Quart. J. Roy. Meteor. Soc., 101, pp. 665–673, 1975.

Soukissian, T., M. Hatzinaki, G. Korres, A. Papadopoulos, G. Kallos and E. Anadranistakis, Wind and Wave Atlas of the Hellenic Seas, Hellenic Centre for Marine Research Publ., 300pp, 2007.

Sun, G., Ranson, K.J., Kharuk, V.I., and K. Kovacs, Validation Of Surface Height From Shuttle Radar Topography Mission Using Shuttle Laser Altimeter, Remote Sensing Of Environment, 88, 401–411, 2003.

Rogers, J.C., North Atlantic Storm Track Variability And Its Association To Both North Atlantic Oscillation And Climate Variability Of Northern Europe, J. Climate, 10, 1635-1647,1997.

Pearson, E.S., Some Notes On Sampling Tests With Two Variables, Biometrika, 21, 337-360,1929.

Pearson, E.S., The Test Of Significance For The Correlation Coefficient. J. Amer. Statist. Ass., 26. 128-134, 1931.

Pirazzoli, P.A., Surges, Atmospheric Pressure And Wind Change And Flooding Probability On The Atlantic Coast Of France. Oceanologica Acta, 23, 643-661, 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

Development of a Web based Decision Support System for trans-boundary environmental risk assessment and management.

Pediaditi K[a]., Stanojevic M[b]., Kouskouna, M[c]., Karydas, C[a]. Zianis D[a]., Boretos, N[d].

[a] *Department of Environmental Management, Mediterranean Agronomic Institute Chania, Chania, Crete, 73100, Greece.(pediaditi@maich.gr, xkarydas@maich.gr,zianis@maich.gr)*

[b] *Mihailo Pupin Institute, Volgina 15, 11060 Belgrade, Serbia (mstanojevic@institutepupin.com)*

[c] *Faculty of Law, University of Athens (EKPA)(mkousk@law.uoa.gr)*

[d] *Department of Information Systems and Technology, MAICh, Chania, Crete, 73100, Greece (nikos@isoteia.org)*

Abstract: A review of the legal requirements for environmental risk assessment indicates that despite the range of international and EU legislation and guidance proposing the need for risk assessment and management a common framework and procedure has yet to be established. This in turn has limited the implementation of environmental risk assessment and management mainly to fields where liability issues can arise such as in cases of contamination hazards etc. In order to enable the formalized and more systematic utilization of risk assessment and management procedures in environmental decision making processes the "*Remotely Accessed Decision Support System for Transnational Environmental Risk Management*" (**STRiM**) project has developed a web based Decision Support System (DSS) for transboundary environmental risk assessment and management. This DSS is implemented as a web based application which enables environmental administrators and decision makers to undertake generic risk assessment and management identifying areas where detailed risk assessment is required as well as appropriate risk management options. The web based DSS serves as an integration platform for the other components of the system: map server, Risk Assessment indicator database and document database. The main functionality of the web based DSS includes qualitative risk assessment for the identified environmental hazards, MCA facilitated options appraisal for the selection of applicable risk management options and automated report generation. The map server is used to handle GIS maps that illustrate various steps of the risk assessment procedure. The web based DSS is implemented using a three-tier architecture, Java EE 5 platform, Java Server Faces and Java Enterprise Beans technologies, and Open Source products for the application server, database and integrated development environment. Within the context of the STRiM project the web application has been trailed successfully in four pilot trials addressing a range of risks, such as forest damage from storms, water pollution from olive mill waste, wetland loss from water abstraction, and damage from flooding.

Keywords: Trans-boundary Risk; Decision Support Systems; Risk Assessment; Risk Management

1. INTRODUCTION

Awareness of the need to face the risks and challenges for balanced and sustainable development in Europe posed by natural and technological hazards in Europe is increasing. There are several elements in EU legislation, policies and programmes encouraging the

introduction of risk management into planning and decision making, (underpinned by the precautionary principle) yet this inclusion is far from complete and proving problematic [Peltonen, 2006; Kriebel et al 2001]. Currently, hazards are being addressed in heterogeneous, fragmented ways and at different levels by existing community instruments. Although the Seveso II Directive is the legislative basis for risk management of major accidents there is no uniform approach within the EU to deal with environmental risk in general, and subsequently no standardised risk assessment framework and decision support tool to enable its implementation, something the STRiM project attempts to address. In the EC adopted legislative proposals for cohesion policy reform (COM(2004) 492-496) risk prevention is mentioned as a priority under all the three objectives of convergence, regional competitiveness and employment and European territorial cooperation. In fact, the third priority territorial cooperation acknowledges the need for risk management at transnational and interregional level. In addition key measures, such as the Directives on Environmental Impact Assessment (EIA), on Strategic Environmental Assessment (SEA) including the Water Framework Directive (WFD) point out the need for the inclusion of risk assessment and management as well as the consideration of transboundary impacts, yet the process of doing this is not specified, subsequently limiting its implementation.

The STRiM project (Remotely Accessed Decision Support **S**ystem for **T**ransnational Environmental **Ri**sk **M**anagement) aims to address these issues by developing a planning relevant framework to enable generic, transboundary environmental risk assessment. Transboundary risks are defined as "*risks that are generated under one regulatory jurisdiction and have significant actual or anticipated impacts in another, regionally or globally, are a source of concern for regulators, politicians and public*" [Tait and Bruce, 2000, pp1]. The challenge is great, considering the acknowledged difficulties of transboundary or cross border assessment in general, such as technical limitations eg data incompatibility, social and political limitations such as soveregneity differing languages, cultures, levels of economic growth, varying regulatory structures, political systems etc [Craik, 2007].

Increasingly, literature attributes the lack of practical implementation of risk management procedures to the lack of science policy integration within existing decision making processes and the policy/decision maker and scientist conflict [Quevauviller et al 2005; Willems and de Lange, 2007] It is argued that policy makers and implementing authorities are in need of Decision Support Systems (DSS) in the field of environmental management which can enable science knowledge transfer simultaneously paying greater attention to potential user needs and to the identification of concrete application contexts [Giuponi, 2007]. Therefore, STRiM, as an INTERREG project whose primary aim is the building of bridges between scientific research and the praxis of planners and multiple other stakeholders [Peltonen, 2006] has focused on the development of a common framework and DSS to enable the integration and implementation of transboundary environmental risk management.

Prior to developing the DSS, the STRiM consortium conducted a literature review and end user interview survey to establish what existing environmental risk assessment DSS tools where available, their use, and limitations. It was established that there is no commonly adopted environmental risk assessment procedure within the EU, like there is for EIA and SEA and subsequently no common decision support tool. However, a web based review of existing risk assessment tools, indicated that there is a plethora of national, both commercial and public, domain specific risk assessment DSS tools. For example, only in Great Britain, regarding contaminated land, there are many used tools (eg CLEA, SNIFFER, RBCA, Risk Assistant, Risc-Human health). Additionally, from the interviews it was established that risk assessment as a procedure and concept is much less understood, and often confused with impact assessment. Recommendations to develop a generic/qualitative environmental risk assessment framework and supporting DSS in order to enable implementation were made.

2. THE STRiM FRAMEWORK AND DECISION SUPPORT TOOL

The STRiM project was developed based on the findings of its predecessor ISOTEIA project (INTERREG IIIB CADSES, 3B093). ISOTEIA concluded that risk management

decision makers and users are in need of a common framework to carry out environmental risk assessment and management at a transnational level as well as user focused decision support tools and guidance (Karydas et al., 2006). In particular end user confusion regarding the implementation of the concept of risk and utilisation of probability components in decision making proved to be an issue in its formalised uptake in environmental management. Based on the above, the STRiM DSS has been developed providing a common framework to carry out generic risk assessment and management, for any environmental risk, regardless of temporal and spatial boundaries.

The STRiM DSS was designed to adhere to the following specifications, which in turn influenced its nature.

a) User focused (simple, cost and time efficient)
b) Participatory (enabling multi-stakeholder collaboration and risk communication)
c) Generic (applicable to any environmental risk)
d) Transboundary (provides the opportunity for the assessment and management of risks regardless of administrative, geographical and political boundaries)

The generic/ qualitative nature and wide scope of this framework is recognised. Indicatively it can be used to inform a decision regarding whether to give planning permission to a small scale olive mill processing unit and similarly to a decision regarding the diversion of a transnational river. The target audience is wide focusing on practitioners, technicians and decision makers, not the scientific community. Typical applications foresee the involvement of a team of decision makers and technicians, including interested stakeholders (which can represent the perspectives of different countries and interests) assisted by the remotely accessed STRiM DSS to arrive at an informed decision regarding identification of hazards requiring detailed risk assessment as well as a prioritised list of the best risk management option to address a given environmental risk. Due to the generic nature of this tool, its quantitative component is undoubtedly limited, enabling users to inherently make subjective decisions using qualitative data. However, this does not undermine the value of this tool, as its primary aim is to introduce the concept and procedure of risk assessment in environmental decision making, providing indications regarding the need for more detailed quantitative risk assessments as well as risk management actions.

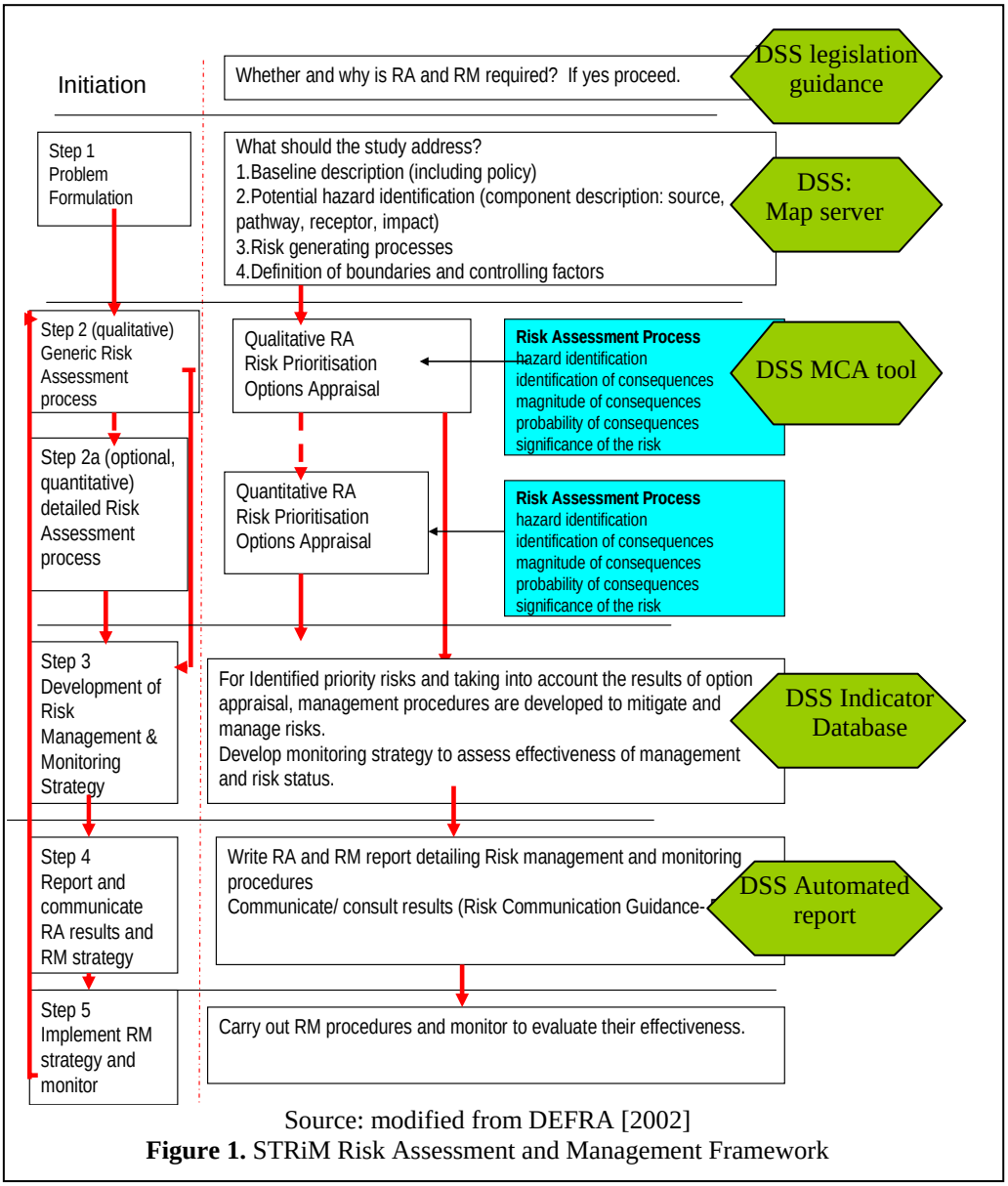
2.1 The STRiM Risk Assessment and Management Framework

A review of risk literature indicates a plethora of definitions of risk [Pediaditi et al , 2005] including of a range of risk assessment and risk management frameworks [Power and McCarty, 1998] none of which are commonly adopted, thus limiting their use in decision making. This issue is also valid for the CADSES region which this project focuses on. The STRiM RA & M framework consists of 5 iterative steps (Figure 1) and is predominantly based on the DEFRA [2002] guidelines due to their focus on risk management and applicability to any type of environmental risk. It needs to be emphasised that although the DEFRA [2002] guidelines theoretically elaborate on detailed risk assessment, the STRiM DSS application focuses on generic risk assessment and is qualitative in nature. Due to the user focused nature of this framework and the intention for its implementation in everyday environmental management decision making which is constrained by limited resources and time emphasis is placed in a) correctly defining the problem and prioritising risks before moving on to data collection, (commonly applied in EIA as scoping) b) considering risks taking into account potential management solutions from the onset c) enabling multi-stakeholder engagement and transparency in determining risk significance. The STRiM DSS as described below provides users with additional tools such as an indicator database for consultation during Step 3 (Figure 1), guidance on participation and risk communication for each step of the process as well as a compendium of potentially relevant environmental legislation (Initiation and Step 1, Figure 1). Users of the DSS having completed the STRiM RA and M process are provided with an automated report, of the procedure followed and results generated, i.e. a prioritised list of the most appropriate risk management options and significant hazards requiring further detailed risk assessment.

2.2 The STRiM software

Technically, STRiM DSS is implemented as a Web-based application enabling its users to create risk assessment and management case studies by carrying out the following five steps, on an anytime-anywhere basis.

- Step 0. **Initiation** consists of the starting point of any RA or RM, which entails the consideration of the need to conduct the assessment.
- Step 1. **Problem Formulation** is composed of four components: baseline description, potential risk identification and components description, identification of risk generating processes and definition of boundaries and controlling factors.
- Step 2. **Generic (Qualitative) Risk Assessment and Management Process** includes qualitative risk assessment of selected hazards and options appraisal of applicable management options.
- Step 3. **Risk Management** evaluates and discusses the preferred risk management options and monitoring strategy.
- Step 4. **Risk Communication** describes the risk communication procedure followed in the case study.
- Step 5. **Risk Monitoring** consist of the implementation of risk management options and monitoring of their effectiveness.



Source: modified from DEFRA [2002]

Figure 1. STRiM Risk Assessment and Management Framework

STRiM DSS facilitates an automatic Qualitative Risk Assessment based on the user's estimation of risk specific magnitude of impacts and three probabilities contributing to risk: a) probability of hazard occurring, b) probability of receptor being exposed and c) probability of harm occurring. It also supports Options Appraisal of different risk management options and enables the selection of the best option using Multi-Criteria

Analysis (MCA) and according to the stakeholders' preferences. Although the current version of STRiM DSS does not support Quantitative Risk Assessment, this feature can be also implemented. The Risk Assessment and Management (RA&M) DSS serves as an integration platform (Figure. 2) for the other components of STRiM DSS: DSS map server (IMS), DSS indicator database and DSS document database.

The RA&M DSS Client Application offers an intuitive, user-friendly interface to facilitate an easy utilization of the application functionality (Figure 3). DSS map server (MS) Client Application is also implemented as a Web application and offers a capability of handling, combining and viewing GIS files in order to create maps needed in STRiM DSS. The business tier contains:

- Working Session, which performs retrieval and storage operations for the Initiation, Problem Formulation, Generic Risk Assessment and Management Process, Risk Management and Risk Communication steps of the STRiM Risk Assessment and Management Framework.
- DSS MS Server Components used to communicate with the underlying UMN MapServer to handle GIS files and fetch data from STRiM GIS.

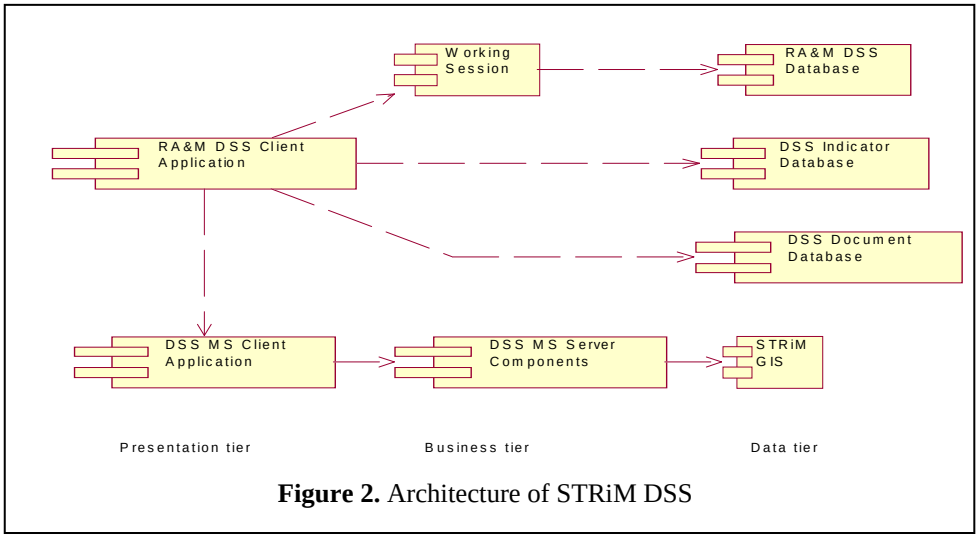


Figure 2. Architecture of STRiM DSS

RA&M DSS Database is a relational database used to store information about the objects in RA&RM DSS Application: users, case studies, baselines, hazards, risk generating processes, boundaries, controlling factors, magnitude of impacts, probabilities, risk prioritization, management options, options appraisal, risk management, risk monitoring. DSS Indicator Database contains a set environmental indicators for risk assessment for different environmental domains collected from a number of existing and commonly used environmental databases. STRiM GIS acts as a spatial database which contains the thematic maps needed for risk assessment. These thematic maps are attached to case studies developed using STRiM DSS. In order to visualize spatial data, STRiM GIS employs MapServer from the University of Minnesota. MapServer is one of the leading packages for web mapping applications, providing feature-rich cartographic output, such as scale bars, legends, reference maps, and labelling. MapServer has the ability to generate thematic maps based on classes and regular expressions. Moreover MapServer implements the open geospatial consortium implementation specifications for open web services (OWS), namely Web Mapping Service (WMS), Web Feature Services (WFS) and Web Coverage Service (WCS). It also supports GML and provides on-the-fly projection capability.

2.3 Geospatial data services and data sharing

The effortless sharing of geospatial data is a major requirement of STRiM GIS. In order to support collaborative decision making among its stakeholders, STRiM GIS data services were developed to overcome most technical impediments to the accessibility of shared data. In Internet Mapping applications, spatial data is expected to be massive, collected by a variety of sources and represented in a multitude of formats. In order to allow effective sharing, STRiM GIS provides users with a number of data services. STRiM GIS data storage is either file-based, or utilizes spatial relational databases. For file-based data, there

are the following basic services: the file management service, the data translation service and the data composition service.

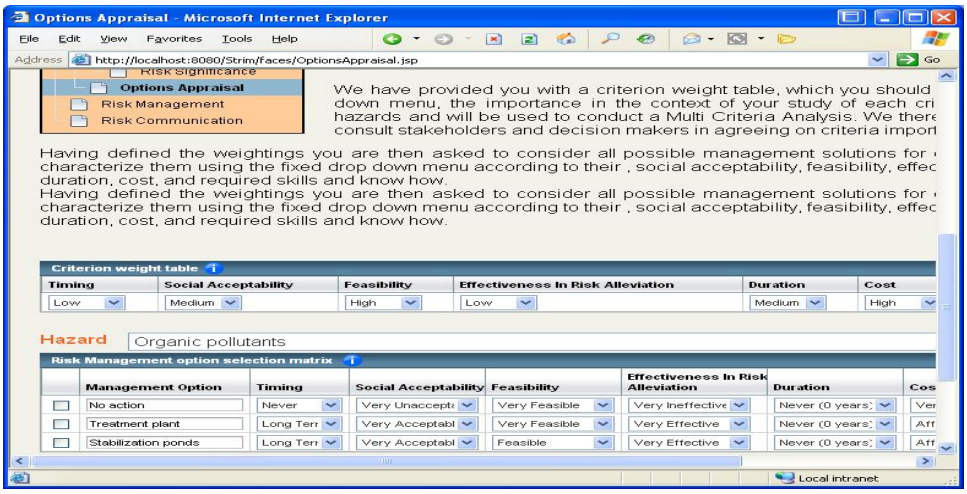


Figure 3. STRiM Risk Assessment and Management DSS user interface

The file management service allows STRiM GIS users to upload, download, delete and inspect spatial data files in a variety of formats. File transfer to and from the server uses standard web mechanisms. Inspection applies to raster images and presents the user with file related metadata, without accessing the bulk of the file. The file translation service is essentially a front-end to the Geospatial Data Abstraction Library (GDAL) for raster data translation and the OGR Simple Features Library for vector data translation (Warmerdam).and support translation between over 40, as of this publication, different geospatial data formats. Typically a user would translate a given raster file into an HDF4 [HDF Group] dataset that is then used for raster analysis and map algebra via the NAP module described below. The base options of the translation facility are quite strict regarding format conversions, and are used as default in this implementation, otherwise errors would occur upon translation. Although not currently implemented, a user should be able to “relax” these default, options by using, among others, the “-not_strict” flag which is more ”forgiving of mismatches and lost data when translating to the output format” (GDAL_translate). A number of spatial database servers can also be supported for storing spatial data. Currently, only simple feature vector data can be stored in spatial databases. The STRiM GIS implementation is based on the PostGIS RDBMS, but using different RDBMS products is straightforward. The data composition service is used to combine together spatial data files or database-stored records. Data from the selected files or database tables is used to define a number of layers, as well as queriable data attributes, via a form-based interface. The final product of this process is a new map, which can then be loaded into the visualization service, to be explored and/or processed by the spatial analysis tools.

Spatial processing in STRiM GIS is currently heavily geared towards processing raster data via map algebra service. providing a plethora of unary and binary arithmetic, logical and fuzzy-logic related operators, as well as several map classification methods. All map algebra functionality on raster data is currently implemented by the Numerical Array Processor library (NAP), a language similar to APL, J, IDL and Matlab (Davis, 2002). In a setting where map algebra operations are computed on the same machine as the web server, it could be impossible to support more than a few concurrent users without severe drop in system responsiveness (Hawick, 2003). To alleviate this, our data analysis services are architected on a client/server model: processing is performed on a number of dedicated compute servers, (possibly) separate from the web server machine. Dispatch is via Simple Object Access Protocol (SOAP), which is also similar to widely adopted standards for Grid computing (Globus and the Open Grid Software Architecture standard). All the aforementioned modules are integrated into and run as threads under the TclHttpd Web Server, a robust, multi-threaded application server, (Welch, 2000).

3. STRiM DSS IMPLEMENTATION & EVALUATION

In order to test the functionality and applicability of the STRiM RA & M framework as well as DSS a number of pilot trials have been carried out (Table 1). An international training seminar to over 60 practitioners is organised for June 2008 where a user focused evaluation of the DSS will be carried out. In this paper preliminary results from the pilot trials which involved end user and stakeholder consultations during the process are presented.

Table 1. STRiM Pilot Trials

No	Pilot Trial Name	Scale	Transboundary	Participatory
1	Keritis watershed risk assessment and management of water pollution (Crete, Greece)	Localised (watershed)	No- Chania Prefecture	Yes (water and regional authorities, geologists)
2	Axios Delta: Risk assessment and management of wetland loss resulting from water management processes	Large scale Axios river basin	Yes- Greece & FYROM	Yes (consultation with water authorities)
3	Assessment and management of flooding risks in the Sava river basin	Large scale Transboundary flood plain	Yes, Bosnia and Herzegovina, Serbia and Montenegro	Yes (consultation with planning authorities)
4	Generic Risk Assessment and Management for storm break in the national parks of Sumava	Large scale	Yes, Bavaria and Czeck Republik	Yes (Administration of National Parks, Forest Owners)

The tangible outputs of the DSS application consist of automated reports which describe the results of each step of the process including a list of risks requiring further detailed study and a prioritised list of risk management options. Within the report a description of risk communication activities is also disclosed. The preliminary evaluation of the framework and DSS against the original specifications (see Section 2) is positive indicating that the STRiM RA & M DSS is in fact a one stop shop to transboundary RA & M in the CADSES region. The pilots illustrated that for the generic assessment a user, without undergoing recommended consultation exercises can effectively carry out and have the automated report within a few hours! In addition, any assessment can be revisited and different risk management options evaluated, illustrating the dynamic nature of this tool. The flexible qualitative nature of the framework which allows the user to select data sets when estimating risk significance overcomes a commonly acknowledged issue of transboundary assessments in general, that of data incompatibility and unavailability. The applicability of this application at different scales was also concluded as feasible, as the DSS structured yet flexible approach to problem formulation enables the assessments to be undertaken at any scale, relevant to environmental planning decision making such as SEA, EIA or River Basin Management Plans (WFD). However, it was noted that at larger scales, the number of hazards (primary and secondary) requiring further detailed quantitative risk assessment was greater than for the targeted smaller scale pilot trails, underlining the value of having designed the framework to be an iterative process (Figure 1). The capacity of the DSS to assign weightings to management criteria prior to undertaking an options appraisal, illustrates the frameworks recognition of the context specific nature of decision making, yet on the other hand the promotion of multi-stakeholder involvement during the assessment, which can take place remotely (something often an issue in transboundary assessment) the value of this application as a participation media is underlined. Transparency, through the automated report structure enables the clear illustration of the uncertainty and subjectivity involved in all such type of generic procedures.

4. CONCLUSIONS

Despite the preliminary implementation results of the STRiM RA & M DSS indicating the fulfilment of original specifications, there are a number of barriers and inherent limitations to the process. Fundamental is the lack of formalised legislative and institutional adoption of this process by CADSES region governments, which may limit its utilisation in the long term. Similar to any generic procedure, the advantage of its wide applicability to any environmental risk at any given planning relevant scale, limits the detail and increases the

level of uncertainty of the results obtained. Nevertheless, the capacities of this application are formidable as they allow any user (which does not have to have a scientific background) to proceed step by step in formulating the problem he or she wishes to address, conduct a preliminary risk assessment and identify areas which require further investigation as well as most appropriate (to the specific context) management options for further consideration. All this can be carried out jointly, yet remotely, in collaboration with stakeholders and decision makers in bordering countries, or regions. Finally, the capacity of the DSS to automatically produce reports which illustrate the assumptions and decisions made throughout the process apart from reducing the bureaucracy and tediousness of the procedure also increases the transparency required for any effective risk communication process.

ACKNOWLEDGMENTS

This research was funded under the INTERREG IIIB CADSES project STRiM (Priority 4, measure 4.2; code: 5D008). The authors would like to acknowledge the contributions of the entire STRiM consortium.(http://www.strim.eu).

REFERENCES

Craik, N., *Transboundary Environmental Impact Assessment in North America: Obstacles and Opportunities, Annual Meeting Washington* D.C., 3-7 January http://www.aals.org/am2007/thursday/craik.pdf , 2007

DEFRA, Department for Environment, Food and Rural Affaires, *Guidelines for Environmental Risk Assessment and Management*, HMSO, 2002.

Gdal_translate http://www.gdal.org/gdal_translate.html

Giupponi, C., Decision Support Systems for implementing the European Water Framework Directive: The MULINO approach, *Environmental Modelling and Software*, 22, 248-258, 2007.

HDF Group http://hdf.ncsa.uiuc.edu/

Karydas, C.G., G.C.Zalidis, L.Kouskouna, E.Feoli, N.G.Silleos, I.Z.Gitas. 'European Environmental Legal Framework and Decision Support towards Sustainable Development: ISOTEIA model', *OTEM Summer School Proceedings*, Baia Mare, Romania, July 2006.

Kriebel, D., Ticker, J., Epstein, P., Lemons, J., Levins, J., Loechler, E.L., Quinn, M., Rudel, R., Schetter, T., and Stoto, M., The precautionary Principle in Environmental Science, *Environmental Health Perspectives*, 109 (9), 871-875, 2001

Pediaditi, K., Wehrmeyer, W., Chenoweth, C., Brownfield redevelopment, integrating sustainability and risk management., in Brebbia, C.A., Popov., V., & Fayzieva, D., (eds), *Environmental Health Risk III*, WIT Press, 2005

Peltonen, L., Recommendation for a risk mitigation oriented european spatial policy. Natural and Technological hazards and risks affecting the spatial development of European Regions, Geological Survey of Finland, Special paper 42, 153-176, 2006.

Power, M., McCarty, L.S., A comparative analysis of environmental risk assessment/risk management frameworks, *Environmental Science and Technology*, 32: 224A-231A, 1998.

Quevauviller, P., Balabanis, P., Fragakis, C., Weydert, M., Oliver, M., Kaschl, A., Arnold, G., Kroll, A., Galbiati, L., Zaldivar, J.M., and Bidoglio, G., Science- policy integration needs in support of the implementation of the EU Water framework Directive, *Environmental Science and Policy*, 8, 203-211, 2005

MapServer, http://mapserver.gis.umn.edu

Willems, P., and de Lange, W.J., Concept of technical support to science-policy interfacing with respect to the implementation of the European water framework directive, *Environmental Science and policy*, 10, 464-473, 2007.

Harvey Davis 2002, the NAP (N-Dimensional Array Processor) extension to Tcl. ActiveState

Coddington, P. D.. Hawick, K. A and James, H. A. Web-based access to distributed high-performance geographic information systems for decision support. In *Proc. Hawaii Int'l Conf. on System Sciences (HICSS'99)*, 1999.

Warmerdam, F. http://www.gdal.org

Welch, B. The Tcl Web Server, In Proc. 7th *USENIX Tcl/Tk Conference* February 14-18, Austin, Texas, USA, 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

Estimating the Pollution Risks to Urban Groundwater from Industry and Sewers

D.N. Lerner[a], B.N. Chisala[a,b], and N.G. Tait[c]

[a] Catchment Science Centre, University of Sheffield, Broad Lane, Sheffield S3 7HQ, UK (d.n.lerner@shef.ac.uk)
[b] Now with Croydon Council, Davis House, Robert Street, Croydon CR0 1QQ (Brenda.Chisala@croydon.gov.uk)
[c] Property and Land Application Team, Royal Borough of Kensington and Chelsea, Horton Street, London W8 7NX, UK (Nigel.Tait@rbkc.gov.uk)

Abstract: Urban groundwater continues to be at risk of pollution by organic chemicals and microbiological contaminants despite its potential economic and ecological value. Assessing these risks is hampered by the large number of potential sources and the general lack of detailed site data, and so there is a likelihood of poor decisions being made which would spoil the overall quality of life. However, we have shown that it is possible to build a GIS-based risk analysis tool. The tool, called the Borehole Optimisation System or BOS, has been validated against a variety of field datasets, and has been shown to make reasonable predictions of risks – that is within 2 orders of magnitude. Better predictions are made when there is a multiplicity of potential sources within a borehole catchment. There are a range of types of risk predictions that can be made with BOS, including analysis of a single site, mapping of risk over a city, and generic risk analysis without a site specific component. These risk analyses are probabilistic to take account of the uncertainties and poor characterisation of the environment. The information can be presented in a simple enough way to support decision-making and help to enhance the quality of life by targeting resources in an efficient manner.

Keywords: GIS; BOS; urban groundwater; groundwater pollution; probabilistic risk modelling.

1. URBAN GROUNDWATER POLLUTION

Urban groundwater is frequently considered to be at significant risk of pollution from the urban activities on the land surface above. The major sources of pollution are (a) industry, through spillages of chemicals and the creation of contaminated land, and (b) leaking sewers.

Industrially contaminated land is widespread, although in principle there are regulations preventing it in most countries. There is often legislation requiring assessment of the risks from any pollution and there may be requirements for clean-up of contaminated sites. However, the focus of the regulatory regime is on individual polluting sites, rather than on the cumulative effect of multiple sites on individual receptors.

Much less legislation focuses on the effects of leaking sewers, despite the health impacts of microbiological pollution being much larger than those of industrial chemicals. There is evidence from the UK and elsewhere that sewers do leak and cause pollution of groundwater, and there have been serious consequences for health in some incidents. The most infamous recent case is Walkerton in Ontario, with 2300 people falling ill and seven deaths [Howard, 2007]. Lerner et al. [1994] and others have catalogued cases in the UK and beyond where significant problems have arisen from sewer leakage. A recent exercise found widespread microbiological pollution in shallow groundwater in Nottingham [Barrett

et al., 1999], and evidence of enteric bacteria and viruses tens of metres below the water table in both Nottingham and Birmingham [Cronin *et al.*, 2003; Powell *et al.*, 2003]. However, there has been no systematic assessment of the risks that such leaks pose to groundwater.

The quality of life can be improved by reducing the risk of pollution affecting human health or ecological quality. However, if money is spent on minor problems of groundwater pollution, it will be diverted from potentially more important investments and it could be argued that the quality of life will be made worse by poor decisions. The urban infrastructure is complex and poorly characterised, and urban hydrogeology is usually poorly characterised as well, due to the costs of investigations. In such circumstances, poor decisions are quite likely to occur due to the lack of good information.

Over a series of projects, we have developed a set of linked tools to estimate the risk of pollution for a potential new user of urban groundwater [Davison et al. 2002; Tait et al. 2004; Chisala et al. 2007; Tait et al. 2008]. The overall package is called BOS (Borehole Optimisation System). The objectives of this paper are to give an overview of BOS and to discuss the validation of risk models. We give some example applications in order to show that it is possible to use information handling to support good decision making in such poorly characterised systems and so enhance the quality of life.

2. BOS: BOREHOLE OPTIMISATION SYSTEM

2.1 Overview of BOS

BOS forecasts water quality at new abstraction borehole locations. The conceptual model is the reverse of the conventional source, pathway and target risk assessment methodology. BOS begins with a user specified borehole (the target), and retraces the flow lines of the capture zone in order to identify the multiple potential contaminant sources situated upstream.

In order to undertake the risk analysis, BOS utilises three discrete component modules. The Catchment Zone Probability Model (CZPM) module identifies the probabilistic surface expression of a borehole catchment. The Land-use Model (LM) module identifies the potential current and historical contaminant sources within the catchment area from either industrial landuses or the sewer network. The Pollution Risk Model (PRM) module is essentially a transport model, and estimates the combined threat posed by the identified potential contaminant sources at the abstraction borehole. The integration of these independent component modules (CZPM, LM and PRM) within a single Graphical User Interface (GUI) was central to the successful development of BOS as a powerful tool for addressing the issues regarding the best use for urban groundwater under conditions of high uncertainty. In this case GIS offers the ideal platform for coupling the diverse components that make up BOS.

BOS seamlessly integrates the CZPM, LM and PRM modules within a GIS based GUI which contains all of the necessary functions with which to undertake the groundwater risk analysis while shielding the user from the complexities of the background process controlling the component modules. The integrated working environment offers either full control to the experienced user or minimal interaction to the novice and creates a powerful decision making tool for establishing the best locations for new boreholes in urban areas.

BOS depends on the user to supply the specific modular datasets for the urban region to be analysed. Thus a MODFLOW groundwater flow model, historical land-use and sewer network shapefiles, land-use and contaminant property databases and surface elevation coverages are essential. Background shapefiles (roads, rivers, railways etc) of the study area are not necessary but can be utilised by BOS, for visualisation purposes, if present. The reliance of BOS on the user to provide these representative datasets means that the application is not restricted to the analysis of a single urban area. Indeed the BOS application can be applied to any region given the appropriate data in the correct format.

2.2 Components of BOS

The CZPM module is based on a three-dimensional finite-difference MODFLOW groundwater flow model. It is described by Davison et al. [2002] and Tait et al. [2004].

The LM module identifies pollution sources within the catchment using spatial land-use information and associated Microsoft Access land-use and contaminant databases. It comes in two versions. The original version handles landuse, and identifies sources of industrial pollution, principally organic contaminants [Tait et al. 2004, 2008]. The newer version has added information on sewer networks in order to estimate loads of microbiological pollutants. Using the limited data that is available on sewer leakage rates, and how these rates are related to sewer age, a model has been developed to estimate leakage rates for each 500x500 m grid square in the city [Chisala and Lerner, 2008a].

The Pollution Risk Model (PRM) module employs a stochastic analytical solute transport model based in a Microsoft Excel spreadsheet with Crystal Ball probabilistic extension [Decisioneering, 1996] and custom PRM add-ins. For organic pollutants, each source term is conceptualised as the dissolution of the contaminant from a non-aqueous phase liquid (NAPL) by recharge, so is represented by the contaminant solubility (C_s), mole fraction (X), recharge rate (R), and the area of the source (A), with the latter providing a way of describing and varying the source strength [Chisala et al. 2007].

Calculating potential attenuation requires the total travel time for contaminants through the unsaturated and saturated zones. The unretarded travel time through the unsaturated zone to the water table is a function of soil moisture content (θ), unsaturated zone thickness (w) and recharge rate. The unretarded travel time (t_s) in the saturated zone is automatically calculated by the groundwater flow model (see above). Biodegradation, treated as a pseudo first order decay process, and sorption are described by the biodegradation rate constant (λ) and retardation factor (R_f) respectively. The resulting pollutant flux at the borehole is summed over all m sources, and is diluted in all the water pumped from the borehole by use of the pumping rate (Q). The predicted concentration at a borehole, C_w, is given by:

$$C_w = \sum_m C_S X \frac{RA}{Q} \mathbf{exp}\left[-\lambda R_f \left(\frac{\theta w}{R} + t_s \right) \right] \quad (1)$$

For microbiological pollutants, this conceptual model of fate and transport was found to be inadequate as it underestimated the concentrations observed in the field [Chisala and Lerner, 2008b]. A revised model which allowed preferential flow of a small fraction of the pollutant load was employed and gave much better results. :

$$C_w = \sum_m F_u \left[0.01 \mathbf{exp}\left(\frac{\lambda R_f t_s}{f} \right) + 0.99 \mathbf{exp}\left(\lambda R_f t_s \right) \right] \quad (2)$$

where F_u is the flux of pollutants reaching the water table, 0.01 is the fraction of the load travelling by preferential pathways, and f is a travel time factor expressing how much faster flow is in the preferential pathway than in the bulk, matrix flow, pathway.

3. APPLICATIONS

3.1 Case study area

The City of Nottingham was selected as a study area where BOS could be applied due to its geological character and the numerous studies on groundwater under the city that have been conducted. This urban area is situated in the East Midlands on one of the most important aquifers in the United Kingdom. The Permo-Triassic Sandstone aquifer provides groundwater resources that have been extensively developed for both public and industrial water supply across the country. In areas where they are not affected by human activity the

aquifer produces a high yield of generally good quality groundwater. However previous studies have identified both localised industrial contamination of groundwater in numerous urban areas by organic compounds and more diffuse inorganic contamination.

This city has been a major industrial centre since the 18th century and has a long and varied industrial history largely originating in the textile industry. By the middle of the 19th century Nottingham's rivers had become too polluted to be used for public water supply. This was a time of growing industry and population contemporaneous with an increased demand for clean water. Thus from 1850 onwards virtually the whole water supply for the region was derived from the Sherwood Sandstones below the city. The impact of the long-term industrialisation in the Nottingham area on the underlying groundwater is indicated in studies by Barrett et al. [2001]. This work revealed the groundwater quality beneath this area to be poorer than nearby rural locations. The deterioration is not great for inorganic species, except for localised pollution incidents, and no trace metals were found. Nitrate concentrations are similar in urban and rural locations, and frequently exceed the drinking water limits. Chlorinated solvent pollution is widespread as are BTEX (benzene, toluene, ethylbenzene and xylenes) compounds originating from fuel and solvent spills. Most water is now withdrawn in the surrounding rural areas. Most of the groundwater abstraction in Nottingham is now used for private industrial use.

The regional hydrogeology and groundwater flow model for the Nottingham area used in this case study has been described by Yang et al. [1999] and Trowsdale and Lerner [2003]. The model is a steady state single layered model covering the urban area and rural locations to the north and east. The landuse component of the LM comprises a series of spatial land-use databases depicting the historical land-use activity in the urban area and a relational database containing the associated industrial and contaminant information. The data inputs for the LM are described in more detail by Davison et al. [2002], and resulted in 6 land-use maps for the years 1901, 1920, 1939, 1954, 1971, and 1991. They contain information on 16 000 landuse polygons for a city of ~250 000 people. The sewer network is represented by the digitised sewer asset map kindly provided by Severn Trent Water, with ages of each area taken from the dates of the housing developments.

3.2 Validation

Most environmental risk models are not validated, which should give users some cause for concern. Of course, it is no easy to validate a risk model because its purpose is to forecast the future, which implies that there are not yet any observations available to check the model forecasts against, and to do so in a probabilistic way. Validating a probabilistic forecast would require enough field observations to calculate statistics and frequencies of exceedances. Nevertheless, given the importance of giving confidence to user by validating a model, we have attempted to validate BOS for a number of different pollutants by a range of approaches, as summarised below:

- Single site, perchloroethene (PCE) a chlorinated hydrocarbon solvent widely used for dry cleaning. At one borehole in the case study area, average concentrations of PCE were available. BOS was run to predict concentrations at this location, and 91 potential sources of PCE were found in the catchment. The 50%ile concentration predicted (49 µg/l) was within a factor of 2 of the observed concentration (33 µg/l) (Tait et al. 2004).

- Multiple sites, PCE. In a later study, information on PCE concentrations was available for 6 pumped boreholes. Differences between predicted and observed concentrations ranged from <1 to 6 order of magnitude (OM), with 4 cases being within 2 OM. The more accurate predictions were for boreholes with higher numbers of potential PCE sources upstream (Tait et al. 2008).

- National statistics for methyl tert butyl ether (MTBE) an additive to petrol to enhance octane ratings and reduce smog-forming emissions. In an attempt to compare BOS against a large dataset, field data from 1100 boreholes in England and Wales was compared with model predictions for 70 sites within the case

study. There is not an exact match between the hydrogeology and pollution scenarios of the two sets, and we expect the field data to show lower concentrations as it includes more rural sites. 96% of the field samples are non-detects, and there were a lot of different detection limits used, which made comparison difficult with the more consistently organised model predictions. Nevertheless, the percentage of exceedances of concentrations from 0.5 to 100 µg/l between the two sets was within a factor of 2 (Chisala et al. 2007).

- Microbiological indicators, 3 sites. Low levels of microbiological pollution were observed at the three locations, two with multi-level samplers which had been sampled on several occasions, while the third had one set of observations from 11 boreholes in a small area. The frequency of field observation of faecal bacteria was compared with model predictions. Provided a preferential flow mechanism was included in the transport model (Eqn 2), there was good agreement for the two multi-level sites, but not for the third location (Chisala and Lerner, 2008b).

These validation exercises gives more understanding of the BOS model and its uncertainties than is usually available for risk models. They show that the model has some predictive power, but that it can only be used as a screening tool. It performs better when there are multiple sources in a catchment, presumably because of the averaging effects.

3.3 Use of BOS to predict risks

The city-wide risk model has four main uses. It can be used to analyse the risks of pollution for a single location, perhaps the site of a proposed supply borehole. In this case, a realistic groundwater flow model can be used, and special attention paid to collecting data about landuses in the probable catchment of the new source. This type of use has been shown in the PCE and MTBE case studies (Tait et al. 2004, Chisala et al. 2007).

A variant on single site analysis is to help interpret the conditions at an existing borehole were pollution is occurring. BOS is able to give a rapid analysis on the potential sources upstream of any site and the probability that they are within the boreholes catchment. This is essentially the type of analysis used in the validation exercises outlined above.

The third use of BOS is to map the risk of pollution across the whole city in order to identify areas of high and low risk. In this case, a standard new borehole is simulated on a grid of positions across the whole city, and the resulting predictions contoured. For the Nottingham case, 1300 locations on a 50 m grid were simulated for the PCE risk map (Tait et al. 2008) and 70 locations for the MTBE map (Chisala et al. 2007).

The final use we have made of BOS is a generic risk analysis, in which broad conclusions on the likelihood and severity of pollution are drawn. This can be done by analysing the statistics of the multiple simulations carried out in a risk-mapping exercise, but can also be done with fewer simulations. For the analysis on the risks of microbiological pollution from sewers, the model was run for a number of randomly selected locations. These showed that there was a significant risk of microbiological pollution in most locations (Chisala and Lerner, 2008b).

4. CONCLUSIONS

Urban groundwater continues to be at risk of pollution by organic chemicals and microbiological contaminants despite its potential economic and ecological value. Assessing these risks is hampered by the large number of potential sources and the general lack of detailed site data. However, we have shown that it is possible to build a GIS-based risk analysis tool. The tool, BOS, has been validated against a variety of field datasets, and has been shown to make reasonable predictions of risks – that is within 2 orders of magnitude. Better predictions are made when there area multiplicity of potential source within a borehole catchment. There are a range of types of risk predictions that can be

made with BOS, including analysis of a single site, mapping of risk over a city, and generic risk analysis without a site specific component.

The case study examples show that a complex and poorly characterised problem such as the risk of urban groundwater pollution can be represented in models. Relatively simple summary outputs can be produced to inform decision makers. These outputs can assist them to make good choices for environmental protection and improvement, and so help improve overall quality of life.

ACKNOWLEDGEMENTS

This research has been financially supported by the Engineering and Physical Sciences Research Council, the Natural Environment Research Council, the Environment Agency, the Ministry of Agriculture Fisheries and Food (now part of Defra) and a Commonwealth Scholarship. The opinions are those of the authors alone.

REFERENCES

Barrett, M.H., Hiscock, K.M., Pedley, S., Lerner, D.N., Tellam, J.H. and French, M.J. Marker species for identifying urban groundwater recharge sources: a review and case study in Nottingham, UK. *Water Research*, 33(14), 3083–3097, 1999.

Barrett, M.H., Yang, Y., Lerner, D.N. French, M.J. & Tellam, J.H. *The impact of cities on the quantity and quality of their underlying groundwater*. Technical report TR P185 by University of Bradford for the Environment Agency, Bristol, 2001.

Chisala, B.N., Tait, N.G. and Lerner, D.N. Evaluating the risk of methyl tertiary- butyl ether (MTBE) to urban groundwater at city scale: Nottingham case study. *Journal of Contaminant Hydrology*, 91, 128-145. doi: 10.1016/j.jconhyd.2006.08.013, 2007.

Chisala, B.N. and Lerner, D.N. Quantity and distribution of sewer exfiltration to urban groundwater. *Water Management, In press*, 2008a.

Chisala, B.N. and Lerner, D.N. *Sewage risks to urban groundwater.* Science Report SC030134, Environment Agency, Bristol. 36 pages. ISBN 978-1-84432-820-8. 2008b.

Cronin, A.A., Taylor, R.G., Powell, K.L., Barrett, M.H., Trowsdale, S.A. and Lerner, D.N. Temporal trends in the hydrochemistry and sewage-related microbiology of urban sandstone aquifer – Nottingham, UK. *Hydrogeology Journal*, 11(2), 205–216, 2003.

Davison, R.M., Whittaker, J.J., Prabnarong, P. and Lerner, D.N. A probabilistic management system to optimise the use of urban groundwater. In: K.M. Hiscock, M.O. Rivett & R.M. Davison (eds), *Sustainable groundwater development*, Special Publication 193, 265-276. Geological Society, London, 2002.

Howard, K.W.F. Microbial pollution of groundwater in the town of Walkerton, Canada. In: J.H. Tellam et al. (eds), *Urban groundwater management and sustainability.* Nato Science Series V74, Springer, Netherlands. 315-330, 2006.

Lerner, D.N., Halliday, D. & Hoffman, J.M. The impact of sewers on groundwater quality. W.B Wilkinson, *Groundwater Problems in Urban Areas*, Telford, London, 64-75. 1994.

Powell, K.L., Taylor, R.G., Cronin, A.A., Barrett, M.H., Pedley, S., Sellwood, J., Trowsdale, S.A., Lerner, D.N. Microbial contamination of two urban sandstone aquifers in the UK. *Water Research*, 37(2), 339–352, 2003.

Tait, N.G., Davison, R.M., Whittaker, J.J., Leharne, S.A. and Lerner, D.N. Borehole Optimisation System (BOS) - A GIS based risk analysis tool for optimising the use of urban groundwater. *Environmental Modelling and Software*, 19, 1111-1124, 2004.

Tait, N.G., Davison, R.M., Leharne, S.A. and Lerner, D.N. Borehole Optimisation System (BOS) - A case study assessing options for abstraction of urban groundwater in Nottingham, UK. *Environmental Modelling and Software*, 23, 611-621, 2008.

Decisioneering. Crystal Ball version 4.0 User Manual: Discover the Risks and Opportunities in your Forecasts. Bright Communication Inc., Colorado, 1996.

Trowsdale, S.A., Lerner, D.N. Implications of flow patterns in the sandstone aquifer beneath the mature conurbation of Nottingham (UK) for source protection. *Quarterly Journal of Engineering Geology and Hydrogeology* 36 (3), 197–206, 2003.

Yang, Y., Lerner, D.N., Barrett, H.M., Tellam, J.H. Quantification of groundwater recharge in the city of Nottingham, UK. *Environmental Geology*, 38(3), 183–198, 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

Spatial Information System for Risk Assessment of Dust Transport in the Neighbourhood of the Surface Coal Mine

L. Matejicek[a], **Z. Janour**[b]

[a] *Institute for Environmental Studies, Faculty of Science, Charles University in Prague, Benatska 2, Prague 2, 12801, Czech Republic (lmatejic@natur.cuni.cz)*

[b] *Institute of Thermomechanics, Academy of Sciences, Dolejskova 5, Prague 8, 18200, Czech Republic*

Abstract: Risk assessment of dust transport from surface coal mines over neighbouring residential zones requires more complex modelling tools in order to analyse satellite images, to attach high resolution aerial images, to create more precise digital terrain models, to integrate time series of meteorological variables, to realize linkage between the spatial database and numerical simulation tools, to provide spatial analysis, and to visualize the results. Thus, a spatial information system based on GIS technology is used for data management, advanced spatial analysis and numerical simulation linked by shared data files. At first, a digital terrain model is created by processing data originating from surface laser scanning, GPS measurements and geodetic surveys. This enables precise modelling of wind flows over potential dust emission sources for an estimation of dust transport. Finally, the simulation outputs are transformed into a set of existing digital map layers, and spatial analysis is carried out for visualization.

Keywords: Spatial Information System; Dust Transport; Risk Assessment; Surface Mine.

1. INTRODUCTION

For the past few decades, various approaches have been developed to perform health risk assessment in compliance with requirements contained in the US EPA methodology as well as with EU instructions [Mower, 1998]. Due to available data, recent projects are mostly focused on short-term direct and indirect effects, such as ingestion through food chains or inhalation, but do not include research for long-term observations and data processing [Finkelman, 2002]. Moreover, evaluations become very complex when several emission sources contribute to the overall concentration of contaminants in the air. Because each emission source has a different emission rate, location and dispersion characteristic, more complex analytical tools are needed in order to manage data, provide analyses, solve numerical models, display results and share data via networks.

The objectives of this paper are to propose and demonstrate a spatial information system that can support risk assessments focused on dust dispersion over a selected surface coal mine and surrounding neighbourhoods in the Czech Republic. In mining areas, suitable remediation procedures have to be carried out to protect the natural environment, because associated impacts on human health have long been recognised. In order to find the best scenarios for remediation, all suitable data have to be collected and analysed.

A prerequisite for exploring the impact of dust deposition is to determine emission sources, spatial distribution and temporal variation [Lu, 2001]. To carry out all these tasks, a spatial information system is needed for data management, spatio-temporal analysis, numerical simulation and final visualization. Despite the existence of software tools for spatial

analysis or temporal analysis, a complex quantitative solution has not been possible until recently. Thus, in order to integrate spatial data and temporal data, a geographic information system (GIS) is used for these tasks [Goodchild, 1996].

2. SPATIAL DATABASE

Nearly all recent GISs can manage data in a relational database, which simplifies data sharing among various software tools and via the Internet. Display and visualization are provided in the framework of digital map layers that include vectors, grids, triangulated irregular networks (TINs) and other objects [Zeiler, 1999; Döllner, 2000]. The spatial data formats offer more complex analyses, ranging from exploratory spatial data analysis (ESDA) to geostatistical interpolation and modelling [Mitchell, 2005; Maguire, 2005].

Data represented by the digital thematic map layers of the surface coal mine and its surrounding areas are stored in the database and complemented by satellite images, aerial images, surface laser scanning, GPS measurements, a geodetic survey, data from local meteorological stations, terrain measurements of dust deposition, and inputs/outputs of standalone modelling tools. The database is created in the framework of the GIS project and is based on ESRI's geodatabase [Arctur, 2004].

3. SPATIAL INFORMATION SYSTEM

In order to support the risk assessment of dust transport, the spatial information system based on ESRI's ArcGIS is created to analyse the surface spatial data, local meteorological observations and the inputs/outputs of numerical modelling. In addition to GIS software, many other software tools have been used for data pre-processing.

3.1 Satellite images

An overall view of the surface mines from a Landsat 7 panchromatic image is shown in Figure 1. The area of interest marked in the middle encompasses approximately 1.25 km^2.

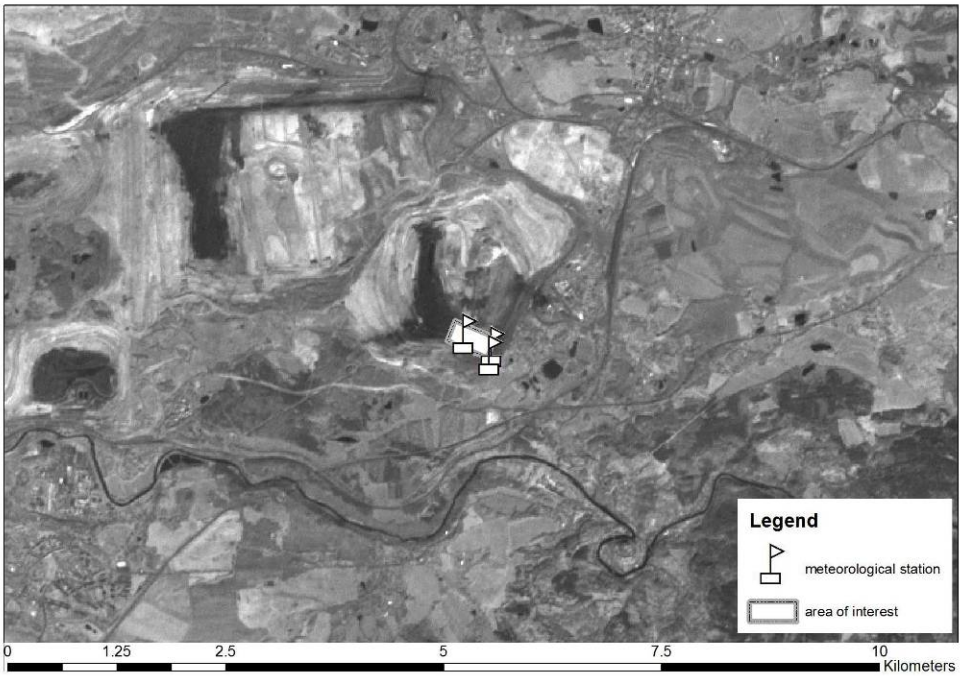


Figure 1. The satellite image from Landsat 7 with the area of interest and the local meteorological stations (panchromatic band 0.520-0.900 μm at 15 meters resolution).

In addition to the panchromatic band at 15 meters resolution, the Landsat 7 Enhanced Thematic Mapper (ETM+) provides 7 bands of multispectral data. While bands 3-2-1 (red, green, blue) create a true colour composite which displays objects similarly to a colour photograph, bands 4-3-2 (near infrared, red, green) and bands 5-4-2 (mid infrared, near infrared, green) are used for classification and identification of geologic rock types and soil boundaries.

3.2 Surface laser scanning

Surface laser scanning is used to capture surface points in an area of temporary coal storage and neighbouring slopes near residential zones, Figure 2. To obtain complete measurements from shielded parts, the final set of spatial points is based on multiple scans from different locations. After spatial transformation, the reduced set is placed into a common reference system and included into the complex DTM.

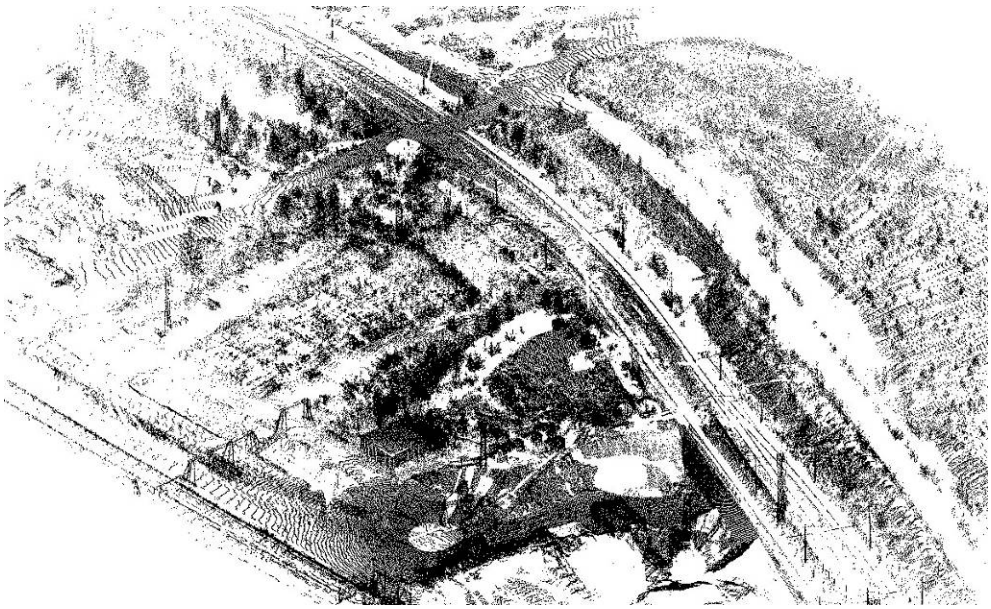


Figure 2. The clouds of points from surface laser scanning for creation of the DTM.



Figure 3. The draped aerial image and the GPS measurements on the DTM.

3.3 GPS measurements and draping the aerial images on the DTM

The geo-rectification of satellite images, aerial images and clouds of surface spatial points is provided by GPS and data from the geodetic survey. GPS is also used for the location of emission sources, slope shapes, roads and temporary storage sites. The draped aerial image on the DTM is shown in Figure 3. This provides a more realistic view of the area of interest and its neighbouring residential zones.

3.4 Time series from the meteorological stations

Time series of meteorological variables captured by local meteorological stations are pre-processed in order to explore the dominant wind direction, wind speed, atmospheric pressure, and to set model parameters. The wind class frequency distribution and the wind rose based on data from all the meteorological stations are shown in Figure 4.

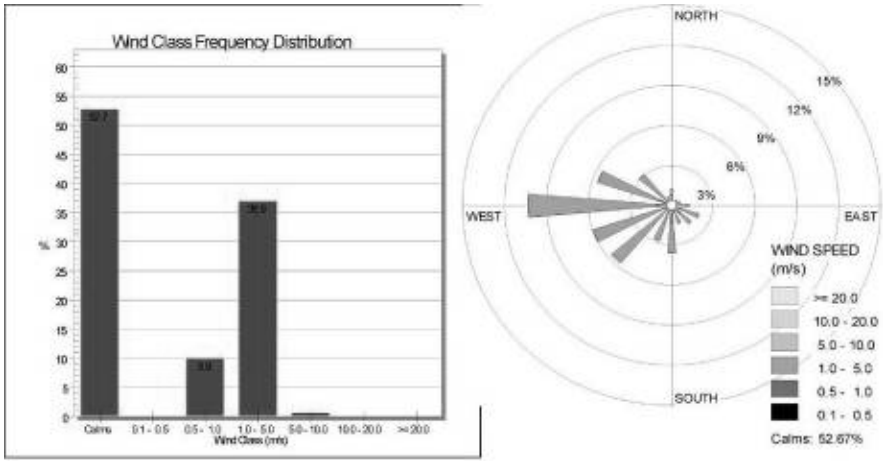


Figure 4. The class frequency distribution of wind speed captured by local meteorological stations, illustrated by histogram and a wind rose diagram.

3.5 Numerical simulation

Due to the time consuming numerical operations, standalone software tools are used for dispersion modelling over a part of the surface coal mine. Data linkage between the spatial information system and the inputs/outputs of the numerical simulation is provided by shared data files. The simulation is based on Reynolds averaged Navier-Stokes (RANS) equations for incompressible flows with turbulent closure of the model by the algebraic turbulence model [Kozel, 2007]. The DTM created in the spatial information system assists in a more precise modelling of wind flows over the surface. The output data are located in nodes of the horizontal layers. Each node contains values of the wind velocity vector, the atmospheric pressure and the dust concentration. The example in Figure 5 illustrates the lowest surface layer dedicated for the display of the wind speed and the wind direction over the area of interest. The project is developed in the ArcGIS environment. In this case, the simulation is carried out for one local dominant emission source, the temporary coal storage place. The spatial domain under consideration is represented by the rectangular grids included into 21 layers over the surface. The whole dataset is stored in the geodatabase and managed by the ArcGIS. In addition to the numerical simulation focused on one dominant emission source (temporary coal storage place), data from other numerical simulations based on short and long distance effects can be included into the scalable GIS environment. The resolution of the spatial data used for wind flow modelling in the local scales is based on surface laser scanning and precise GPS measurements complemented by aerial images. The middle scale studies focused on more emission sources with short and long distance effects are supported by spatial data obtained from standard digital thematic maps extended by satellite images. The limits of numerical simulations depend on used computer programs.

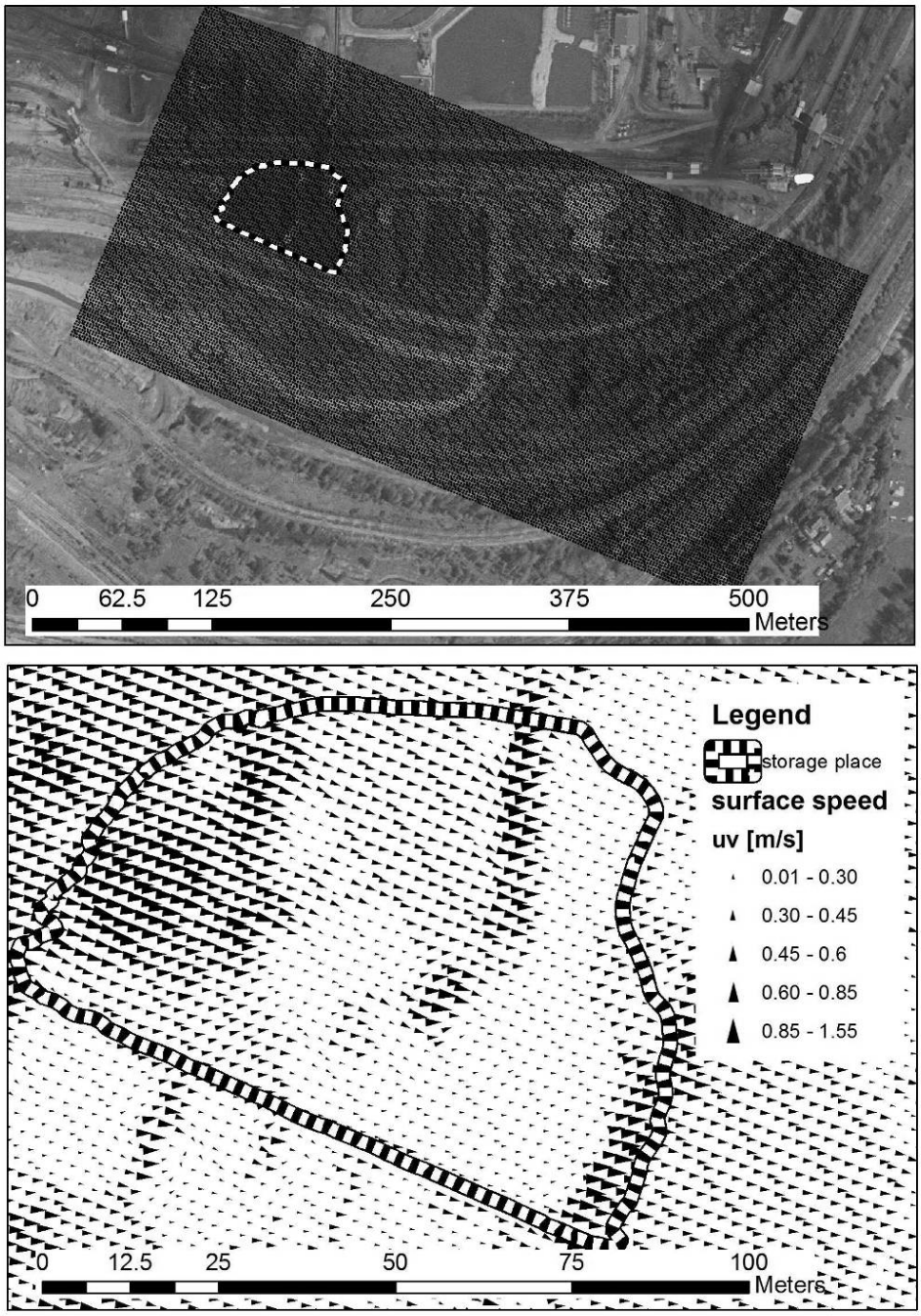


Figure 5. The spatial domain for numerical simulation of wind flows in the local scale complemented by the display of the wind speed and the wind direction over the surface of the dominant emission source, the temporary coal storage place.

4. CONCLUSIONS

The presented spatial information system based on ESRI's ArcGIS can manage and analyse a wide range of data in order to support risk assessment methods focused on dust transport in the area of the surface coal mine. The demonstrated task deals with an estimation of dust emission and dust transport from one of the dominant sources, the temporary coal storage site. Until recently, many studies have focused on estimates of dust rates from various surface emission sources or mining activities [Chakraborty, 2005; Ghose, 2000], on dispersion modelling over simplified surface objects (stockpiles and man-made barriers)

[Badr, 2005], and on numerical simulations based on US EPA models (ISCST3, ISC-PRIME, AERMOD, AERMOD-PRIME, or CALPUFF) [Wang, 2005]. But, the absence of advanced spatial analysis tools and visualization methods has not allowed more complex explorations in the framework of risk assessment [Matejicek, 2007]. In the case of the presented spatial information system, dust emission, dust transport and potential dust deposition can be studied together with other spatial data sources (satellite images, aerial images, existing thematic map layers, post-processed GPS measurements, pre-processed time series of meteorological variables, advanced methods for creation of the DTM-surface laser scanning). Thus, the scenarios of various site configurations, including man-made barriers and stockpile geometries, can be tested in order to minimize dust transport over the neighbouring residential zones.

ACKNOWLEDGEMENTS

This paper was carried out in the framework of a project supported by the Academy of Sciences in the Czech Republic (AVCR 1ET400760405). The spatio-temporal data and terrain measurements were processed by ArcGIS and ERDAS Imagine in the GIS Laboratory at the Faculty of Natural Science, Charles University in Prague supported by the Ministry of Education, Youth and Sports.

REFERENCES

Arctur, D., and M. Zeiler, Designing Geodatabases: Case Studies in GIS Data Modeling, ESRI Press, Redlands, CA, 2004.

Badr, T., and J.T. Harion, Numerical modeling of flow over stockpiles: Implication on dust emissions, *Atmospheric Environment*, 39, 5576-5584, 2005.

Chakraborty, M.K., M. Ahmad, R.S. Singh, D. Pal, C. Bandopadhyay, and S.K. Chaulya, Determination of the emission rate from various opencast mining operations, *Environmental Modelling & Software*, 17, 467-480, 2002.

Döllner, J., and K. Hinrichs, An object-oriented approach for integrating 3D visualization systems and GIS, *Computer & Geosciences*, 26, 67-68, 2000.

Finkelman, R.B., W. Orem, V. Castranova, C.A. Tatu, H.E. Belkin, B. Zheng, H.E. Lerch, S.V. Maharaj, and A.L. Bates., Health impacts of coal and coal use: possible solutions, *International Journal of Coal Geology*, 50(1-4), 425-443, 2002.

Ghose, M.K., and S.R.Majee, Sources of air pollution to coal mining and their impacts in Jharia coalfield, *Environmental International*, 26, 81-85, 2000.

Goodchild, M.R., Steyaert, L.T., and Parks, B.O. et al. (eds), GIS and Environmental Modeling: Progress and Research Issues, GIS World Books, Fort Collins, CO, 1996.

Kozel, K., T. Bodnar, E. Gulikova, and V. Pisa, Numerical simulation of flow and pollution dispersion over obstacles in complex terrain, Colloquium Fluid Dynamics 2007, Institute of Thermomechanics AS CR, 43-44, 2007.

Lu, H., and Y. Shao, Toward quantitative prediction of dust storms: an integrated wind erosion modelling system and its application, *Environmental Modelling & Software*, 16, 233-249, 2001.

Maguire, D.J., M. Batty, and M.F. Goodchild, GIS, Spatial Analysis, and Modeling, ESRI Press, Redlands, CA, 2005.

Matejicek, L., and Konradova, J., Presentation of environmental information to public with spatial information systems, 6th International Symposium on Environmental Software Systems (ISESS'07), 2007.

Mitchell, A., GIS Analysis: Spatial Measurements & Statistics: Volume 2, ESRI Press, Redlands, CA, 2005.

Mower, B., A multiple approaches to acute human health risk assessment, *Waste Management*, 18, 377-384, 1998.

Wang, L., D.B., Parker, C.B., Parnell, R.E. Lacey, and BW.Shaw, Comparison of CALPUFF and ISCST3 models for predicting downwind odor and source emission rates, *Atmospheric Environment* 40, 4663-4669, 2006.

Zeiler, M., Modeling our world, ESRI Press, Redlands, CA, 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

Cost Effectiveness Analysis for Renewable Energy Sources Integration in the Island of Lemnos, Greece

A. Angelis-Dimakis[a], P. Trogadas[b], G. Arampatzis[a] and D. Assimacopoulos[a]

[a] *National Technical University of Athens, School of Chemical Engineering, Process Analysis and Plant Design Department, 9 Heroon Polytechniou, Zografou Campus, GR-15780, Athens, Greece. e-mail: assim@chemeng.ntua.gr*

[b] *Department of Chemical and Biological Engineering, Illinois Institute of Technology, 10W 33rd st Perlstein Hall 127, Chicago, IL 60616*

Abstract: The development of more efficient and least cost energy management interventions is of great importance for isolated energy systems. Islands are typical examples of isolated regions, often highly dependent on imported fossil fuels but with a significant and often unexploited Renewable Energy (RE) potential. This paper presents a least cost planning approach towards the integration of Renewable Energy Sources (RES) in such systems, which is applied to the island of Lemnos, Greece. The approach involves the application of Cost-Effectiveness Analysis (CEA) and Incremental Cost Analysis (ICA) for screening possible alternatives and determining the most economically efficient and effective plan for their implementation. The objective of the application of the proposed approach in the specific case study is to meet through the use of RE technologies all the additional electricity and thermal energy demand, compared to 2007. Various supply side options are evaluated, and an implementation plan is derived. The results indicate that the excess of both electricity and thermal energy demand can be met in the near future without any significant changes in existing infrastructure, while other options should be considered for a more extended time horizon.

Keywords: Renewable Energy Sources (RES); Cost Effectiveness Analysis (CEA); Isolated Energy Systems; Incremental Cost Analysis (ICA); Implementation Plan.

1. INTRODUCTION

Isolated energy systems are characterized by geographical discontinuity, which causes difficulties of interconnection to the main electricity grid and increases the cost of transporting fuel. Furthermore, when there is limited potential for indigenous fossil fuel production, all the above, along with the increase in energy demand, can create problems that affect sustainable energy management.

Islands are typical examples of isolated regions, often highly dependent on imported fossil fuels, but also often having significant potential of Renewable Energy Sources [Rei et al., 2002]. According to the European Island Agenda [1997], the continuing exploitation of non-renewable energy sources is a provisional solution, inadequate to address energy problems in the long term. Thus, the higher penetration of RES is the only choice that could contribute towards energy autonomy, more stable energy prices and sustainable economic growth in these regions.

This paper outlines a methodological approach for planning efficient and least-cost integration of RES in isolated energy systems. The approach involves the application of

Cost-Effectiveness Analysis (CEA) and Incremental Cost Analysis (ICA) for screening possible alternatives and determining the most economically efficient and effective plan for their implementation, taking into account projections of future energy demand. The benefits of the proposed approach, as compared to optimization techniques, comprise the ease of implementation and the ability to define the optimal interventions within a complex range of alternatives through a transparent process. It should be noted that CEA and ICA do not identify a unique or "optimal" solution, but can lead to better-informed choices among alternative solutions, providing a basis for comparison of the relevant changes in costs and outputs on which such decisions should be made [Yao, 1992]. In such analyses, costs are typically calculated as the direct financial or economic costs of implementing a proposed measure, with effectiveness being defined in terms of some physical measure of environmental outcome [RPA, 2004]. Thus, the two methods provide results that can easily be interpreted and evaluated by policy makers and/or a wider audience.

Furthermore, and with regard to the specific goals of energy planning, the selection of CEA over traditional cost-benefit analysis allows addressing the different benefits of RE integration swiftly and objectively. Through the choice of appropriate indicators, local benefits associated with improved environmental quality, economic growth, job creation, increased control of energy production and energy supply security can easily be taken into account, while at the same time avoiding the time-consuming and often biased procedure of assigning monetary values to benefits.

Following this brief introduction, Section 2 presents the proposed methodology, and in Section 3 the methodology is applied to the island of Lemnos in Greece and the results are discussed. Finally, Section 4 summarizes the conclusions and makes suggestions for further research.

2. METHODOLOGY

Cost-effectiveness analysis (CEA) is a form of economic analysis that compares the relative expenditure (cost) and output (effectiveness) of two or more courses of action (options). The final result is a set of solutions (combinations of options) which can achieve the objectives set at the minimum cost. The identification of these solutions is performed through a relatively easy procedure, which consists of nine standard steps grouped in four tasks [Orth, 1994]. The 1st task focuses on the estimation of the cost and effectiveness of all available options, the examination of their compatibility and the formulation of the alternative solutions by combining these options. The measure of effectiveness is chosen to reflect the objective set as closely as possible. All cost and output estimates need to be measured over the same time period and in the same unit of measurement. That is, outputs and costs can be estimated either on an average annual basis or on a total output and cost basis [Robinson et al., 1995].

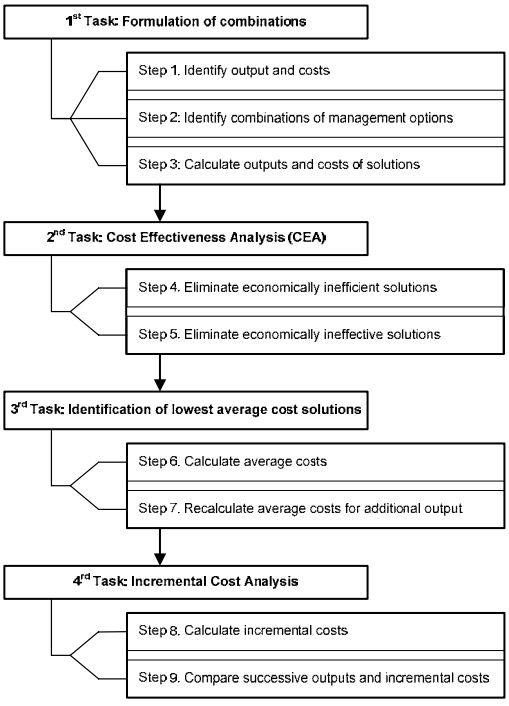


Figure 1. Cost-effectiveness analysis steps

The 2nd task performs the cost-effectiveness analysis, in order to eliminate inefficient or ineffective solutions. Inefficient are the solutions which, for the same amount of output, have greater cost than others, whereas ineffective those that for less output have the same or higher costs. The 3rd task calculates the average cost of the cost-effective solutions, in order

to eliminate those that have lower total cost but are relatively inefficient in output. Although this task is optional, it can help to eliminate distortions in the ICA, which is performed in the 4th task. The overall aim of ICA is to reveal changes in cost as levels of outputs increase, in order to determine whether the next level of the output is economically effective. Therefore, in this last task, successive solutions are compared against their incremental cost, in order to establish whether the next level of output is worth the additional monetary cost. The incremental cost (often mentioned also as marginal cost) is the change in cost that results from a decision, and is calculated by dividing the difference in cost between two successive solutions by their difference in output. The final step is the development of the incremental cost curve, a column chart that presents the incremental cost of the cost effective solutions. The incremental cost curve makes the relationship of cost and output for each alternative, as well as the variation in cost and output across alternatives, more visually apparent, in order to ascertain whether the next level of output is economically effective [Gerasidi et al., 2003].

According to Robinson et al. [1995], the following guidelines related to outputs, costs and the incremental curve can be used to assist in the decision making process:

(a) Curve Anomalies. Abrupt changes in the incremental cost curve, such as a breakpoint, a spike or a peak, which may indicate a sharp incremental cost increase, are potential decision points.

(b) Output Target. If a study has established a specific resource target to be met, then a decision rule could be developed to partially or fully meet that target.

(c) Output Thresholds. In some cases, it may be necessary to initially produce a minimum (or maximum) base amount of output and any lesser (or greater) amount would not contribute to the achievement of the objectives set. If such thresholds exist, they can be utilized to identify the range of acceptable solutions.

(d) Cost Affordability. If implementation funds are a constraint, then decision makers can review the curves that will help them identify the best investment for the funds available.

3. CASE STUDY

In this paper, the methodology described above is applied to the island of Lemnos, Greece. The island is located in the N.E. Aegean Sea (Figure 2), has a total land surface of 478 km^2 and a population of 18,104 (Census of 2001). The island is not connected to the mainland electricity grid. Its power system consists of a thermal power plant (5 fuel oil-fired engines with a total installed capacity of 19,640 kW and a diesel-fired engine of 1,500 kW) and a wind park of total installed capacity 1,140 kW. All the non-renewable primary energy resources are imported from the mainland. Hence, the use of RES in the island would be an important step towards ensuring energy autonomy and security.



Figure 2. Location of Lemnos Island

Previous works have estimated the electricity and thermal energy demand for 2007 at 53,530 and 67,170 MWh respectively [Angelis-Dimakis, 2005]. An assessment of the RE potential of the area has pointed towards the possibility for substantial integration of RE into the island's energy system. In more detail, the meteorological data of the island indicate fairly strong winds throughout the year (mean annual wind speed ~5m/s) and long sunny periods (mean daily solar radiation on horizontal surface ~4.8 kWh/m^{2}d). Geothermal resources also exist in Lemnos, with temperatures of hot waters reaching 50^{o}C.

The main agricultural product of the island is wheat, which leaves a great amount of biomass residues (almost 25,000 tons per year) that could be utilized for heating purposes [Trogadas, 2005]. The RE technologies chosen are wind turbines, photovoltaics, solar water heating systems, ground source heat pumps and biomass heating systems.

The overall goal of the case study is to define an implementation plan that could satisfy all the additional, as compared to 2007, electricity and thermal energy demand through the use of RE technologies. Three different scenarios have been formulated for the energy demand projection. The Business As Usual (BAU) scenario was based on historical data, assuming different energy demand growth rates for each sector according to the Public Power Corporation's forecast for electricity demand [Karalis et al., 2000]. The High Energy Demand (HED) scenario assumes the establishment of a University Department on Lemnos, an option which is currently under discussion, and a subsequent increase in population and energy demand. Finally, the Low Energy Demand (LED) scenario assumes that population increases at declining rates and that lower growth rates will be observed over the coming years. All the above scenarios have been formulated and analysed through the Long range Energy Alternatives Planning (LEAP) software, a package developed at the Stockholm Environment Institute for analysing energy balances and forecasting future energy demands.

Table 1. Alternative energy management options examined

Actions	Alternative Options	Energy Produced (MWh)	Annual Cost (€)	Code
Electricity Demand				
Wind Turbines	1×850 kW	2460	134280	A1
	3×850 kW	7230	290181	A2
	5×850 kW	11800	443902	A3
Photovoltaics (PV)	Myrina Hospital	47	23559	E1
	Myrina Town Hall	31	15689	E2
	100 Houses	360	207800	E3
	250 Houses	900	519500	E4
	500 Houses	1800	1039000	E5
Thermal Energy Demand				
Biomass Heaters (BH)	Moudros High School	149	7803	B1
	Moudros Lycee	62	6576	B2
	Moudros TEE	124	7241	B3
Geothermal Heat Pumps (GHP)	"Myrina Beach" Hotel	2230	68343	C1
	"Lemnos Village" Hotel	740	22964	C2
Solar Water Heaters (SWH)	100 Houses	240	12500	D1
	250 Houses	600	31250	D2
	500 Houses	1200	62500	D3
	1000 Houses	2400	125000	D4

Feasible and acceptable options were selected and evaluated (Table 1) after site investigation and consultation with local stakeholders. Wind turbines can be installed in the existing wind park, taking into account the limitations of the autonomous network concerning the total installed capacity. For this purpose, a medium size wind turbine is selected and all alternative options have been formulated so that the total installed capacity does not exceed 40% of the island's peak load. The remaining RE technologies can be installed in public buildings and households of the island's two largest towns Myrina, the capital, and Moudros. Moudros, located in a rural area, has been chosen as the suitable site for biomass exploitation technologies, in order to minimize the corresponding transportation costs. On the other hand, Myrina, where all the public services and hotels are

located, has been chosen for the installation of solar and geothermal technologies. The number of additional solar water heaters and photovoltaics was calculated taking into account (a) the number of available houses (~5000) and (b) the share of houses that already have a solar water heater installed (around 40%), further assuming that solar technologies can be installed in at least 50% of the remaining houses. As solar water heaters are a very popular technology in Greece, it was also assumed that they would have a larger penetration than PVs, accounting for 2/3 of the units installed.

Cost estimates presented in Table 1 correspond to the total annual cost for the implementation of each option, which comprises: (a) the amortization of the initial investment cost and (b) the annual operation and maintenance costs. Investment costs include design, transportation and installation costs, where applicable, and equipment purchase costs.

In order to account for all potential benefits from the exploitation of RES on the island's energy balance, the effectiveness of each option was expressed in terms of the annual renewable energy delivered by each option. The annual energy production for the renewable energy technologies was calculated using RETScreen, a standardized and integrated renewable energy project analysis software developed by the Natural Resources Canada's (NRCan) CANMET Energy Technology Centre - Varennes (CETC-Varennes).

As the target of the analysis is to satisfy both electricity and thermal energy demand, two separate analyses were developed, the Electricity Production Analysis (EPA) and the Thermal Energy Production Analysis (TEPA), using software developed at the Environmental & Energy Management Research Unit (EEMRU), an educational and research unit in the School of Chemical Engineering at the National Technical University of Athens (NTUA).

Figure 3 and Table 2 present the results of the Cost-Effectiveness Analysis. Out of the total of 65 (EPA) and 161 (TEPA) possible and acceptable solutions (combinations of options) that emerged as a result of Task 1, four and six solutions respectively have been identified as economically efficient and effective (Task 2). It should be noted that each point of Figures 3a and 3b represents the cost and effectiveness of the respective solution. The curve joining all the cost effective solutions is referred to as the Cost Effectiveness Frontier (CEF). All the remaining solutions located above and to the left of the CEF are economically ineffective or inefficient.

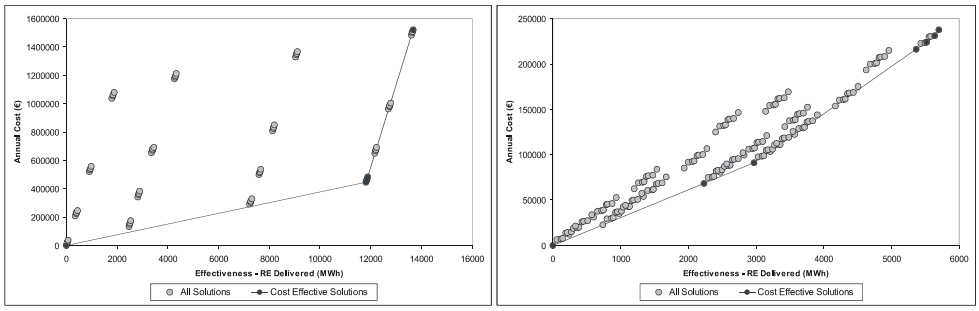


Figure 3. Estimated outputs and costs of all solutions and cost effective solutions for (a) electricity production analysis and (b) thermal energy production analysis.

Figure 4 presents the results of Tasks 3 and 4 of the ICA. Each column within the graph represents the incremental cost and incremental output associated with the respective solution. The difference in incremental cost between two successive columns indicates the per unit additional cost that should be paid in order to reach the next level of output. Results from the ICA show that in the set of cost effective solutions for meeting electricity demand there is a sharp incremental cost increase from the first (5×850 kW Wind Turbines) to the second level of output, which involves the introduction of photovoltaics in public buildings. This implies that additional demand-driven management options need to be examined before further capacity expansion is decided. On the contrary, in TEPA transition from one level of output to another is smoother, with the sharp increase in incremental cost being observed only for the production of the last 124 MWh.

Table 2. Cost effective solutions

Electricity Demand		Thermal Energy Demand	
Solution	Description	Solution	Description
EL1	5×850 kW Wind Turbines (A3)	TH1	GHP in "Myrina Beach" Hotel (C1)
EL2	EL1 + PV in Myrina Hospital (E1)	TH2	TH1+GHP in "Lemnos Village" Hotel (C2)
EL3	EL2 +PV in Myrina Town Hall (E2)	TH3	TH2 + SWH in 1000 Houses(D4)
EL4	EL3 + PV in 500 Houses (E5)	TH4	TH3 + BH in Moudros High School (B1)
		TH5	TH4 + BH in Moudros Lycee (B2)
		TH6	TH5 + BH in Moudros TEE (B3)

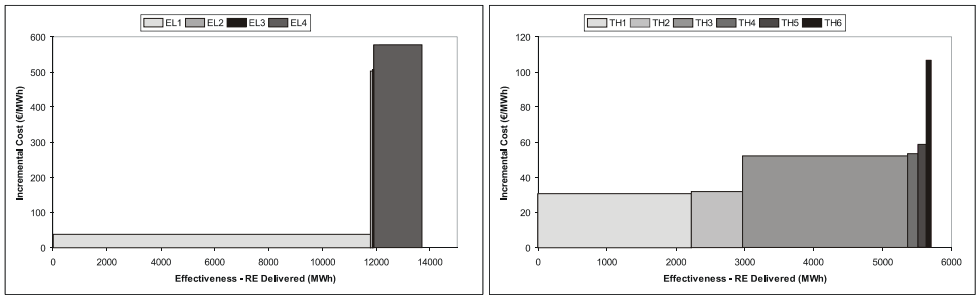


Figure 4. Incremental cost analysis for the cost effective solutions

While formulating an implementation plan, emphasis is placed on reaching the output targets set, i.e. meeting through RES all electricity and thermal needs exceeding the 2007 demand level. The ability of meeting the extra demand is examined for each of the three scenarios and the results are presented in Figures 5, 6 and 7. The curve expresses the estimated energy demand, electrical or thermal, in figures (a) and (b) respectively. The columns for every year correspond to the total energy that can be supplied, i.e. the sum of demand that is met in the base year and the additional demand that can be met through the implementation of the chosen options. The gap between the curve and the columns represents the demand that is still unmet.

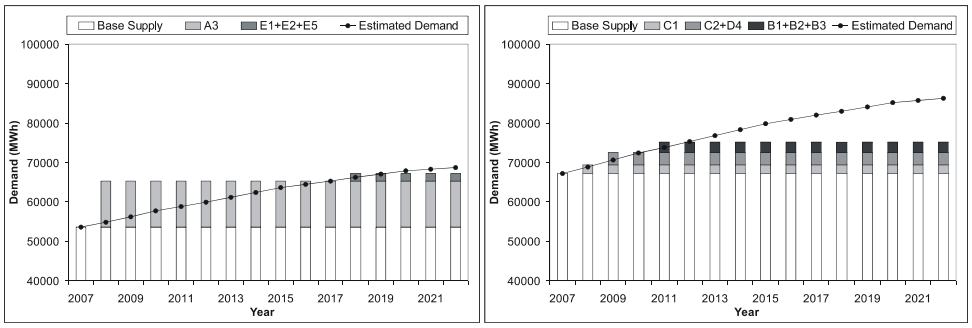


Figure 5. Formulation of the implementation plan for (a) electricity production analysis and (b) thermal energy production analysis. (BAU Scenario)

Results show that electricity demand can be met satisfactorily through RE technologies at least for the next decade, even under the most pessimistic forecasts. The total cost incurred is approximately 500,000€ for the BAU and LED scenarios, and increases up to three-fold in the HED scenario. On the contrary, the satisfaction of thermal energy demand is feasible only up to 2010 in the worst scenario (HED), or up to 2013, in the best scenario (LED). In a decade, the unmet thermal energy demand will be ranging from 5%, in the case of optimistic forecast (LED), to 16%, for the HED scenario. Meeting this demand would require (i) demand-side interventions (increasing efficiency in energy use), and (ii) measures aimed at inducing behavioural change towards lower energy consumption. A provisional plan for the implementation of the different options in each scenario is presented in Table 3. The interventions up to 2012 are of first priority. Given that the implementation of the remaining options corresponds to a sharp increase of incremental

costs, their implementation should be re-assessed, taking into account the actual evolution of energy demand by 2012.

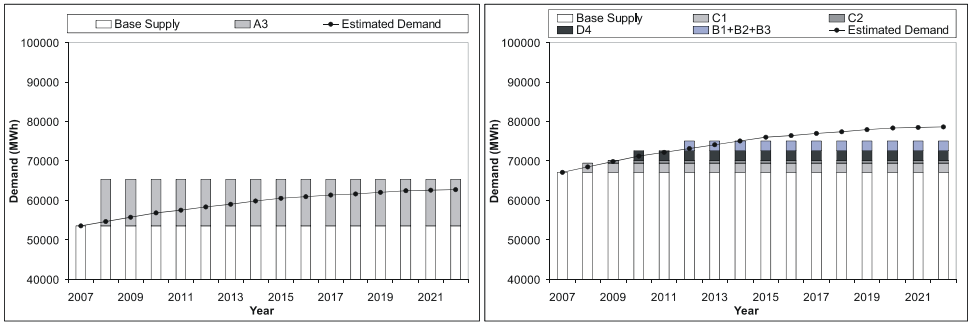


Figure 6. Formulation of the implementation plan (a) electricity production analysis and (b) thermal energy production analysis. (LED Scenario)

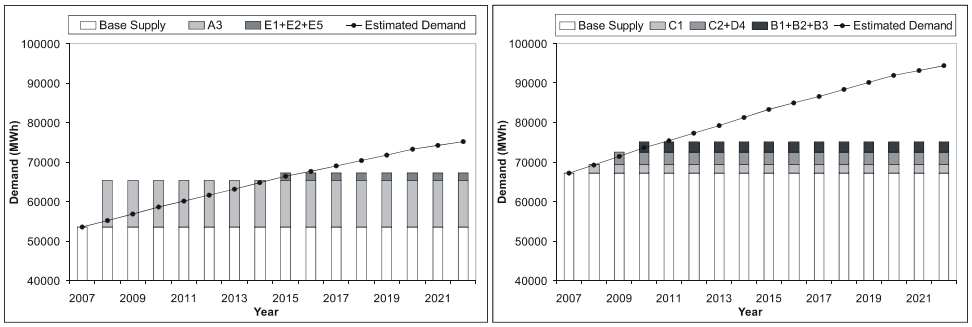


Figure 7. Formulation of the implementation plan for (a) electricity production analysis and (b) thermal energy production analysis. (HED Scenario)

Table 3. Implementation Plan for each scenario

Year	BAU	LED	HED
2008	5×850 kW Wind Turbines & GHP in "Myrina Beach" Hotel	5×850 kW Wind Turbines & GHP in "Myrina Beach" Hotel	5×850 kW Wind Turbines & GHP in "Myrina Beach" Hotel
2009	GHP in "Lemnos Village" Hotel & SWH in 1000 Houses	GHP in "Lemnos Village" Hotel	GHP in "Lemnos Village" Hotel & SWH in 1000 Houses
2010		SWH in 1000 Houses	BH in Moudros High School, in Moudros Lycee & in Moudros TEE
2011	BH in Moudros High School, in Moudros Lycee & in Moudros TEE		
2012		BH in Moudros High School, in Moudros Lycee & in Moudros TEE	
2015	PV in Myrina Hospital, in Myrina Town Hall & in 500 Houses		
2018			PV in Myrina Hospital, in Myrina Town Hall & in 500 Houses

4. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

This paper introduces a methodological approach for planning efficient and least cost integration of RES in closed and isolated energy systems, based on cost effectiveness and incremental cost analysis. The approach is applied to the island of Lemnos, where results show that the additional, compared to 2007, electricity and thermal energy demand can be

met in the near future, without significant changes in existing infrastructure. However, for a more extended time horizon, other options should be considered.

The proposed approach provides a consistent and easy-to-apply methodology for assessing energy management options according to their annual energy production and cost. As all selected options involve integration of Renewable Energy Sources, the annual energy production indicator does not only address the overall energy balance issue, but is also linked to wider environmental and socio-economic objectives. It should however be noted that the overall approach can easily be adapted to specifically address multiple objectives, through the selection of appropriate indicators representing economic, social or environmental criteria.

The environmental impact of the selected options has not been directly taken into account in this work; however environmental aspects, such as reduced emissions, protection of vulnerable ecosystems, etc., can be incorporated in the analysis to indicate trade-offs among the possible options. Individual decision makers can use these data to make planning decisions, depending on local development priorities and conditions.

Furthermore, an area of future research is to incorporate uncertainty in estimations of effectiveness and cost. According to Haimes [2004] uncertainty might stem from at least five different sources: (a) model topology, (b) model focus, (c) model parameters, (d) data used and (e) human subjectivity. In this regard, stochastic elements can be introduced to identify the impact of uncertainty in both the selection and the time-plan for the ranking of the different options.

REFERENCES

Angelis-Dimakis, A., Regional Energy Planning, Case Study: County of Lemnos, Dissertation, School of Chemical Engineering, National Technical University of Athens, Greece, 2005.

Gerasidi, A., P. Katsiardi, N. Papaefstahtiou, E. Manoli, and D. Assimacopoulos, Cost-effectiveness for water management in the island of Paros, Greece, paper presented at The 8th Inter-national Conference on Environmental Science and Technology, Lemnos, Greece, September 2003.

Haimes, Y., Risk Modeling, Assessment and Management, ISBN 978-0-471-48048-8, J. Wiley & Sons, 2nd Edition, 2004.

INSULA, UNESCO, European Commission, Consell Insular de Menorca, First European Conference on Sustainable Island Development, The Minorca Commitments, European Island Agenda, Ciutadella Declaration, Minorca, 1997

Karalis, G., A. Zervos, G. Panaras, and P. Garofallis, Towards 100% RES supply in the island of Lemnos-Greece, paper presented at the International Conference: RES for island tourism and water, Crete, Greece, 2000.

Orth, K., Cost-effectiveness analysis for environmental planning: Nine EASY Steps, U.S. Army Corps of Engineers, Water Resources Support Center, Institute for Water Resources, Alexandria Virginia, 1994.

Rei, P, JP. Fonseca, N. Duic, and MG. Carvalho, Integration of renewable energy sources and hydrogen storage in the Azores archipelago. *Proceedings of the International Conference on New and Renewable Technologies for Sustainable Development*, Ponta Delgada, Azores, 2002.

RPA, CEA and Developing a Methodology for Assessing Disproportionate Costs, Final Report for Defra, WAG, SE and DOENI, July 2004.

Robinson, R., W. Hansen, and K. Orth, Evaluation of environmental investments procedures manual, U.S. Army Corps of Engineers, Water Resources Support Center, Institute for Water Resources, Alexandria Virginia, 1995.

Trogadas, P., Study of Alternatives Scenarios for the Introduction of RES in Lemnos' Energy System, Dissertation, School of Chemical Engineering, National Technical University of Athens, Greece, 2005.

Yoe, C., Incremental Cost Analysis Primer for Environmental Resources Planning (draft). Prepared for the U.S. Army Corps of Engineers, Institute for Water Resources, Alexandria Virginia, 1992.

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

Implementation of a GEOdatabase to administrate global energy resources

M. Biberacher[a], **S. Gadocha**[a], **S. Gluhak**[a], **N. Dorfinger**[a], **D. Zocher**[a]

[a] *Austrian Research Centers GmbH – Research Studios, 5020 Salzburg, AUSTRIA*

(markus.biberacher@researchstudio.at)

Abstract: The future development of our energy system is still unclear. This uncertainty is based on numerous aspects. In addition to the currently most relevant climate debate this will also be a question of available technologies as well as of available resources. Hence the global energy resources and their distribution will be one of the most influencing factors in the design of our future energy system. To model and administrate global energy potentials a geographical (GEO) database has been implemented. The intention to implement such a global GEO database is based on the fact that numerous aspects describing our energy system rely on spatial characteristics. Particularly renewable energy resources depend on numerous datasets, the most important ones are topography, land cover, climate, population distribution and precipitation. With respect to these datasets the theoretical and furthermore useable global potentials of renewable energy carriers were estimated, in this paper the computation of wind power is shown exemplarily. When modelling renewable energy potentials special attention has to be paid to the fact that renewable energy potentials are often not additive, what means that the land surface is only available once and therefore several potentials exclude the option to yield another potential on the same area. For the estimation of the highest possible energy potential on a certain area, a competition analysis regarding the utilization of the land surface was carried out.

Keywords: Energy resources, GIS, Global Database, modelling.

1. INTRODUCTION

Numerous statistics give numbers on energy potentials on country level (IAEA 2005, BP 2006, EIA 2005, WEC 2004) but only few studies reflect the real spatial and temporal distribution of - especially renewable - energy carriers. One work treating this issue quite fundamentally has been done by Hoogwijk [2004]. Nevertheless also in this work the chosen spatial resolution is quite rough (0.5°). Since primary geographic information like topography, land cover, climate are available in a quite high spatial resolution for the whole globe (up to 1km), an estimate on the related renewable energy potentials can be evolved in the same high spatial resolution. That issue is treated in the current paper and a global GEOdatabase including conventional and renewable energy potentials is developed.
The GEOdatabase not only includes the consideration of the global spatial distribution of single conventional and renewable energy resources but also the influence of their spatial distribution in the context of the complete energy system and the coupling with spatial energy demand structures.

2. BASIC FRAMEWORK

The intention of the GEOdatabase is to consider all relevant energy resources with their geographical distribution in one common database. As the input all relevant raw data like solar insolation, wind speed, precipitation, topography, land cover, etc. are used on a global scale. These datasets are included in the database, each describing one aspect of the energy

system. The datasets are linked via the common geographical reference, whereby mainly renewable energy resources have been investigated. Therefore general data – describing land cover issues, climate conditions or population distribution and statistical datasets – are utilized to generate spatial estimates on harvestable renewable energy potentials as well as conventional resources. A scheme of the Geodatabase is shown in figure 1.

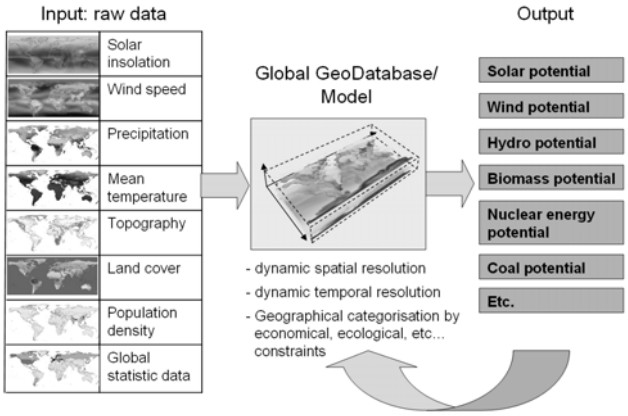


Figure 1: GEOdatabase framework to model and administrate global energy potentials

The challenge for the integration of raw data and the modelling process is the dynamical spatial resolution. For the renewable energy resources the data are proceeded and included into the database on a raster basis of different resolution up to 1 km , whereas other energy sources like nuclear or fossil energy are handled on a country or regional level. A further crucial point within the modelling of renewable energy potentials is the temporal resolution of their availability, which could be hourly, daily, seasonal or annual, depending on the considered energy carrier.

Hence the GEOdatabase includes a 3D data framework that enables the administration and treatment of two space and one time dimension.

3. GENERAL MODELLING APPROACH

3.1 Relevant Data

The global energy resources depend on numerous datasets which represent physical, ecological and economical restrictions on the available potential. Non-renewable resources are mainly determined by their global spatial distribution by type of resource (conventional/non-conventional), level of confidence that the respective resource exists as well as by recovery costs. Regarding renewable energy potentials the most significant general influencing factors - next to specific data like solar insolation or wind speed - are topography, land cover, climate and population distribution. Although these restricted data are not sufficient to provide an accurate description of the real situation, they offer a good basis to estimate the renewable energy potentials. By implementing and interlinking these datasets on a global scale, the potential energy harvest of a respective energy resource in different areas of the globe can be estimated in a proper way with respect to local/regional conditions and included in one common database.

- **Topography**

Topography is considered as one of the main influencing factors regarding the spatial distribution of energy resources. The physical availability of resources is influenced by topography as well as a possible harvest or depletion of energy resources. That affects all forms of renewable energy carriers as well as the depletion of conventional energy resources. Hence the topography has to be included in the modelling approach.

- **Land cover**

A land cover classification which distinguishes 13 land cover classes – representing the generalised main land use categories - was used [Hansen et al, 1998]. These land cover classes are a major input to extrapolate available renewable energy resources as they are

correlated to the possible energy potential for the decision where an energy potential could be harvested.

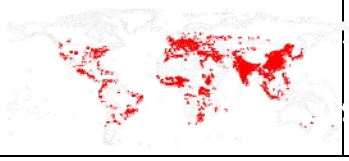
- **Population density**

The global population is also considered as relevant for the estimation of energy potentials. Next to the factor of reachability also the aspect of land use competition is an aspect that is mainly influenced by the distribution and density of the population. Data on population distribution were derived from the Center for International Earth Science Information Network (CIESIN) [2005].

- **Accessibility**

Derived from the mentioned primary datasets the so called "accessibility" can be identified as a secondary dataset. This variable is determined by the global spatial distribution of demand pattern that has to be satisfied by the spatial distribution of supply pattern. On a global scale, some regions might be identified with a high theoretical renewable energy potential, but are far away from already built or economically feasible infrastructure. To estimate such effects a sensitivity analysis is carried out. This analysis deals with the spatial distribution of regions which enclose a certain threshold of inhabitants within a certain area as shown in table 1. Although “accessibility” is defined by several aspects, whereas population density can be identified as the most significant one. Therefore accessibility was modelled under the simplified assumption that only population density affects the accessibility of renewable energy potentials. This approach to define accessibility is applied to categorise the different energy resources.

Table 1: Distribution of accessible areas depending on inhabitants within a certain area

	radius of 100 km	radius of 200 km
Inhabitants in area > 10 * radius²		
Inhabitants in area > 100 * radius²		

3.2 Estimation of global energy resources

The estimation of the energy potentials is accomplished in a top-down approach. This means that in a first step the physically available energy potential of the respective energy carrier is calculated, which is then reduced to a realisable potential by further constraints like technical efficiency factors or politically and economically influenced aspects. So far solar, wind and hydro power, biomass, fossil and nuclear resources have been included in the database. While the renewable resources are modelled on a spatial resolution of 1 km² for the whole globe and a time resolution varying from hourly to yearly cumulated values, the conventional resources are treated on a national cumulated level reflecting the known remaining recoverable resources.

3.3 Global Wind Power Potential - Case study

3.3.1 Methodology

The calculation of the global wind power potential is based on the average wind speed, topography, land cover, population and the energy harvest of a typical wind turbine. Therefore the physically available potential of wind power is computed in the first step, which is done by calculating the energy harvest gained from a 2.3 MW turbine depending on the average wind speed [wind data from NASA, 2007]. This is accomplished on a

temporal basis of 3 hours because the diversification of the wind supply is quite high as can be seen in figure 2 and a high temporal resolution has to be used for further investigations in terms of suitability for covering the energy demand on time.

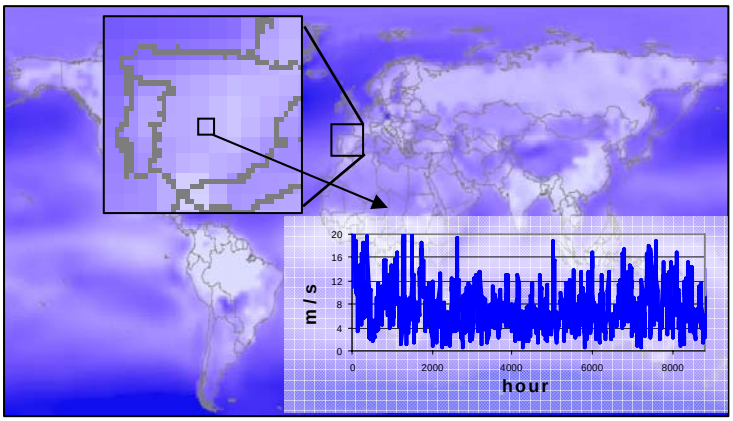


Figure 2: Temporal diversification of wind potential

The size of a 2.3 MW turbine is chosen because it is considered as state of the art. Regarding the size of a turbine and resulting wind shadow effects, a 2.3 MW turbine and resulting distances between two turbines could be accepted as reference to estimate a harvestable wind power potential per area. The power curve of the wind turbine as a function of wind speed is shown in figure 3 [Danish Wind Industry Association, 2007]. In sequence the energy output for the physically available global potential from wind power per surface area can be generated. For this the real surface area is considered. A further restriction that has to be taken into account is land cover. It is assumed that for an intensive utilization of wind power the average power density sticks out to be 4.6 MW/km^2 [Hoogwijk, 2004]. The estimation of the average upper limit of installable wind turbines depending on the designation of the surface was performed on the basis of land cover categorisation. With the combination of the upper limit of wind turbines per area and the energy harvest of a 2.3 MW turbine a global theoretical potential is identified. This potential is then further restricted by neglecting single quadrants, depending on topography. An upper threshold of 2000 m refers to high mountain areas that are not reasonable for wind turbine installations. A lower threshold of – 40 m refers to offshore regions, which will be too deep for turbine installations [Hoogwijk, 2004].

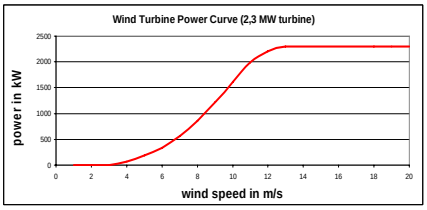


Figure 3: Power curve of a 2,3 MW wind turbine

3.3.2 Categorisation of wind potentials

Limiting the global wind power potential to the restricted usable potential for energy issues provides the upper limit of the usable global wind power potential. This potential has to be categorised, which is carried out by the assumption of two attributes for the global wind power potential: *Availability* and *distance to population.*

The *availability* is defined as the annual average energy harvest of single wind turbines, translated to corresponding full load hours. Three levels are distinguished: *low (<800 full load hours / year), mean (800 - 3000 full load hours / year)* and *high availability (>3000 full load hours / year).*

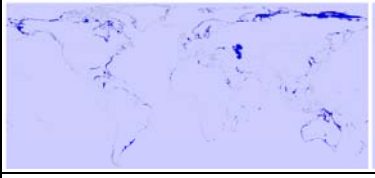
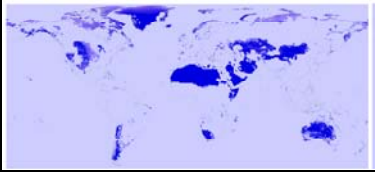
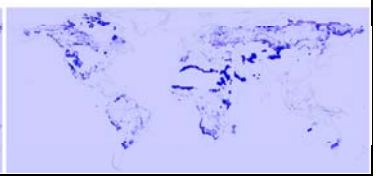
The attribute *distance to population* is chosen as the cumulated population higher than 100000 within a circle of 200 km diameter. Four categories are distinguished: *offshore*

short distance, *offshore large distance*, *onshore short distance* and *onshore large distance*. The categorisation is based on the global population grid and the elevation.

That leads to four different categorisation classes, based on the permutation of on-/ and off-shore with two accessibiliy classes.

Dividing these categories further on into three natural availability classes *low*, *average* and *high*, 12 individual categories could be distinguished. Each single geographic location is assigned to one of these categories [Biberacher et al., 2007]. Four of these categories are shown exemplarily in table 2.

Table 2: Categorisation of wind potentials

Availability of wind power potential		
	mean	low
Off shore potential far away from populated areas		
On shore potential far away from populated areas		

As exemplarily outlined for the global wind power potential also the other mentioned potentials are investigated.

3.4 Hydro power

For the estimation of the hydro power potential datasets of the global distribution of precipitation and topography [Worldclim database, 2007] of the globe were used. The description of topography allows the calculation of the drainage system for drain water, while the precipitation data provide the base for the quantitative estimation of drain water. By interlinking both datasets it is possible to calculate the gravity potentials of cumulated rain runoffs over different terrains. This potential is then further reduced by several constraints, for example the restriction that only potentials beyond a certain threshold are considered. Further restrictions are caused by issues like water back hold.

3.5 Biomass

Regarding biomass, woody (forestry plantations, natural forests and natural woodlands) and non woody biomass resources (grassland and cropland) were taken into account. Processed waste or by-products of agro-industrial activities are not considered. For the estimation of the biomass potential the most relevant datasets are land cover, climate data and population data. For the different land cover classes a mean net production of biomass following Heinloth [2003] was set as well as a factor of possible energetic use, also depending on population density. As far as the net production of biomass is concerned the climatic parameters precipitation and annual mean temperature are considered as the main influencing factors [Lieth, 1972; Grieser, 2006] and are

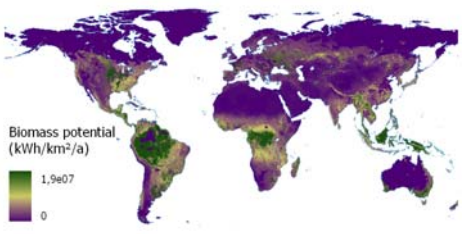


Figure 4: Biomass production in kWh/km^2/a

therefore included in the modelling process. The biomass production - restricted by the accessibility - in energetic units is shown in figure 5. The possible share of biomass that can be used for energy production was based on Heinloth [2003] and own assumptions. As a last step an energy conversion factor was included to estimate the global biomass potential.

3.6 Solar power

Solar energy potentials were calculated mainly based on global irradiation [Ramachandra and Shruthi, 2007] taken from NASA [2007], topography and land cover. It is estimated that the upper limit of available surface share for solar collectors is determined by the real land cover. For example is assumed that 1% of the urban and built up area can be used for the installation of solar collectors following Sörensen [2001]. Furthermore two constraints were included in the computation to generate different categories of potentials, availability – depending on the annual global insolation per m^2 - and accessibility.

3.7 Nuclear resources

Uranium, Lithium and Thorium resources are included in the database on a national level. These resources were categorised regarding the level of confidence that the resource exists, given how much exploitation has been done as well as by recovery costs [IAEA, 2006]. Regarding Uranium only conventional uranium resources, where uranium is the main product of extraction processes are considered. As far as Thorium and Lithium are concerned all currently known resources were included in the database according to IAEA [2006].

3.8 Fossil resources

All currently known and proved recoverable resources of crude oil, gas and coal were added to the database. These data were estimated based on a literature research [BP, 2006; WEC, 2002; EIA, 2004]. The preserved data were evaluated and thereafter the fossil resources were defined.

3.9 Competitive land use

Special attention has to be paid to the fact that renewable energy potentials are often NOT additive potentials. This means that land surface is only available once and therefore several potentials exclude the option to use another potential on the same area.

To estimate the highest possible energy potential derived from a certain area a competition analysis – regarding the utilization of the surface – has to be carried out. That means that land surface dedication is NOT considered to be changed in order to increase the utilisable energy potential (otherwise in most cases solar power would be the most competitive energy carrier). The challenge of a competition analysis is to identify the maximum renewable energy potential for each location on the globe.

To simplify the analysis a possible double usage per raster cell is neglected and only the most promising renewable energy carrier – depending on land cover classes (GLCF - Global Land Cover Facility) [Hansen et al., 1998] - is considered to determine the upper bound of utilisable energy potential per raster cell.

This analysis refers to primary energy units and therefore lines out which energy density per location could possibly be harvested if a technology without efficiency losses would be available. This could also refer to an energy carrier that is not the most competitive one at this location regarding current available technologies. Based on this approach a maximum upper limit of the renewable energy potential derived from the competition of biomass, solar energy and wind energy is determined on a global scale (see figure 6).

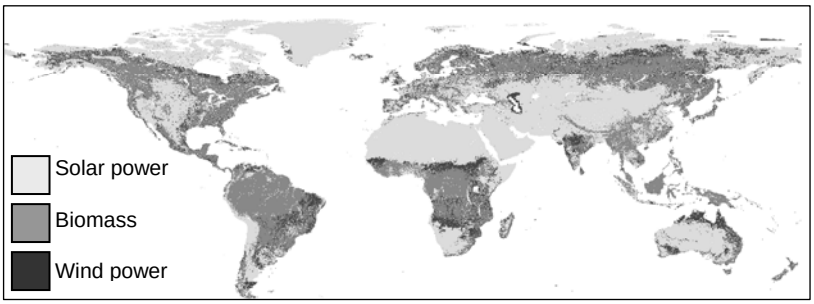


Figure 5: Geographic distribution of competitive renewable energy source (referring to primary energy units). Each grey value identifies one resource as the most competitive one.

This competitive land use calculation can be part of a dynamic parameterisation, in terms of regional deviations, land use adaptations and further aspects. Table 3 outlines numbers derived for the different potentials within single world regions. These numbers are subject to an individual parameterisation and have to be evaluated under these assumptions made.

Table 3: Land use competition regarding a parallel utilization of dedicated area.

Region	Solar total	Solar under competition constraints	Wind on-shore total	Wind under competition constraints	Biomass total	Biomass under competition constraints	Single additive total Potentials	Maximal additive Potential under competition constraints
	PJ / Year	PJ / Year	PJ / Year	PJ / Year	PJ / Year	PJ / Year	PJ / Year	PJ / Year
Other Asia	484128	482656	81832	7186	53345	38311	619304	528160
Africa	4622148	4621507	483224	42930	92707	53334	5198080	4717793
Former Soviet Union	816120	811217	240451	34819	47840	19120	1104412	865202
East Europe	40032	39550	12496	2308	7034	1703	59562	43560
West Europe	119232	114530	58698	17708	15433	4504	193363	136854
Australia	457632	456232	157889	28552	19649	5033	635170	489816
Middle East	1556784	1556010	147834	1501	4468	392	1709086	1557904
Central & South Amarica	393336	381989	100872	40338	142852	116093	637060	538427
Canada	288036	284818	97542	12773	15890	10145	401468	307764
China	790740	790542	99547	4928	31777	10224	922064	805705
India	139428	137182	18792	5152	15962	5702	174182	148061
Japan	7272	7142	4219	1930	1775	1084	13266	10152
South Korea	1584	1548	670	238	558	331	2812	2117
Mexico	78480	78145	15088	3564	7096	4028	100663	85745
USA	432288	429534	127055	18432	34564	13068	593906	461038
World	484128	482656	81832	7186	53345	38311	619304	528160

Since the database forms one common platform to model and administrate all energy potentials and accompanied influencing factors, the comparison, regarding differences and interactions between single potentials can be investigated in a convenient way.

4. CONCLUSIONS

The database provides a suitable and flexible platform for the standardized administration of data on renewable, fossil and nuclear energy resources. Some data, especially those referring to renewable energy potentials are included on a raster basis, others - mainly nuclear and fossil energy resources - are provided on a country or regional basis. Also a dynamic temporal resolution is included in the database to illustrate if and how the possible energy potential might alternate in an annual, seasonal or even hourly period.

One influencing factor for the future energy supply is the accessibility of energy carriers. Especially for renewable energy carriers their vicinity to demand is important, as their energy density is not very high. Accessibility also refers to the distance to existing or economically feasible infrastructure and therefore the inclusion of all relevant data on energy potentials in the database can show preferable areas for locations of energy supply. Although the data included do not provide an accurate description of the real situation, they offer a good basis to approximate the global energy potentials. Since primarily global datasets are utilized in the approach shown, a good reliability regarding comparisons of

different world regions is ensured. A possible adjustment to regional datasets will be investigated in further steps.

The innovation of this approach lies in the model based correlation of single datasets in order to estimate utilizable potentials dynamically - dependent on differing input assumptions. Furthermore this dynamic treatment of single energy resources enables their investigation regarding different land use competition aspects by joining the single results. On a global scale the model based estimations on energy potentials - depending on dynamic constraint assumptions - reflect stated numbers in common literature, e.g. WEC [2004]. Statements on a subregional or even local scale are accompanied by a high uncertainty. This uncertainty is based on the fact that primarily global datasets are investigated. Furthermore the location dependent variance of datasets, also linked to the spatial resolution, is quite high.Strategies for an optimal interaction of different energy resources for the satisfaction of the still globally rising energy demand can be derived from these data. Further investigations will also treat ecological and economical impacts on the exploitation of resources, namely referring to learning curves of technologies and therefore possible deviations in future estimates of utilizable potentials.

REFERENCES

Biberacher, M., Modelling and optimisation of future energy systems using spatial and temporal methods, VDM Verlag, ISBN 978-3836424356, Germany, 2007.

Biberacher, M., Gluhak, S., Zocher, D., Dorfinger, N., Labriet, M., Lechon, Y., Cabal, H., EFDA Technology Workprogramme 2006; EFDA-TIMES MODEL: Resource potentials update, Final Report: Appendix B (GIS Approach), October 2007, to be published

BP, Brithish Petroleum, Statistical Energy Review, 2006; http://www.bp.com/ (12.06.2007)

Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). 2005. Gridded Population of the World Version 3 (GPWv3). Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University.
Available at http://sedac.ciesin.columbia.edu/gpw. (22.03.2007).

Danish Wind Industry Association,
2007,http://www.windpower.org/en/tour/wres/pow/index.htm (03.03.2007)

EIA, International Energy Annual 2005, http://www.eia.doe.gov/ (27.05.2007)

Grieser, J., Gommes, R., Bernardi, B., The Miami Model of climatic net primary production of biomass, FAO of the UN, Italy, 2006

Hansen, M., R. DeFries, J.R.G. Townshend, and R. Sohlberg (1998), UMD Global Land Cover Classification, 1 Kilometer, 1.0, Department of Geography, University of Maryland, College Park, Maryland, 1981-1994

Heinloth, K., Die Energiefrage, second revised edition, Vieweg Verlag, Braunschweig, ISBN 3-528-13106-3, 2003

Hoogwijk, M., On the global and regional potential of renewable energy resources, ISBN 90-393-3640-7, Netherlands, 2004.

IAEA, Uranium 2005: Resources, Production, Demand; OECD, International Atomic Energy Agency (“The Red Book”), 2006

Lieth, H., Modelling the primary productivity of the world, Nature and Resources, UNESCO, VIII, 2; (1972) 5 – 10

NASA, Surface meteorology and solar energy 6.0, NASA Langley Atmospheric Science Data Center, distributed Active Archive Center, http://eosweb.larc.nasa.gov/ (04.09.2007)

Ramachandra, T.V., Shruthi, B.V., Spatial mapping of renewable energy potential, *Renewable and Sustainable Energy Reviews*, 11(7), 1460 – 1480, 2007

Sørensen, B., GIS management of solar resource data, *Solar Energy Materials & Solar Cells*, 67(1–4), 503 – 509, 2001

WEC, World Energy Council, Survey of Energy Resources, 2004; http://www.worldenergy.org/ (27.05.2007

Worldclim database, http://www.worldclim.org/current.htm (01.02.2007)

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 Model Platform for Sustainable Energy Planning. Colombian case.

R. Quijano[a] and J. Domínguez[b]

[a] *Economy, Environment, Energy Model Analysis Team. National University of Colombia. Sede Medellín Carrera 80 #65-223(rquijano@unal.edu.co).*
[b] *Renewable Energies Division. Energy Department, CIEMAT. Avd. Complutense, 22, 28040 Madrid. Spain (Javier.dominguez@ciemat.es)*

Abstract: Colombia has relied for long on a generous endowment of fossil fuels (oil, coal and gas) and hydropower to meet domestic energy needs and also to contribute substantially to the balance of trade in international markets. But this situation is about to change due to the announced loss of self-sufficiency in oil (forecasted by 2010-2011) unless a dramatic last minute change in proven reserves occurs. Likewise, Colombia's hydro potential (the largest energy source for power generation in the country) is facing difficulties to add further representation within the generation mix. These difficulties include increasing environmental and social costs associated with large hydro projects and the likely impacts of climate change and climate variability (drastic increases in surface temperature in the Andes and increases in intensity and frequency of ENSO -El Niño/Southern Oscillation- signals driving prolonged periods of drought). On the other hand, the country is richly endowed with wind and solar energy, resources that still have to participate more significantly in the national energy mix. Optimization of the nation's energy options, however, is not a trivial exercise. Decreasing reserves of natural gas, dwindling oil resources and vulnerable hydropower generation due to climate variability will likely force the power sector to increasingly seek alternative options to the current power mix. In this sense, National University of Colombia and CIEMAT have been planning a project in order to build a new model tool for energy planning. This tool will be developed by integrating environmental and social constraints as well as renewable energy sources (RES) (according to the current Colombian National Power Plan 2006 – 2025). This RES integration will be made using geographical information systems (GIS) technology. In this paper, the project is presented, as well as the state of the project, the basis of the model, the software to be used and the designed platform.

Keywords: sustainable energy, Environmental Decision Support Systems, renewable energy, geographical information systems, multi-criteria decision analysis.

1. INTRODUCTION

Colombia's primary energy supply is equivalent to 74 MMTOE/year. About half is produced by hydropower and the balance is obtained by fossil fuels (oil, coal and gas), 70% of which is imported. The energy demand makes up by large and growing requirements by the transport sector, followed by industry and the domestic sector. The annual electricity consumption is 49 GWh and the average electricity consumption per capita is 1113 kWh/yr (although this consumption is rising, between 1999 and 2004 decreased). CO2 emissions are 56 MMT (1.26 t CO2/capita), or less than half of world's average. Colombia's energy intensiveness is 0.1 TOE/ thousand -2000 US$PPP [WORLD-BANK, 2006] (see Table 1). This is equivalent to 10.1 dollars of product (GDP) generated per unit of energy use (2000 PPP$/kg oil equivalent), a high rate that shows Colombia's efficiency in energy use

compared to a 6.2 for the Latin American and Caribbean region and 5.2 for high income countries [WORLD-BANK, 2006].

Table 1. Colombia Energy Balance. (Source: Ministerio de Minas, 2006 and others. Elaborated by authors).

Energy Source for electricity power generation	Generation GWh (2005)
Hydropower	39,576.1
Thermal	9519.4
Renewables	49.48
Exports	1,757
Imports	36.9

Colombia is particularly vulnerable to the impacts of climate change. The first national communication (NC 1) to the United Nations Framework Convention on Climate Change [UNFCCC, 2005] indicates the high vulnerability of Colombia to expected impacts from climate change [IDEAM, 2001] identifying health, high mountain habitats, and insular and coastal regions as the topics of primary concern. More recently, studies commissioned as part of the preparation of the second communication have confirmed and indicated in more detail trends and impacts in these areas. These vulnerabilities find echo in the findings of the Intergovernmental Panel on Climate Change (IPCC). The energy sector in Colombia is also vulnerable to the ongoing climate modifications associated with global warming. Major vulnerabilities related to the energy sector are described below:

- Coastal energy infrastructure is vulnerable to extreme weather events and coastal flooding potential, major power plants are located in coastal zones that may become susceptible to floods with projected sea level rise and changes in storm surges.

- Hydropower potential is vulnerable to changes in precipitations and warm temperatures leading to high ratios of evaporation. Ecosystem changes driven by warming temperatures are anticipated to negatively affect water retention rates, increase pace of sedimentation and reduce runoffs.

- Rapid glacier retreat will affect water regulation in the Andes and thus reliability of dependent installed hydropower capacity.

- Biggest intensity and frequency of ENSO signals will reduce rainfall and affect the share of hydropower.

In summary, decreasing reserves of natural gas, dwindling oil reserves, vulnerable hydropower generation due to weather variability (El Niño(a) events) and climate change (temperature and precipitation changes), will force the power sector to increasingly seek alternative options to the current power mix. In this sense, several options are included like increasing the hydropower generation with additional safety margins to cope with climate considerations, increasing the use of coal or alternatively, increasing the reliance on clean non-conventional energy sources; or a combination of some of the previous ones.

This situation makes advisable the development of a new energy model that integrates not only the purely technician-economic variables but also social and environmental aspects [WORLD-BANK, 1992]. This model should be developed from a perspective that will lead to cope with the necessities of most possible Colombian territory.

In this context, the project has innovative characteristics that represent a strong difference with respect to other projects or initiatives, because the main goal is “to integrate” several approaches in order to claim a complete and integral vision of energy problem from geographic, environmental, social and technologic point of views

.

Next, we will describe the main aspects of the project plan, indicating the initial hypothesis, the proposed work methodology, the contributions that are hoped to reach with the development of the project, as well as its current state.

2. HYPOTHESIS AND GOALS

The aim of the project is to design and build a methodology for energy and environmental planning in an integrated platform. It will include the evaluation of environmental impacts generated by energy production and use, the simulation of an energy matrix based on renewable energies and the required information for the decision-making in the design of energy and environmental policies in the medium and long term.

The guidelines will be based on international criteria about sustainable development. These lines will be developed applying different techniques of energy-cycle analysis, in order to get a global assessment at National scale [IAEA, 2005, UN_DESA, 2001, Vera, I*., et al.*, 2007].

Models are usually developed to address specific questions and therefore they are only suitable for the purpose they were designed for. Incorrect application of a model may result in significant misinterpretations. This fact cannot be ascribed to poor model functioning but, as the World Bank [WORLD-BANK, 1991] –among others– claims, the misuse could be responsibility of the model users. Our project will be based on two different types of models. On one hand, general purpose models (forecasting, exploring, backcasting), and on the other, more specific purpose ones (e.g., demand-supply analysis, environmental impact analysis, etc), which will be the most meaningful within our project.

3. METHODOLOGY

The methodology focuses on three main issues:

- The energy load and supply and the environmental impacts produced by the energetic activity, the renewable energy in the national energy matrix and the promotion of sustainable energetic resources.
- The integration of the renewable energy resources (RES) based on geographic information systems (GIS)[Dominguez, J*., et al.*, 2007].
- The sustainability of the energy process based on the environmental impact, externalities and Life Cycle Analysis (LCA) assessment [Lago, C*., et al.*, 2007] and the definition of a sustainable matrix [Smith, R*., et al.*, 2000].

The first issue focuses on the integrated planning models (load-supply) and its conception. This vision should include the definition of the requirements for the development of activities that are very intensive in terms of energy demand. The project includes the analysis of the different models that can be useful for the *formulation line* (a specific line of strategic planning in Colombia) and its application. Decision makers and analysts require a comprehensive analysis of different elements and components, many times at odds, that provide objectivity and truthful information to check the approach, correlation and the validity of the relationships among the same ones [Frangopoulos, C. A*., et al.*, 1997].

The second one is related to the integration of RES. GIS technology will be used in order to overcome the current supply-demand model [Dominguez, J. y Amador, J., 2007]. By means of the geographic analysis, we will build a strategic planning based on constraints, simulating scenarios of supply–request with different technological and available possibilities. The evaluation of resource potential (solar energy, wind and biomass) will be emphasized in order to improve the penetration of new energy sources. The result should be a technological mix, sustainable and hierarchical, build upon resources, potentials and technologies. The potential of renewable energy that can be introduced into the energy

basket will be considered in the sustainable matrix along with the respective assessment of LCA and the up-to-the-minute models of planning with new scenarios.

Finally, we will build a sustainable matrix, in order to intersect the environmental impacts and the external costs or externalities associated with the energy production cycles. We will count on the compilation of the original information with the results of the LCA of the several energy alternatives and the available information of externalities' assessment in the development of the matrix [Lago, C*., et al.*, 2007]. In addition, we will take into account sustainable indicators [IAEA, 2005, UN_DESA, 2001, Vera, I. y Langlois, L., 2007] for the comprehensive assessment of scenarios. The decision-making analysis, with an objective assessment, will be made for all the activities of the energy cycle. The matrix will be set up with variables, indicators, index and criteria, all of them delimited under the same technological, environmental, social, economic and cultural dimensions. The introduction of the energy consumption structure, can modify the composition of the energy structure creating sustainable scenarios [Smith, R*., et al.*, 2000] and closing the cycle of the comprehensive model, this way will feedback the results of the matrix. The change of the energy structure shows alternatives and elements in order to develop a sustainable environmental energy policy new.

4. CONTRIBUTIONS

The result of the present research will be a simulation tool for energy and environmental planning. This tool will join GIS, externalities assessment, LCA and a *sustainable energy matrix*. In several aspects, this approach is innovative compared to the current sectorial visions. The integrate conception of the model, the inclusion of social and territorial aspects, the prospective vision based on local and renewable resources as well as environmental issues structured in a sustainable matrix, could convert this develop in a powerful tool for integral energy planning that contributes to the takeoff of clean energy in Colombia.

5. CURRENT STATE OF THE PROJECT

During last year, National University of Colombia (UNAL) and Centre of Energy, Environmental and Technological Researches (CIEMAT), have been planning a project about the viability of building a new model tool for energy planning, integrating environmental and social constraints (land use, resources competition, natural protected areas… in function of the different parameters of each technology) as well as using geographical information systems (GIS) technology in order to integrate renewable energy sources.

The work flow chart of the model has already been defined (see figure 1), the framework of the integrated platform has been designed, the election of the software has been made and a preliminary draft of the sustainability matrix has been designed.

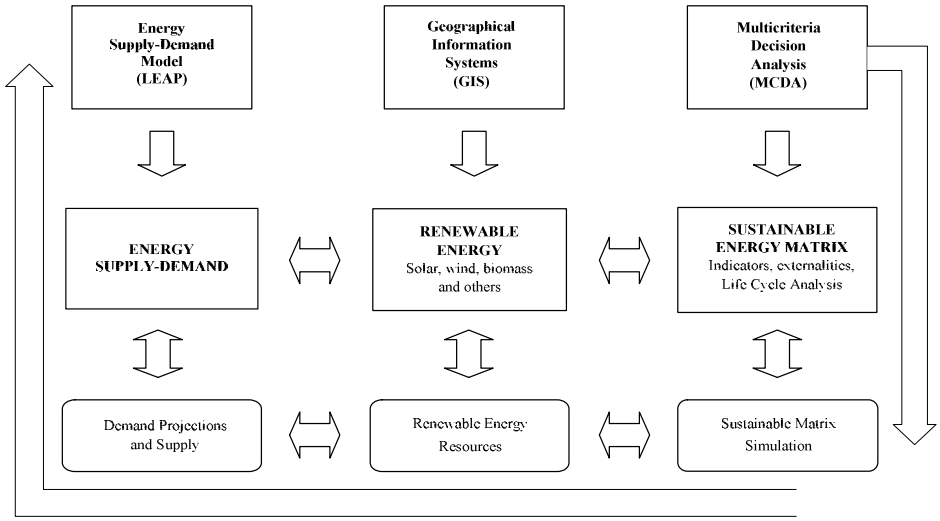


Figure 1. Flow chart of the designed model.

As we describe in the methodology section, the conception of the model is structured in three parts. These are: demand-supply models as LEAP, renewable energy integration based on GIS and sustainable matrix for LCA and externalities results. The entire model will be running until the feedback returns a sustainable energy mix, in order to define a sustainable energy supply. The model integrates some of the most important technologies in the evaluation of the energy system, environmental aspects and spatial analysis.

We will use the software LEAP for the supply-demand model, ARCGIS for GIS integration and, at the final step, software for the multi-criteria decision analysis (MCDA) in order to support the sustainable matrix.

The sustainability matrix will be defined according to the indicators included in the Energy Indicators for Sustainable Development (EISD) [IAEA/IEA, 2001] across to dimensions social, economic, technical and environmental. These four dimensions are further classified in themes and sub-themes. That some indicators can be classified in more than one dimension, theme or sub theme, given the numerous interlinkages among these categories. In addition, each indicator might represent a group of related indicators needed to assess a particular issue.

Table 2. Sustainability matrix example.

Theme/Sub-theme	**Energy indicator**	
General - Equity		
Accessibility		Share of households (or population) without electricity or commercial energy, or heavily dependent on non-commercial energy
Affordability		Share of household income spent on fuel and electricity
Disparities		Household energy use for each income group and corresponding

Theme/Sub-theme	**Energy indicator**	
		fuel mix
Overall use		Energy use per capita
Overall productivity		Energy use per unit of GDP
Supply efficiency		Efficiency of energy conversion and distribution
Production		Reserves-to-production ratio
End use	1.1	Industrial energy intensities
	1.2	Agricultural energy intensities
	1.3	Service/ commercial energy intensities
	1.4	Household energy intensities
	1.5	Transport energy intensities
Diversification (fuel mix)	2.1	Fuel shares in energy and electricity
	2.1	Non-carbon energy share in energy and electricity
	2.3	Renewable energy share in energy and electricity
Prices	3	End-use energy prices by fuel and by sector
Imports	4	Net energy import dependency
Strategic fuel stocks	5	Stocks of critical fuels per corresponding fuel consumption
Environmental Dimension	6	Other
Economic Dimension	7	Other

Theme/Sub-theme	**Energy indicator**	
Social Dimension	8	Other
Political and Institutional Dimension	9	Other

In our opinion, the first steps for the development of a solid project have been defined. In this sense, besides the aspects described previously, contacts with multiple agents potentially involved in the project have been made (energy corporations and governmental agencies in charge of necessary geographic information for the project).

6. CONCLUSION

According to the Colombian National Power Plan 2025, there should be available consistent and powerful tools for planning like the ones shown in this paper. These tools will ensure that the energy resources would meet the national request and would guarantee the sustainability of the energy sector in the long term.

The tool will provide enough elements for the formulation of coherent energy and environmental policies. These policies will be then based on a sustainable energy basket, which would not only meet society demands while improving its quality of live but also would promote the use of national resources minimising external energy dependence.

ACKNOWLEDGMENTS

International Cooperation Spanish Agency (AECI) has founded this proposal in the framework of PCI-INTERCAMPUS as a cooperative project between National University of Colombia and CIEMAT (Spain).

REFERENCES

Dominguez, J. and Amador, J. Geographical information systems applied in the field of renewable energy sources. *Computers & Industrial Engineering. A Cluster on Planning and Management of Energy and Infrastructure Engineering Projects*. *52 (3)*, 322-326, 2007.

Frangopoulos, C. A. and Caralis, Y. C. A method for taking into account environmental impacts in the economic evaluation of energy systems. *Energy Conversion and Management Efficiency, Cost, Optimization, Simulation and Environmental Aspects of Energy Systems*. *38 (15-17)*, 1751-1763, 1997.

IAEA. Energy indicators for sustainable development: guidelines and methodologies. 2005.

IAEA./IEA Energy indicators for sustainable development: methodologies. 2001.

IDEAM. Primera Comunicación Nacional al UNFCCC. 2001.

Lago, C., Cabal, H., Sáez, R. and Montserrat, F. Análisis del Ciclo de vida de Combustibles alternativos para el Transporte. 2007.

Smith, R., Mesa, O., Dyner, I., Jaramillo, P., Poveda, G. and Valencia, D. *Decisiones con múltiples objetivos e incertidumbre*, Universidad Nacional de Colombia. Facultad de Minas., Medellín (Colombia), 2000.

UN_DESA. Indicators of sustainable development: guidelines and methodologies. 2001.
UNFCCC. United Nations Framework Convention on Climate Change. 2005.
UPME. National Power Plan 2006 - 2025. 2006.
Vera, I. and Langlois, L. Energy indicators for sustainable development. *Energy. Third Dubrovnik Conference on Sustainable Development of Energy, Water and Environment Systems*. *32 (6)*, 875-882, 2007.
WORLD-BANK. Assessment of Personal Computer Models for Energy Planning in Developing Countries. 1991.
WORLD-BANK. Environmental Assessment Electrical power Colombia. 1992.
WORLD-BANK. World Development Indicators (WDI). 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

GIS modelling of forest wood residues potential for energy use based on forest inventory data: Methodological approach and case study application

L. Panichelli and E. Gnansounou

Laboratory of Energy Systems- Swiss Federal Institute of Technology, LASEN-ICARE-ENAC, Station 18, EPFL, CH-1015 Lausanne, Switzerland (luis.panichelli@epfl.ch)

Abstract: This paper presents an approach to perform geo-referenced estimations of forest wood residues availability for energy use based on forest inventory data integration into a GIS. Three different estimation methods are described. The first one evaluates biomass availability based on the application of biomass expansion factors to stem volume data of the forest inventories. The method accounts for forest dynamics and assigns management treatments in function of forest properties. The second method estimates available forest wood residues applying biomass production by tree, derived from field studies, to the inventoried tree species. The third method links inventory data with national statistics of final cuttings of commercial tree species. Useful biomass potential is then estimated based on ecological, logistic and economic constraints. The methods were tested in a case study in Northern Spain where optimal facilities location based on marginal delivery costs and resources competition between facilities were found. Results are presented for three different scenarios. Biomass resources estimations under the different methods result in significant differences. GIS maps of useful biomass availability estimations are presented giving an idea of the optimal locations for bioenergy facilities based on resource availability.

Keywords: GIS; forest wood residues, biomass, Northern Spain, forest inventories.

1. INTRODUCTION

Estimations of aboveground biomass have gained importance in recent year as they are used to evaluate the carbon content in forests as a need for Greenhouse Gases Emissions Inventories in order to achieve the Kyoto Protocol target in committed countries (Brown, 2002). As present forest inventories only account for stem volume wood, total biomass estimations have to be made in order to consider the whole tree carbon content (Chhabra *et al.*, 2002). These methods are based on extrapolations of forest inventory data and are used to estimate the non-commercial available biomass, as this information is not often given at present in forest inventories and they are the potential source to supply bioenergy facilities.

Biomass-to-energy projects are highly geographically dependent (Noon and Michael, 1996). Estimations of feedstock availability allow to determine bioenergy facilities optimal location and size, and to estimate biomass supply costs from forest parcels to the power plant (Voivontas, 2001). Consequently, inventory data needs to have a geographical support in order to estimate biomass delivery costs and potential facilities location.

Even though different biomass estimation methods exists, their applications to bioenergy systems rely generally on the application of one method and constraints for resources use as

feedstock for an energy facility are not always included. In this paper we present an integrated approach that combines three biomass estimation methods including limitations of the resources availability to calculate the useful biomass potential.

Depending on data availability, the structure of the forest inventory data and the purpose of the estimation, different methods can be applied. Methods based on allometric equations are not included because it was difficult to integrate individual tree data into the GIS environment. Consequently, proposed methods are based on the use of biomass expansion factors (BEF) and residues production ratios applied to inventory data at the parcel level. Estimations based on wood stem volume assume that the logs will be cut and used for other purposes while the remaining biomass (branches and leaves) will be used as feedstock for a bioenergy facility.

2. GENERAL OVERVIEW

The main characteristics of each method are presented in Table 1. The first method (M1) estimates biomass availability based on the application of biomass expansion factors to stem volume data of the forest inventories. The method accounts for forest dynamics and assigns forest management treatments (FT), such as thinnings and final cuttings, in function of forest development stage (FDS), rotation period and tree species. The second method (M2) estimates available forest wood residues applying residues production data by tree, derived from field studies for each forest treatment, to the inventoried tree species present in the forest parcel. The third method (M3) links inventory data with national statistics of final cuttings of commercial tree species by applying BEF to the commercial stem volume of exploited tree species in the parcel.

Table 1. Comparison of biomass estimation methods.

	Classification			
Method	Input data source	Input variables	Estimation	Growth
M1	NFI	VCC (T) + IAVC (T)	BEF	Dynamic
M2	NFI	VCC (T), Na	BUP	Static
M3	NFI + NFS	VCC (R)	BEF	Semi-dynamic

NFI: National Forest Inventory, NFS: National Forest Statistics, VCC (T): NFI wood stem volume in the parcel, IAVC (T): NFI wood stem volume annual increment, VCC (R): NFS wood stem volume from commercially exploited species in the parcel, Na: Number of trees, BEF. Biomass expansion factors, BUP: Biomass unitary production derived from regional field studies, Dynamic: forest annual growth is considered and FDS is projected to the life time of the energy facility, Static: Estimations are based on present FDS, Semi-Dynamic: Forest growth is indirectly considered in NFS wood stem volume.

Useful biomass potential is then estimated based on ecological, technical and economic constraints. The first ones account for soil protection and biodiversity preservation. The second ones reflect biomass losses during logistic operations, forest accessibility, and recovery rates. Finally, economic limitations consider resources competition and production costs.

3. THEORETICAL BIOMASS POTENTIAL ESTIMATION

The theoretical potential is the quantity of biomass from forest management activities (expressed in $ton.yr^{-1}$ or $MWh.yr^{-1}$) that a region can produce in a predefined period of time. This potential represents the quantity of forest wood residues (FWR) generated in a region, and can be considered as the upper bound of the bioenergy feedstock that can be obtained from forests in the area. This potential is a function of the forest type, the main tree species present in the forest, the exploited tree species, and the forest treatments applied to them.

Forest type depends on the fraction of the parcel covered by forest. This data is provided in the NFI at the parcel level (1 km^2) and classifies forests from dense to disperse woodlands. Disperse forests do not generate much residues and its collection can be costly, so biomass

estimations are recommended to be done from dense forest types. Tree species differ in geographical location, rotation period, growth, forest management practices, aerial biomass, heating value and packing density. Location is given in the NFI. However, the other data has to be linked and estimated from other sources. Some forest inventories contain data on forest management practices that can also be linked. On the other hand, this information has to be validated with current practices in the study region.

Theoretical biomass potential is obtained as the sum of the biomass quantities present in each parcel for the whole region.

M1. Assigning forest management based on forest dynamics

For each FDS (regeneration, sapling, pole, and timber) an average age class is assigned based on the rotation period of each selected tree species in the parcel. Then, the plantation is projected for the next 20 years (the life time of the bioenergy facility) and FT are assigned in function of tree age. As a consequence of forest growth, FDS will change over time. First thinnings are applied to sapling parcels, commercial thinnings to pole and timber parcels and final cuttings to timber parcels. No treatment is applied to regeneration parcels and only one FT is assumed for each FDS. Each FT is applied to a certain percentage of the parcel according to current practices. Available biomass is then projected till the moment when the FT is applied by multiplying the steam volume and its annual increment by their BEFs. Specific BEFs are applied to calculate the useful fractions (all aboveground biomass, branches, and leaves). Forest residues production for each tree species by parcel is estimated as the sum of the biomass generated in each FT. The annual harvestable biomass ($ton.yr^{-1}$) at each forest location for a 20 years period is obtained (Solla-Gullón, 2006).

Table 2 shows an example for *Pinus radiata*. In this case, (rotation period: 32 years) the present regeneration surface (average age: 2.5 years) will develop until a timber stage of 22.5 years. That is to say, theoretically, this surface will undergo a first thinning at 8.5 years that will affect 40% of the surface and a commercial thinning at 16 years that will affect 37% of the surface. The resulting biomass amount in the first thinning will be the result of multiplying 0.40 by the projected steam volume at that age and by the respective BEFs. In the same way, biomass amounts would be estimated for the rest of the FDS and tree species.

Table 2. Example for *Pinus radiata*

FDS	Age Range		Mark Age	Age Evolution	Mass Status							Forest Treatment							Applied Surface					
					2.5	8.5	16	26	32	2.5	8.5	2.5	8.5	16	26	32	2.5	8.5	8.5	16	26	32	2.5	8.5
R	0	5	2.5	22.5	R	S	P					NT	FT	CT					0.4	0.37				
S	5	12	8.5	28.5		S	P	T					FT	CT	CT				0.4	0.37	0.35			
P	12	20	16	32	4		P		T					CT		FC				0.37		1		
T	20	32	26	32	14				T	R	P					FC	NT	FT				1	0	0.4

M2. Field studies data

Field studies allow validating modelled estimations with current values in the study region. These values are linked to the inventoried number of trees present in each parcel for each FDS, obtaining the biomass availability by tree species in each forest location.

The biomass obtained in each parcel i for all FT and all tree species a is calculated as: $BT_i = \sum_{a=1}^{A} \sum_{FT=1}^{T} BT_{i,aFT} . S_{aFT} . N_a$, where, $BT_{i,aFT}$ is the biomass quantity generated per tree, per FT and per tree species in parcel ($ton.tree^{-1}$), S_{aFT} is the surface of the parcel where the FT is applied (%) and N_a is the number of stands of that species in the parcel. In order to consider forest dynamics Na should be projected for the life time of the energy facility.

M3. Final cuttings of selected tree species

Final cuttings produce most of the wood residues generated by a forest plantation during its rotation period (Kallio and Leinonen, 2005). Moreover, these residues represents the most economically and technically biomass fraction to be used in a bioenergy system. As biomass availability is dependent on forest commercial exploitation, available amounts of forest residues are subject to final cuttings of commercial wood. Scenarios derived from this methodology are assumed to be the most realistic as they take into account the real quantity of biomass that is produced annually in the region. On the contrary, only biomass from final cuttings can be estimated as it is difficult to obtain statistics of biomass volumes generated in other FT from exploited species.

The biomass estimation method is based on the link between geo-referenced inventory data and national statistics of commercial cuttings. The average volume of commercial wood (VFC) annually cut from each tree species in each province is obtained and assigned to the forest parcels depending on FDS. For short rotation period species all FDS are considered, as it is assumed what those parcels are the ones to be exploited in the next 20 years. For medium rotation period species pole and timber stages are considered and for long rotation period species only the timber stage is considered. VFC for each tree species and province is divided by the number of forest parcels containing the selected FDS, and assigned to the geo-referenced points where those combinations (species-FDS) are present. Then, BEFs are applied to convert VFC of each parcel and tree species into annual biomass quantities (d.m. $ton.yr^{-1}$). No biomass increments are considered as VFC already represents the real annual wood volume available in each forest location.

4. USEFUL BIOMASS POTENTIAL

Useful biomass potential is defined as the biomass quantity ($tons.yr^{-1}$ or $MWh.yr^{-1}$) that can be used for energy after applying constraints. These limitations accounts for protected areas, land ownership, terrain slops (accessibility), soil protection, recovery rates, biomass use competition, and operational costs.

Environmental sustainability is accounted for by leaving aside protected areas from the available forest resources and by leaving a fraction of forest residues in the parcel for soil protection. Public and private forest management may introduce limitations for biomass extraction. Some parcels may not be accessible or operational due to terrain slope. Different production methods and related costs are linked to terrain slopes. Moreover, not all harvested biomass can be collected. Recovery rates of logging residues depend on the logging method, the harvesting season and if the residues are harvested fresh or dried. Competition for the biomass resource should be assessed in the region for the life time of the energy facility, considering feedstock for other uses and for other energy facilities. The availability reduction will depend on biomass price and geo-political decisions. Operation costs accounts for the biomass collection, haulage and packing. Considering the collected biomass fraction, the terrain slope and the forest management strategy different production costs are obtained

The constraints can be applied by GIS overprint of layer or by calculations at each parcel of the NFI.

5. CASE STUDY

5.1. Description

In the framework of the BIOFOREST project- Waste wood to energy in Northern Spain-, energy facilities were planned to be installed in the Basque Country and its neighbour provinces. The covered area is 41’201 km^2 from which 15’024 km^2 (36%) are dense-wooded forests. The purpose was to recover 84 d.m. $kton.y^{-1}$ of forest wood residues from commercial forest management to produce torrefied wood and supply a 22 MWe gasification unit. The objective of the study was to define the optimal location for bioenergy facilities based on biomass availability and delivery costs. The overall approach

and costs data are detailed in Panichelli and Gnansounou (2007). For this paper, we focus on the estimation of the biomass useful potential.
Three scenarios were modelled and biomass potential was estimated using the three methods described. Scenario E1 estimates final cuttings' biomass of short rotation species, scenario E2 assess final cuttings' biomass of exploited species with technological improvements (bailing system with chipping at power plant) and E3 estimates first and commercial thinnings and final cuttings biomass of short and long rotation period species. No biomass estimations for E3 under method M3 were made, due to lack of data.

5.2. Database development

The software used was the ArcGIS® version 9.1. Supporting software such as Microsoft Access® and Microsoft Excel® was also included for data processing and modelling. The source data was the CAD format National Cartographic Database 1:200'000 produced by the National Geographic Institute. This database is available for each province under study. Forest accessibility was assessed based on the Digital Terrain Model with a 1:200'000 resolution from which terrain slopes were calculated.

Forest inventories were obtained from the Biodiversity Database of the Spanish Environmental Ministry. The decision to work with the most present data has lead to use different forest inventories. However, due to their different structure it was not possible to create a geo-referenced database relating tree species with diametric classes and forest location for the whole region. For these reasons, the BEF method was applied to calculate total harvestable biomass. Data from the Third National Forestry Inventory (NFI3, 2006) was available for the regions of Navarra, La Rioja and Cantabria. For the other regions (Bizkaia, Alava, Gipuzkoa and Burgos) the Second National Forestry Inventory (NFI2, 1996) was used. Forest inventories have a 1: 50'000 resolution. In the first case geo-referenced information by forest parcel (points representing 1 km^2) includes the three more relevant species, with their respective forest type and ownership. A set of tables is also provided containing location of the sample points, number of trees, stem volume (under and over bark), stem volume annual growth, and complementary data. The NFI3 provides a large amount of information. But, georeferenced data is limited and covers the same variables as in the NFI2. Data is presented by parcel and by FDS.

Main tree species for the studied region have been selected in function of overbark stem volume (VCC) production ($m3.province^{-1}$). National forest statistics from the National Institute of Statistics between 1996 and 2003 were processed and an annual average value was obtained. Tree species VCC productions were increasingly ranged and summed up to reach 80% of the VCC production of the province (representing the main forest exploited species). Twelve tree species were selected from which biomass availability was estimated. Depending on the FDS of the tree species different biomass fractions have been considered.

A geo-referenced database integrating both NFI was created with the forest parcels that are wood dense forests and contain at least one of the twelve selected species.

5.3. M1. Theoretical forest management.

It was assumed that only one forest treatment is applied in each FDS. Selected treatments were: first thinning, commercial thinning, and final cuttings. Pruning was not included. Forest management data for each tree species was taken from the `Repoblación y Manejo Forestal´ Manual of the Gipuzkoa Government.

5.4. M2. Field studies data.

Field studies have been carried on in the Navarra region by the CENER in the framework of the Bio-South project (Techno-economical assessment of the production and use of biofuels for heating and cooling applications in South Europe). Biomass generated by forest treatment and by tree species was used to make estimations of the biomass theoretical availability under each scenario.

5.5. M3. Final cutting of commercial tree species.

Stem volume final cuttings' statistics were used to define biomass availability based on actual forest exploitation. The information was presented by province and by tree species. Statistics were taken from the Annual Spanish Statistics. For *Eucaliptus globulus* and *Populus nigra* all FDS were considered. For *Pinus radiata* pole and timber stages were taken into account. Only timber stage was considered for *Quercus* spp, *Fagus silvática* and other *Pinus* with long and medium rotation period.

5.6. Useful potential

Protected areas in the region (wetlands, National Parks and Biosphere Reserves) have been left aside from the availability of FWR. No constraints were found for land ownership and parcels with terrain slopes above 20% were considered but assigned a higher operational cost. 30% of total harvestable biomass at each parcel was left aside for soil protection. Recovery rates were assigned in function of the efficiency of forest residues production activities. Estimations were based on data for large scale production of wood chips from forest residues (Elsayed *et al.*, 2003). Total recovery rates considering soil protection and process efficiency are 61% for fresh residues and 46% for dried residues. No other bioenergy facility using forest residues exists in the area. Board industry and fire wood consumption compete for the forest residues in the region. For scenarios E1 and E2 competitors demand was assumed to be covered by the forest residues provided by the other FT. For scenario E3, 30% of the recovered biomass was left aside to supply competitors based on statistics of biomass destination (NFS) of coniferous and deciduous trees at provincial level. Results at the parcel level were aggregated at the municipal level for visualization purposes.

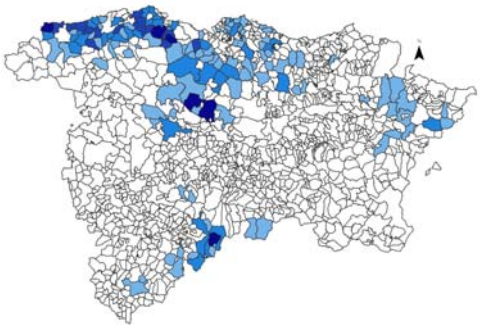
5.7. Results

Methods M1 and M2 have resulted in similar theoretical biomass estimations for scenarios E1 and E2. For E3, a significant difference exists due to biased input data for long rotation period species. Method M3 shows the theoretical potential, based on actual forest exploitation and represents the real biomass availability from final cuttings. At present, approximately one third of the theoretical final cuttings are done, giving a biomass potential of 75 - 100 d.m. ton.yr^{-1} of dried and fresh FWR, respectively, based on actual forest exploitation. For scenario E1useful potential of the studied region is 231 - 314 d.m. ton.yr^{-1} and 174-236 d.m. ton.yr^{-1} of fresh and dried FWR, respectively (Fig. 1).

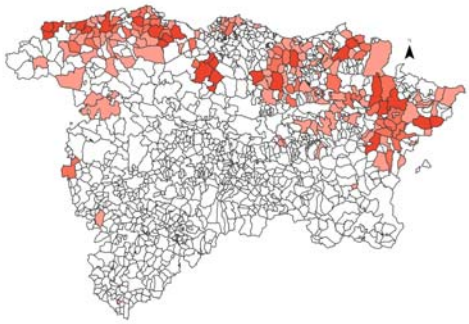
Geographical distribution of biomass resources for energy use is presented in figures 2a-f. Results show that the north-east part of the region presents the higher potential for siting an energy facility. This is due to the availability of short rotation period species that produce larger quantities of biomass during the life time of the plant.

5.8. Discussion

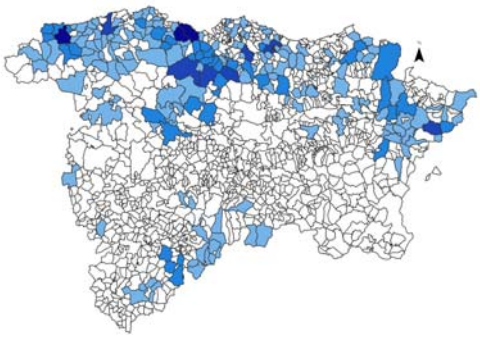
Biased results were obtained from each estimation method. Uncertainty in method M1 is linked to the use of national BEF that may not correspond to the regional values. However, more specific information was not available, and so, the BEF used to perform the National Greenhouse Gases Inventory were used. Forest dynamics was considered based on forest annual increment, as a practical way to account for forest development based on available data. Using a specific forest dynamic model may lead to more accurate estimations, but may also need specific skills and more detailed data, not always available. While M2 allows validating estimations based on regional data, field studies are costly and time consuming. For this case study, the number of samples was too low. This was a source of uncertainty as stands volume significantly vary in a forest. Moreover, projecting the number of trees in a parcel seems more difficult than projecting stem volumes, to account for forest dynamics. For this case study, estimations were done based on the present number of forest stands, and this has lead to introduce uncertainty in the estimation.



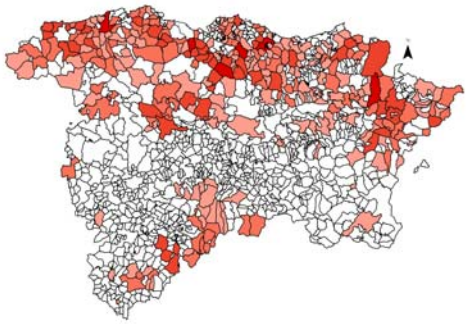
a) E2 M1dried FWR (dm ton.yr^{-1})



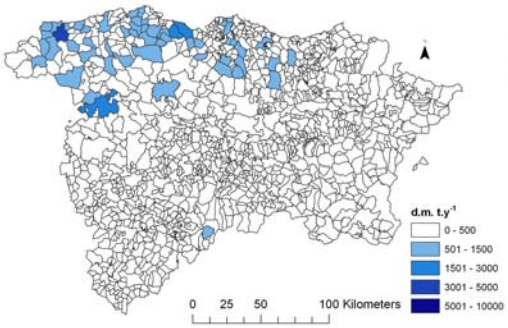
d) E1 M2 fresh FWR (MWh.yr^{-1})



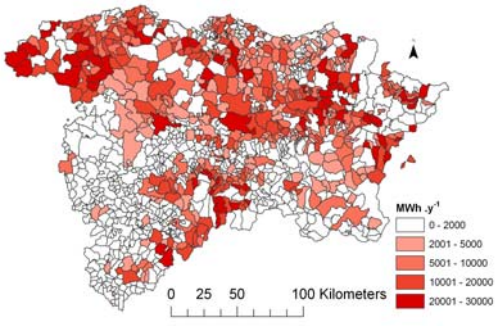
b) E2 M2 dried FWR (dm ton.yr^{-1})



e) E2 M2 fresh FWR (MWh.yr^{-1})



c) E2 M3 dried FWR (dm ton.yr^{-1})



f) E3 M2 fresh FWR (MWh.yr^{-1})

Figure 2. Geographical distribution of biomass resources for energy use. Figures a-c: Estimations for E2 under different methods for dried FWR. Figures d-f) Estimation under M2 for different scenarios for fresh FWR.

Estimations of forest residues for energy use are available from the Spanish Renewable Energy Plan (MITYC, 2005) for the whole region and for the province of Navarra (CENER, 2005). Even though the theoretical potential value is within our values, results are difficult to compare as different species and assumptions were considered. Significant biomass reductions are obtained when applying constraints, limiting the resources availability. These limitations are not considered in previous studies.

Method M3 allows to indirectly accounting for forest development, projecting statistic values of commercial final cuttings for the life time of the energy facility. For this case study, an average value of historical data was used as final cuttings' values by province and tree species do not significantly vary over time. Moreover, the estimations reflect the present forest exploitation as commercial statistics are used. On the other hand, geographical representation of forest resources is less accurate as estimations are not based on the stem volumes present in each parcel, but by distributing final cutting volumes of selected species between the forest parcels in the province with suitable FDS. This will be a

source of uncertainty when estimating operational costs. The same limitation as for method M1 applies concerning the use of BEFs from other region.

6. CONCLUSIONS

Biomass estimations under the different methods have conducted to significant different results in estimating biomass useful potential, giving evidence of the influence of the selected method on forest wood residues estimations for energy use. The use of commercial statistics that reflect the real forest management, the application of constraints on biomass use and the accounting of forest dynamics are essential to perform accurate estimations. These factors may significantly reduce the theoretical biomass potential. More efforts should be done to link forest inventory data with GIS, to develop regional BEF and forest dynamic models, to generate accurate data for FT others than final cuttings, in order to decrease uncertainty of the final results. While for policy and management strategies methods based on forest inventories stem volume are recommended, for practical and commercial applications methods based on commercial forest exploitation should be preferred.

7. REFERENCES

Brown, S., Measuring carbon in forests: current status and future challenges, *Environmental Pollution* 116, 363–372, 2002.

CENER, 2005. Bio-South: Techno-economical assessment of the production and use of biofuels for heating and cooling applications in South Europe. EIE/255/S07.38609. Proceedings from "II Jornadas Proyecto Bio-South", Pamplona, 24 Nov. 2005.

Chhabra, A., Palria, S., Dadhwal, V.K., Growing stock-based forest biomass estimate for India, *Biomass and Bioenergy* 22, 187 – 194, 2002.

Elsayed, M.A., R. Matthews, N.D. Mortimer, Carbon and energy balance for a range of biofuels options. Sheffield Hallam University Report B/B6/00784/REP, 2003.

MITC (Ministerio de Industria, Turismo y Comercio), Plan de energías renovables en España 2005-2010, 2005.

Noon, C.E., Michael, J.D., GIS-based biomass resource assessment with BRAVO. *Biomass and Bioenergy* 10, 101-109, 1996.

Panichelli, L., Gnansounou, E., GIS-based approach for defining bioenergy facilities location: A case study in Northern Spain based on marginal delivery costs and resources competition between facilities, *Biomass and Bioenergy*, (in press), 2007.

Sabaté, S., Gracia, C., Vayreda, J., Ibáñez, J., Differences among species in aboveground biomass expansion factors in Mediterranean forests. Working paper, Center for Ecological Research and Forestry Applications (CREAF), Universitat Autònoma de Barcelona, Spain, 2006.

Solla-Gullón, F., 2006. Unidad de Gestión Forestal Sostenible, Escuela Politécnica Superior de Lugo, Universidad de Santiago de Compostela. Personal communication.

Voivontas, D., Assimacopoulos, D., Koukios, E.G., Assessment of biomass potential for power production: A GIS based method, *Biomass and Bioenergy* 20, 101-112, 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

Research activities in renewable energy sources integration with GIS at CIEMAT

J. Domínguez[a] , I. Pinedo[a] and J.A. Gonzalez[b]

[a] Renewable Energies Division. Energy Department, CIEMAT. Avd. Complutense, 22, 28040 Madrid. Spain (Javier.dominguez@ciemat.es)
[b] National University of Tucuman (Argentina)

Abstract: Since 1994, CIEMAT has been involved in several research activities regarding the integration of renewable energy sources (RES) as well as resource assessment based on Geographic Information Technologies. One of the first projects was SOLARGIS which aimed to demonstrate the viability of applying GIS to the RES integration in rural electrification programs. This research line has been improved since then, through the permanent actualization and update of the methodology. In this sense, the collaboration with UPM (Polytechnic University of Madrid) has been decisive. Nowadays, CIEMAT GIS team is applying this methodology to Latin America countries like Cuba or Colombia, in the general framework of Renewable Energy Integration GIS Project (INTI-GIS). In addition to GIS application for rural electrification research line, our team has also accomplished resource assessments and siting studies for distributed electrical generation based on RES. On this topic, CIEMAT has developed several projects like MERSOTERM, SOLBIO and others related to biomass or wind resource assessment. Finally, in order to promote RES, our team has collaborated in the accomplishment of projects aimed to define a feasible technology mix based on renewables, like SIGER project, or Decision Support Systems (DSS) supporting the decision-making for a sustainable development based on high penetration of RES in Latin-American countries like Colombia. According to our experience, we strongly believe that GIS technology constitutes a powerful tool in order to promote and integrate renewable energy sources in the path of a new energy international model more respectful with the environment and also with the sustainable development of the society.

Keywords: GIS; rural electrification; renewable energy; decentralized generation.

1. INTRODUCTION

Renewable energy sources (RES) are characterised by its strong temporal and spatial variability, in contrast with the distribution of the so-called fossil fuels. This fact makes it possible to find at least one local source of green energy in almost every location. This advantage of the RES, compared to the conventional sources, has also the disadvantage of making the harvesting system a bit more complex than the one using conventional fuels characterised by the specific distribution of supply and demand, with a concentrated generation model and huge consumption points situated far away from resources and power generation.

The typical heterogeneous spatial distribution of RES makes them appropriate and near to demand sources of energy in terms of electricity production applied to rural electrification or distributed generation.

The variability and complexity of supply and demand system based on RES makes powerful tools like GIS very adequate. There are numerous references in the past and current bibliography that defend and prove this fact.

The first approach to the state of the art related to those issues makes us think in a three group classification: decision support systems (DSS), distributed electricity generation based on RES and decentralised applications in rural electrification projects [Dominguez, J*., et al.*, 2007]. The first one, DSS, relays on the implementation of a regional database that supports the energy planning. The distributed generation systems (DG) focus on the definition of the best sites depending, among others, on the resource availability and distance to the grid. Finally, the rural electrification projects concentrate on the localisation and characterisation of the potential demand becoming therefore a bit more complex from a methodological point of view.

Although it covers all the possible research cases, this classification could seem artificial since most of the projects combine aspects from two or three of the formerly mentioned groups. To sum up, it should be emphasised the diversity found in applications and technologies as well as the world wide research studies located along the five continents. In Spain, there are leading groups thanks to its international applications. It is highly advisable to support the international collaboration by means of the creation of research network and technology transfer specialised on this issues.

Next, we will describe several research activities of CIEMAT GIS team in the area of rural electrification, distributed generation and support to RES promotion.

2. GIS AND RURAL ELECTRIFICATION. SOLARGIS MODEL METHODOLOGY AND EVOLUTION

One of the first projects where CIEMAT used Geographical Information Systems (GIS) as the main analysis tool was SOLARGIS [Solargis-Team, 1996]. The mayor goal of SOLARGIS project, titled *Integration of renewable energies for decentralized electricity production in regions of EEC and developing countries,* was the demonstration of GIS viability in contributing to the RES integration in rural electrification programs.

This research line has been improved since then through the permanent actualization and revision of the methodology and software (developed originally with ARCINFO^{TM}). The collaboration established with UPM (Polytechnic University of Madrid) stood the framework for a significant improvement. [Amador, J*., et al.*, 2005] The objective was to minimise and/or control the uncertainties through the in depth study of the technical and economic parameters involved, focusing on load treatment, results control and interface design. Special mention needs to be paid to the spatial sensitivity analysis implemented to check the stability of the results [Amador, J*., et al.*, 2006].

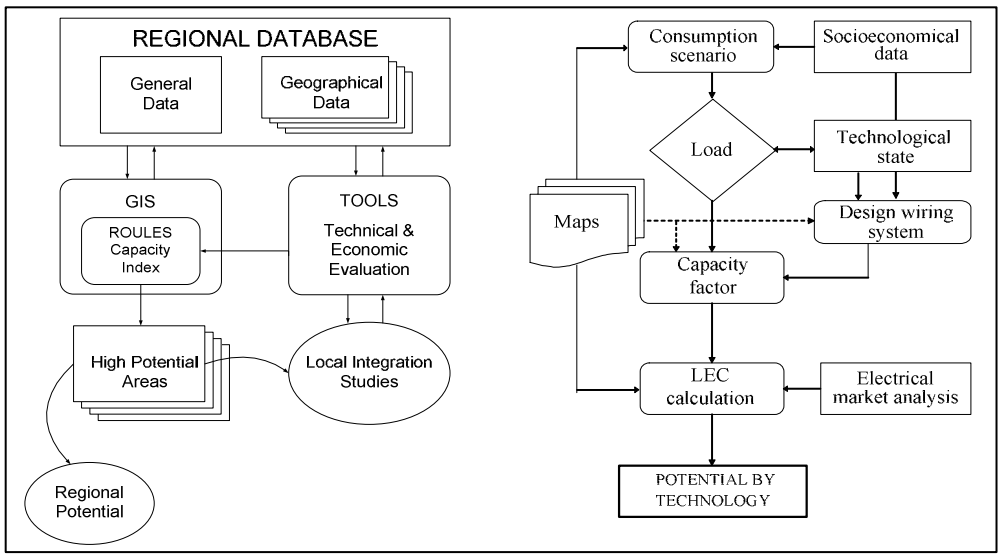


Figure 1. SOLARGIS model evolution.

Nowadays, CIEMAT GIS team keeps on working in the development of SOLARGIS concepts in the framework of INTI-GIS Project. This conception includes new advances in methodological questions as well as the incorporation of new renewable technologies and the application to new areas and, specially, to Latin America countries like Cuba [Pinedo Pascua, I., *et al.*, 2007] or Colombia and working hardly in the adaptation of the software to the new ARCGIS™ concept.

As it can be noticed in Figure 1, the SOLARGIS model has changed drastically since the first version developed within JOULE frame to the version implemented by CIEMAT-UPM collaboration. In the first case, some constrains were imposed to the regional data base using GIS. These constrains were mainly based on the capacity index, which was defined on the basis of a technical and environmental assessment. Using the tools available within the GIS environment and according to the established rules, High Potential Areas were highlighted. Later and focusing on these areas, both the regional potential and integration studies were performed at a local scale. The local scale studies were not elaborated within JOULE program but were developed by the different participants.

Regarding the CIEMAT-UPM SOLARGIS model (Figure 1, right), the demand characterisation becomes the fundamental axis and the methodology is fully implemented in a GIS environment. The output of the model is the definition of the most competitive technology in each location.

This new conception of the model was applied to Lorca municipality (Murcia, Spain), in which the number of isolated houses from national grid were numerous. There were also previous assessment studies about wind and solar resources made by CIEMAT.

Besides the definition of a new methodological structure, the result of the project (Figure 2) showed the spatial distribution of the most competitive technologies (PV, wind, connection to existing grid, individual and central diesel).

One of the most relevant aspects of this project was the development of a spatial sensitivity analysis methodology based on three stages. The first one included the determination of the main influencing parameters by a sensitivity analysis of each technology LEC values. The second one was the spatial sensitivity analysis of the potential of the studied area in respect to the parameters found in the previous stage, selecting the most significant variables in the distribution of the rural electrification potential. Finally, the third one was the study of the spatial behaviour of the variables of the previous stage in order to determine the ‘‘stability’’ of the obtained result.

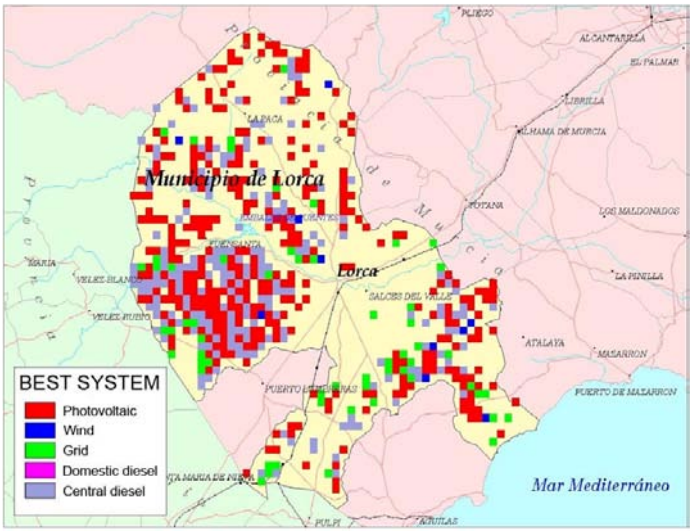


Figure 2. Application to Lorca. Best System.

Although the improvements in the methodology have been really important since the first version, we would like to point out some aspects that could help in updating it. First of all, the method needs to adopt the new sources of geographical information, making the application accessible and useful for a wider and wider range of potential users. This would involve not only using some of the data sources available through Internet but also the development of an application based on those principles.

Secondly, the application should be more dynamic and easily adaptable to other geographical environments. In order to reach this goal, we are working on the

simplification of the required parameters and also trying to provide the application with more dynamism, in terms of including new technologies by means of making the application to be able to read and incorporate other applications output like HOMER or RETScreen.

Finally, we would like to conclude this section presenting the recent application of our model to a rural electrification study in a Cuban municipality. The aim of this study was to define which technology is the most appropriate to be adopted by particular group of un-electrified communities, consistently with the main objectives of SOLARGIS methodology, but also to adjust the whole process to the particularities of the Cuban context.

In this study, several technology systems capable of meeting the existing demand were compared at regional scale: individual systems to electrify single households (PV and diesel) and central systems to electrify communities (diesel and connection to the existing grid). Not only economic data is taken into account. Distance to the existing grid, electricity demand, solar radiation, and technical parameters are also required by the application. In order to assess the distance, information from GTOPO30 DEM and MV digitalized network were used. The slope was calculated and used as a cost layer in order to approximate the results to real distance. The rural load per household was defined in earlier studies [Shiota, A., 2002] on the basis of interviews to the inhabitants of already electrified communities. Considering number of appliances, power consumption and ratio of use, the average consumption per household was obtained. The global annual solar irradiation data was obtained from SWERA project, freely available from their webpage [SWERA-Project]. These data provide monthly and annual average daily total solar resource expressed in terms of kW/m^2 per day for each month with a resolution of 10 by 10 km on a flat surface (see Figure 3). The solar data in tilted surface was calculated through a routine implemented in MATLAB.

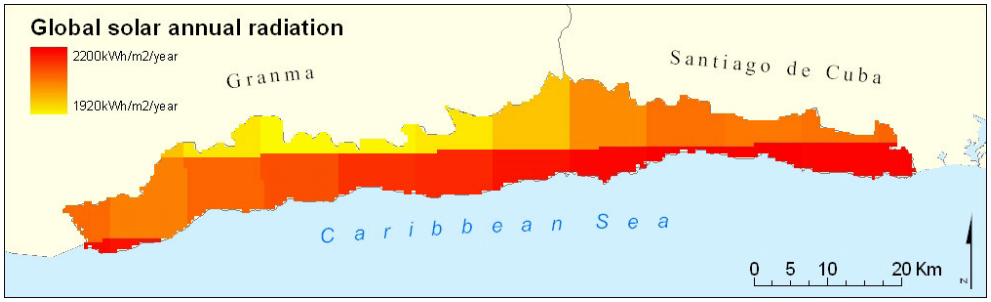


Figure 3. Solar energy values based on SWERA project.

The values obtained show a very high potential of remote sites to be electrified by means of photovoltaic systems. In more than 90% of the considered communities individual photovoltaic systems represents the most competitive way of electrifying the households.

The next steps of our GIS team at CIEMAT are related to the current project we are developing in Colombia, as well as updating the application to the new version of ARCGIS 9.2 and working on a user manual.

3. GIS AND ELECTRICITY DISTRIBUTED GENERATION WITH RENEWABLE ENERGIES

In addition to GIS application for rural electrification research line, CIEMAT GIS team has also accomplished resource assessments and siting studies for distributed electricity generation based on RES and also several projects analysing the possibilities of a high penetration RES technology mix in Spain (SIGER) or Decision Support Systems (DSS) assisting the decision-making for a sustainable development in Latin-American countries like Colombia. This section focuses on this type of studies, presenting firstly MERSOTERM and SOLBIO [Domínguez, J., *et al.*, 2000], as examples of project for distributed generation with renewables, and later the SIGER project [Dominguez Bravo, J., *et al.*, 2007], which focuses on the definition of a feasible technology mix based on

renewables. The DSS for a sustainable development in Colombia is described in another contribution made to this congress.

The MERSOTERM project, "Solar Thermal Electricity generation market with Tower technology. Application to Spain", was commissioned by INABENSA. The study was based on the most recent viability studies about Concentrating Solar Power (CSP) Tower technology deployment in several sites, analyzing the different market possibilities that could arise and the areas with higher potential. It also included a GIS analysis of Spanish case with a proposed solar power tower plant design. The project aimed firstly to spread information about this technology and its potential by solar resource availability and development, which finds ideal conditions in Spanish territory, and secondly to support the deployment of a demonstrative CSP tower plant which would prove the viability of this technology and would became the starting point of the explosion of the CSP market in the solar-belt countries [CIEMAT, 1997].

In order to evaluate the geographical potential in Spanish sites and express it in terms of generated electricity cost, five variables were used (Table 1). These variables were related to infrastructures and also to all the resources required by the technology.

Table 1. MERSOTERM, original variables.

Variable	Value	Analysis	Information Source
Distance to roads	< 10 km	Buffer	IGN-BCN 1000
Distance to water stream	< 10 km	Buffer	IGN-BCN 1000
Distance to Transport Electric Network	< 10 km	Buffer	REE- Peninsular Electric System Map
Slopes	< 4%	MDT→ SLOPE →RECLASIFY→ RASTER to VECTOR	IGN-MDT 1000
Generation cost by kWh	< 18c€/kWh	HELIOSAT	METEOSAT

The obtained results showed two appropriate localization areas: Southwest of Spain (with lower costs along Guadalquivir Valley) and Southeast, characterised by less homogenous spatial continuity although some individual sites obtained more potential than in the previous case (higher radiation values lead to cheaper costs per kWh). Nevertheless it should be mentioned that the precision gave by the working scale it is not enough to define the local production accurately and therefore, the data on the following maps should just be understood as zone references.

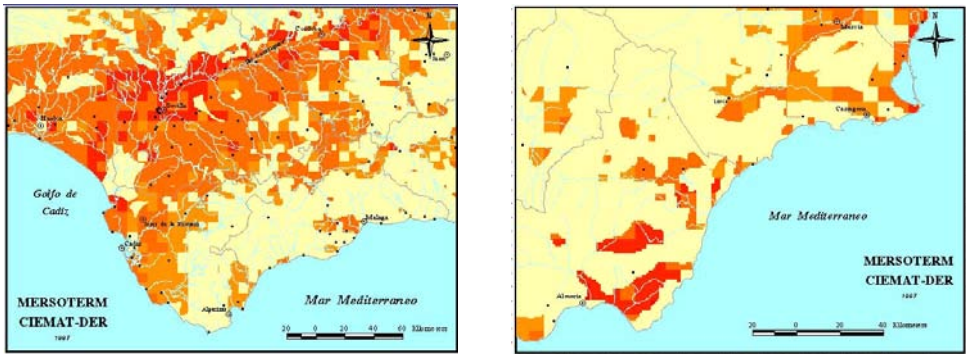


Figure 4. MERSOTERM: first potential areas for Solar Tower technology in Spain.

Based on the conclusions obtained in MERSOTERM project, SOLBIO project aimed to assess the possibilities of installing a solar/biomass Hybrid CSP adding a 300,000 MWh/year biomass plant to the previously proposed CSP design. The analysis extend was reduced to Andalusia autonomous community since the best sites for CSP deployment were mainly found there.

The main source for the analysis was the data provided by the Institute of Statistics of Andalusia about "Agrarian surface according to its land use", at municipal scale and

expressed in Hectares. All the areas dedicated to grass crops, olive grove, vineyard, fruit trees and forestry were included in the study.

The methodology had the following steps:

1. Geographic Database creation, including surface data per municipality and land use.
2. Biomass waste evaluation according to biomass type.
3. Energy power calculation at municipal scale.
4. Neighbourhood analysis to evaluate the areas with higher biomass energy use potential.

The SOLBIO Project results (Figure 5) show that the areas with higher productivity potential are located in the Guadalquivir Valley and in the Betic System, although the region as a whole enjoys important presence of local resources.

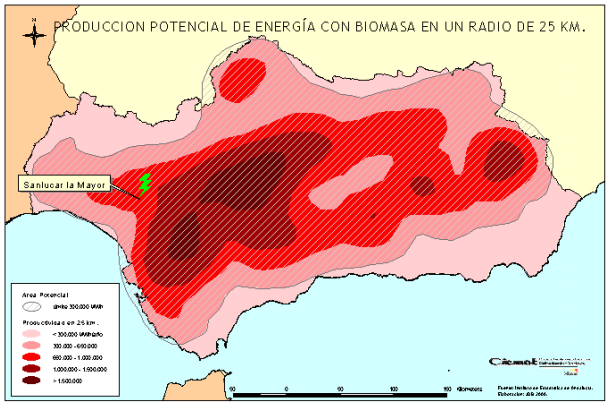


Figure 5. Biomass Potential Generation and available area in Andalucía.

A zonal analysis, constrained by the availability of biomass within 25km from the plant site and a power generation ceiling of 300,000 MWh/year, proves that most of the studied territory is viable for the installation of Hybrid solar-biomass in terms of resource availability.

In order to optimise and update the results of this project, it would be necessary to revise cultivation area surface and to check the uncertainty associated to the used data sources.

Another project with special relevance for our work was SIGER, "Technical analysis of the introduction of high percentage of renewable energies in Spanish electrical generation system". Greenpeace commissioned Technology Research Institute at the Pontificia Comillas University to carry out a study to assess ceilings for the potential and generation of renewable technologies in Spain. CIEMAT assumed the responsibility of accomplishing the required spatial analysis, using GIS as the main tool, which linked to the technical analysis the specific constrictions imposed by territory (natural and anthropogenic) and not just designed to cover a certain demand. Therefore, GIS spatial analysis took into account local conditions producing a more accurate assessment than evaluations made upon ''virtual'' electrical spaces [García Casals, X., *et al.*, 2007].

A scenario for 2050 in Spanish peninsula was taken into account in order to set the energy demand and give some percentages of the demand that could be covered by renewable energy technologies. It should be remarked that a conservative approach was taken in terms of technology, which implies that technological improvements would lead to even higher generation ceilings than the estimations presented.

GIS was used to add the spatial dimension to the analysis. GIS tools were used firstly, to analyse spatial distribution of renewable energy resources; secondly, to impose spatial restrictions related to land suitability (land use, slope, potential productivity, etc.); and thirdly, to calculate capacity and generation ceilings based on available sites.

The results were both cartographical (see Figure 6) and numerical, used to estimate capacity and generation ceilings of renewable energy technologies according to prospected Spanish conditions in 2050. The estimated ceilings show far greater renewable potential than the targets set by long term policies. The approach could be applied in other countries or regions to assess the maximum contribution that RES could have in a more sustainable generation mix. This type of studies would help not only supporters of increasing renewable energies participation in national generation mixes, but also governments to set development targets more adapted to renewables potential. In the case of Spanish

peninsular, a 100% renewable generation mix would be viable to meet total energy demands projected for 2050.

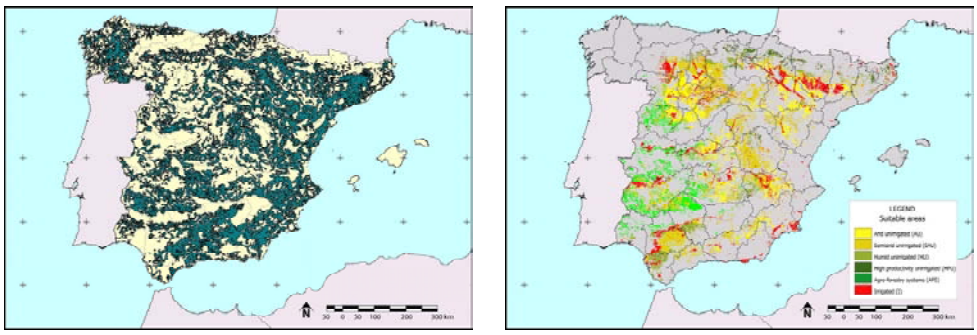


Figure 6. Suitable areas for CSP, energy crops and wind farms over complex terrain.

Next, we present the evolution lines of our research topic and scheduled studies in short term and for the near future.

4. EVOLUTION LINES

The evolution of this investigation field can be analyzed considering three aspects: the evolution of GIS, the development of renewable energy technologies and the integration of models and systems.

Considering the state of the art, we can guess that the evolution of the GIS application to the renewable energy integration will be oriented to a bigger integration and specialization in technologies as much as in data sources. Also, the combination of these aspects with the technological evolution of the GIS implies, in our opinion, that there will be an increasing interest on transferring the knowledge and results of this type of projects to the society (like SWERA or PV-GIS). In order to succeed, the application of space data infrastructures (SDI) and the use of the data obtained from remote sensors are immediate requisites that mark clearly the evolution from the perspective of the innovation of the geographical information technologies (GIT).

Regarding the integration of technologies for the use of RES, the challenge is related firstly to the widespread consideration of the aspects not directly linked to the electricity generation process; secondly to the integration of multiple sources and the design of an energy mix; thirdly to the characterization of new systems and, lastly, to the RES evaluation and prediction based on multiple sources of information. Besides, it should not be forgotten the current improvements of the PV, CSP, offshore wind and biofuels technologies, associated not only to the resources but also to the transformation and distribution processes.

The integration of models and technologies based on the use of GIS as the main tool, allows to combine several elements related to the energy production process such as life cycle analyses, environmental impacts assessments, prediction and impact on the grid, etc… supporting more complex and complete analysis integrating offer-demand scenarios, technological and environmental aspects.

To sum up, the evolution of this multidisciplinary topic will be influenced by different aspects such as: the diffusion of results through SDIs; the resource assessment based on remote sensing technology; the integration of different sources of GIS data; the development of innovative tools from functional and methodological perspectives; the increasing viability of transferring the generated knowledge to the industry and decision agents; the collaboration and research net creation; the integration of RES in applications like heat-cold production, transport and water desalination; etc. There is no doubt about the good prospects of this research field, which also demands from all of us constant upgrade and the necessity of establishing collaborations with multidisciplinary teams in order to reinforce the cross-cutting approach.

5. CONCLUSION

Along this paper, we have tried to synthesize the main arguments that encourage us to keep on working and investing on this research line, which also proves our concern about climate change and our commitment in the support of renewable energies. The geographical specificity of renewable energies has been explained, justifying the necessity of using a multidisciplinary and integrating approach. Both reasons support the use of GIS.

The work with this type of tools should be considered as the laboratory where the researcher approaches the challenge from a territorial perspective, considering all the agents involved in the energy system. Moreover, GIS should act more prominently in the post-process stage, encouraging its capabilities in the field of technology and knowledge-transfer to the society and decision support.

That is the reason why the integration concept should be driving the planning process. In order to this planning process to succeed, it is necessary to consider the difficulties that would be faced which are basically related to the reliance on the sources of information (which could be reinforce through the encouragement of metadata and SDIs creation) and on the application itself (which could be increased by a proper analysis of the problem, the selection of the appropriate tools and an accurate sensibility analysis of the results).

As a conclusion, it must be stated that the use of geographical information technologies applied to renewable energies regional integration could and must be encouraged. This support will clearly contribute to the creation of leading groups with high technological potential and to the promotion of RES through an in depth knowledge of the regional reality.

REFERENCES

Amador, J. and Dominguez, J. Application of geographical information systems to rural electrification with renewable energy sources. *Renewable Energy. 30 (12)*, 1897-1912, 2005.

Amador, J. and Dominguez, J. Spatial analysis methodology applied to rural electrification. *Renewable Energy. 31 (10)*, 1505-1520, 2006.

CIEMAT. Mercado de producción de electricidad solar térmica con tecnología de torre. Aplicación a España. 1997.

Dominguez Bravo, J., Garcia Casals, X. and Pinedo Pascua, I. GIS approach to the definition of capacity and generation ceilings of renewable energy technologies. *Energy Policy. 35 (10)*, 4879-4892, 2007.

Dominguez, J. and Amador, J. Geographical information systems applied in the field of renewable energy sources. *Computers & Industrial Engineering. A Cluster on Planning and Management of Energy and Infrastructure Engineering Projects. 52 (3)*, 322-326, 2007.

Domínguez, J. and Marcos, M. J. Análisis de la producción potencial de energía con Biomasa en la región de Andalucía (España) utilizando Sistemas de Información Geográfica. *Cybergeo. 142* 16, 2000.

García Casals, X. and Domínguez Bravo, J. Renewables 2050: a report about the potential of renewable energy in peninsular Spain, paper presented at ISES Solar World Congress 2007: Solar Energy and Human Settlement, 2597-2601, 2007.

Pinedo Pascua, I., Domínguez, J. and Rodriguez, M. Assessment of solar electrification in cuba: a case study based on GIS for rural development, paper presented at ISES Solar World Congress 2007: Solar Energy and Human Settlement, 167-171, 2007.

Shiota, A. Recommended Guidelines for the use of renewable energy in rural electrification. 2002.

Solargis-Team. Solargis Handbook. Guidelines for the elaboration of regional integration plans for decentralized electricity production with renewable energies. 1996.

SWERA-Project, available on line. http://swera.unep.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

An Environmental Decision Support System for an Efficient Monitoring and Planning of Woody Biomass for Energy Production, with Remote Sensing Aid

F. Frombo[b], M. De Martino[a], R. Minciardi[b] , M. Robba[a,b], and S.B. Serpico[a]

[a] *Interuniversity Centre of Research in Environmental Monitoring, Via Cadorna 7, I-17100, Savona (Italy).*, michela.robba@unige.it, francesco.frombo@unige.it, mdemartino@unige.it

[b] *Department of Communication, Computer, and Systems Sciences, Via Opera Pia 13,16145, Genova (Italy),* michela.robba@unige.it, francesco.frombo@unige.it

Abstract: The proposed paper describes the architecture of a GIS-based Environmental Decision Support System (EDSS) that integrates software tools and sensors, such as: remote sensing for environmental monitoring and support to mathematical models, optimization models (for planning plant size, kind, and biomass harvesting), dynamic forest growth models, and models for carbon dioxide balance calculation. The proposed EDSS is innovative as regards the current literature because it will be a tool for operational monitoring, cost minimization and efficient planning intervention, in accordance with natural and anthropogenic events that occur on the sites. The features of the EDSS and of the user-friendly interface are showed in connection with a case study.

***Keywords**:* Environmental Decision Support System; Environmental Modelling; Remote Sensing Data; Geographic Information system.

1. INTRODUCTION

The sustainable biomass exploitation for energy production requires the effective monitoring, design, planning and management of the related biomass supply-chain, including the integration and processing of different typologies of digital data (topographic maps, forest inventories, remote sensing, GIS data), the optimization of the logistics operations and of the energy conversion process, the assessment of the environmental impacts, the ability to combine economically appealing and environmentally sustainable strategies, as well as the participation in the decision of the different actors and stakeholders. Nowadays, there are several technological solutions that are available to produce energy from biomass. In addition, there are also several methodologies, available from the scientific community, to manage digital information based on Geographic Information Systems (GIS) and Decision Support Systems (DSS) to support decisions and planning (Freppaz et al., 2004; Voivontas et al., 2001).
On the other hand, the statistical and information sources about forest resources are heterogeneous and often fragmented, and difficult to be interpreted. Moreover, the in situ campaigns for forest inventories and resource quantification have very high costs. In this framework, tools that are able to interpret data, quantify environmental impacts, minimize costs for planning and management strategies are necessary. Among them, Remote Sensing (RS) can be a valuable aid to make available accurate information about the static and dynamic characteristics of the vegetal biomass. In the current literature, there are several works that show the importance of RS techniques for forest biomass assessment and

exploitation. Among them, Saatchi et al. (2007) show how radar sensors from airborne or spaceborne platforms have the potential of providing quantitative information about the forest structure and biomass components, Zheng et al. (2004) estimate the aboveground biomass using Landsat 7 ETM+ data, and Baltzer et al. (2003) examine the retrieval of tree growth from multi-temporal spaceborne L-band synthetic aperture radar (SAR).
In particular, the RS can represent a tool of pre-analysis, monitoring and evaluation of biomass resources for energy production. This tool can provide a basic set of thematic data to be supplied to the proposed EDSS, with the aim to replace part of the "traditional" cartographic and inventory information that often are available but too old and fragmented to be considered reliable.
Another important characteristics of the proposed EDSS is the user-friendly interface that easily allows scenario customization. Moreover, the EDSS is able to take into account different issues (environmental, energetic, economic and social) for the sustainable planning of the forest biomass use.
The potential users of the EDSS are Public Authorities that want to locate, size, and choose the technology for a plant, and want to define the harvesting plan, taking into account economic costs and environmental sustainability. The EDSS can also be used by a private company that needs to make a feasibility study for a plant in a territory, and that needs to evaluate the sustainable options and to minimize costs.
In the following sections, the EDSS architecture and the features are described in detail. Then, the different modules that can be used for different planning strategies are presented. Finally, a specific section is dedicated to the RS module and its integration into the EDSS.

2. THE EDSS ARCHITECTURE

The developed EDSS allows the planning of the biomass use in a given region, in order to minimize costs and guarantee the environmental sustainability.
In particular, the architecture is based on the following modules:

1. the GIS-based interface – a data aggregation system of territorial information that allows mapping data, working with different layers, and creating different scenarios (plant location, industrial biomass sources, forest parcels selection);
2. the database - for a suitable management of the information, the data planned in the GIS module and the results deriving from the optimization module are stored in a relational database;
3. the Remote Sensing (RS) module - strictly integrated with the GIS-based interface and the database system – it includes the necessary algorithms to elaborate the satellite data and to integrate them with the forest models and the optimization tools. In particular, this module is devoted to the evaluation of the current state of biomass resources, the vegetation species, the ground slope, the road network, etc. The multitemporal remote sensing analysis can also detect the natural or anthropogenic modifications that occur in the study areas;
4. the optimization module, subdivided in strategic planning and tactical planning to face the different aspects of the planning process.

The users can view the territory via a GIS oriented interface. The territory is divided in parcels, characterized by a homogeneous biomass type. As a first step, the users can customize the problem and create scenarios.
By default, the system appoints as eligible all the parcels. However, it is possible to eliminate those parcels that cannot be considered for harvesting because, for example, they are hardly reachable, environmentally protected, or the owner does not allow their use. Moreover, it is possible to add other biomass collection sites, such as, for example, biomass deriving from agriculture/industrial production, by inserting location, biomass quantity, and purchase cost. As regards plant typologies, it is possible to pick the plant directly on the map through the interface. Moreover, one can choose the plant typology and change parameters. After that, a procedure calculates the distance of the centroid of the selected parcels from the first available road and from the plant location. All data and calculations are stored in a database. Then, the optimization procedure is called. When the optimization procedure ends, the optimal results are shown on the map. For a suitable management of the information, the data planned in the GIS module and the results deriving from the

optimization module are stored in a relational database. Communication with the database is managed by a proper ODBC (Open Database Connectivity) interface, while the optimization module is called within the MS Visual Basic 6.0 program by a specific Lingo 8.0 component.
Figure 1 shows the interface that allows integrating all the modules. In particular, on the top of the interface there are utilities that can help to save files, modify parameters directly linked to the database, zoom, spam, parcels characteristics interrogation, insertion of plant location and industrial biomass sources data and location (using images that can be placed directly on the map), parcels selection for the specific Scenario. On the left there is a column where it is possible to select the following layers: provinces, communes, biomass parcels, territorial images. Once a layer is selected, it appears on the right, directly on the territorial area. It is also possible to decide which layer wants to be brought to front or behind other layers (by simply putting the layer on the top of the legend).

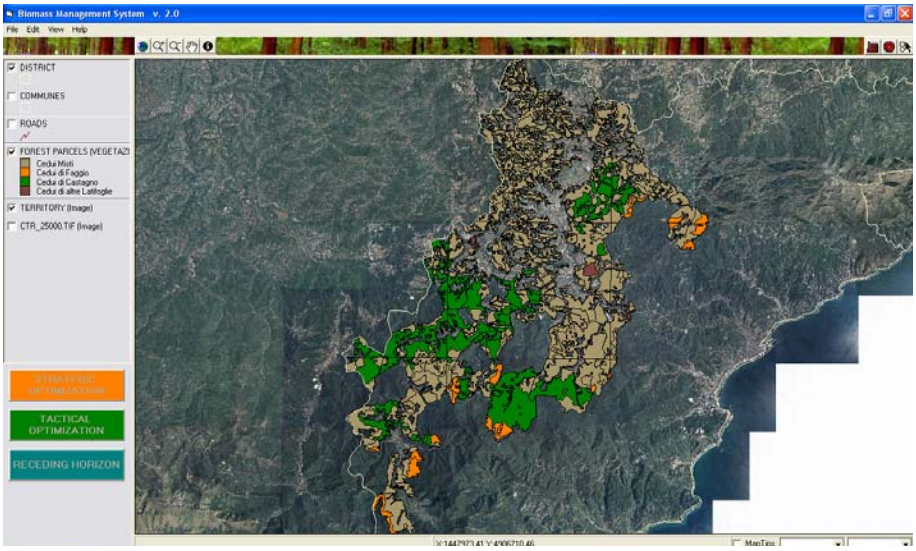


Figure 1. The software interface: the Val Bormida (Savona Province, Liguria Region, Italy) study area.

3. THE STRATEGIC PLANNING DECISION TOOL

3.1 The Decision Problem Description

Strategic planning level decisions refer to plant sizing and location, to the choice of the technology to be adopted, to the degree of exploitation of forest resources. Such a model corresponds to a linear programming problem. It is supposed that there is a unique plant location, which has been already selected, and that the matter of the decision problem is relevant to the plant size and the harvesting to be yearly performed, taking into account constraints deriving from Italian regulations over forest exploitation. The different plant technologies can be taken into account in the decision problem as they determine the values of specific parameters. Running the model for the various technological options, the user can obtain the optimal results for combustion, gasification, and pyrolisis processes and compare results on the basis of economic and environmental considerations. The decision variables are those quantities that represent the decisions to be taken, and are necessary to formalize the objective function and the constraints. In particular, for the strategic planning decision model, they are the annual biomass quantity harvested in the *j-th* slope class of the *i-th* parcel, and the plant capacity, expressed in terms of the maximum developed thermal power for the *k-th* technology. The objective function takes into account the costs and the benefits of the decision problem. In particular, in order to evaluate the overall cost function to be minimized, it is necessary to consider felling and processing, primary transportation (i.e. the transportation from the parcel to the first available road), transportation (i.e., from

the first available road to the plant location), purchasing and plant costs, as well as the benefits deriving from the sales of the products.

3.2 Inputs and parameters in the EDSS

The EDSS dynamically changes the parameters related to plant characteristics when a specific technology is selected. Specifically, the plant technologies considered in the EDSS are: Grate Firing Combustor and Steam Cycle – GFC; Fluid Bed Combustor and Steam Cycle – FBC; Fluid Bed Gasification and Gas Engine – FBG; Fast Pyrolysis and Diesel Engine – FP. The Scenario is set taking the image showing the plant and putting it in the desired location. When the picture of the plant is placed by the user in the desired map, another window appears (on the top-left). Here it is possible to insert the name of the plant and to select the plant typology among the four different technologies. The technology brief description appears when the kind is selected. On the bottom of the window the plant location coordinates appear. Clicking the "Ok" button, the system calculates the necessary distances between the parcels and the plants, and selects the parameters relevant to the specific plant.
The same procedure can be followed for agro-industrial point sources of biomass. The wheel image on the top-right of the interface is selected and located on the map on the desired point locations. A window appears in which data about biomass type, quantity, heating value, moisture content and price are asked.
All Scenario data are transferred to the database and the software is ready for the use of the optimization modules.

3.3 EDSS results presentation: application to a case study

The system has been applied to the consortium of municipalities in the mountain community of Val Bormida, inside the Savona Province, to evaluate the sustainability of a power plant. The study area has a high tree density index, with respect to the Italian tree index; in fact, it is covered for a big part of its area by natural forest vegetation (mostly homogeneous hardwood forest). The total study area is about 28000 hectares, excluded the parts of forest territory that can not be exploited for the presence of legislative constrains (natural parks and protected area) and/or because they have been characterized by forest fires and hydro geological disasters or they are areas of great bio-naturalistic importance.
Figure 2 reports the results for the strategic optimization problem for GFC directly on the interface from where it is also possible to perform queries on the obtained results, directly working with the map. All results are stored in the database module. The selected parcels for the Scenario are shown in different colours in a thematic map, on the basis of the degree of use they are subject to. In fact, the thematic map represents, for each parcel, the quantity of harvested biomass per the available biomass that can be harvested (considering law limits and silvicultural knowledge for each biomass type). Three different main numerical outputs can be found on the interface: the Scenario Characteristics, the Scenario Detail, and the Parcels Detail. They can be viewed by clicking on the correspondent label on the interface. In Figure 2, the Scenario Detail is showed.

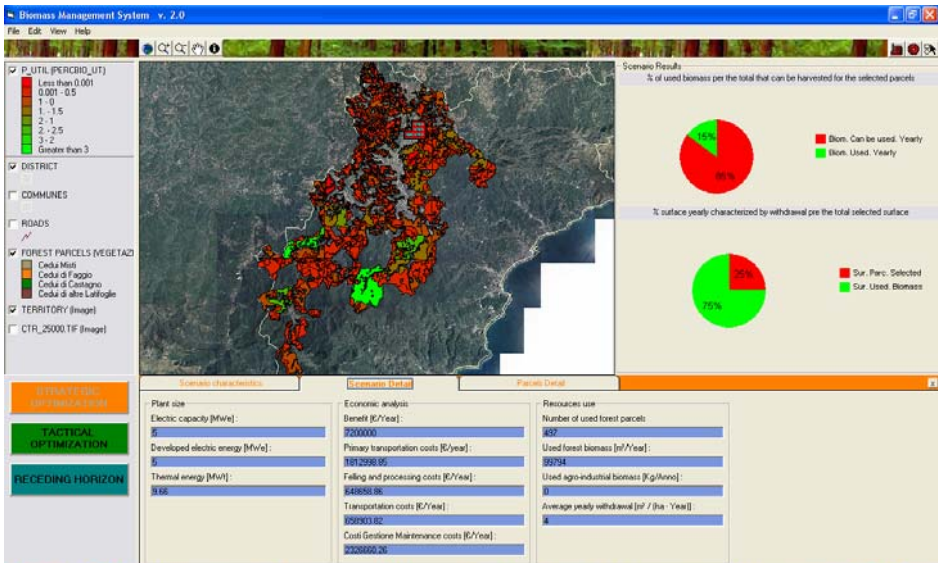


Figure 2. The optimal results for the strategic planning (GFC technology): "Scenario Characteristics" and "Scenario Detail".

4. THE TACTICAL PLANNING DECISION TOOL

4.1 The Decision Problem Description

Tactical planning level decisions refer to a medium-short term horizon, and are generally considered within a discrete-time setting, assuming that the decisions of the strategic planning level have been established (i.e., plant capacity and location are fixed). In this case, a dynamical formalization of vegetation growth models and carbon sequestration models is needed. In this work, the control variables are represented by the yearly amount of biomass harvested from the *i-th* parcel. The state variables, necessary to represent the system state, are the amount of biomass present in each parcel at a specific time, and the average age of a forest parcel. Three different types of cost, related to biomass harvesting, have been considered. Two state equations are then formalized as a function of the control variables. The objective function is then composed by: cost of forest biomasses felling and processing, cost of forest biomasses primary transportation, transportation cost from the landing points to the plant. Finally, a calculation of the carbon dioxide balance is implemented as a function of the state and control variables.

4.2 Inputs and parameters in the EDSS

At the tactical planning level, the creation of the Scenario is the same as the one described at the Strategic level. One can also choose to use the results for the optimal capacity selected by the strategic optimization. However, a possibility is given in order to select a desired plant capacity for the Scenario. When the tactical optimization problem is clicked, a window appears that allows the user to choose a plant capacity. From the same window, it is also possible to set the optimization time horizon. Data are then stored in the database. All parameters related to forest growth models and carbon sequestration models are instead in the database and it is not possible to modify them through the interface. Only an expert user can modify them through the database, according to the forest characteristics of a specific case study.

For the tactical planning model two possible optimization procedures are called: the tactical planning optimization, and the receding horizon control scheme. In the first case, the optimization horizon is selected and the optimization tool is called once, while, in the second case, a receding horizon procedure is called and also the simulation horizon can be chosen.

4.3 EDSS results presentation: application to a case study

The optimization has been run for the Val Bormida case study. Figure 3 reports the results for the tactical optimization problem. Also in this case, results for Scenario characteristics, Scenario detail, and Parcels detail can be calculated. Moreover, results for calculated emissions are reported.

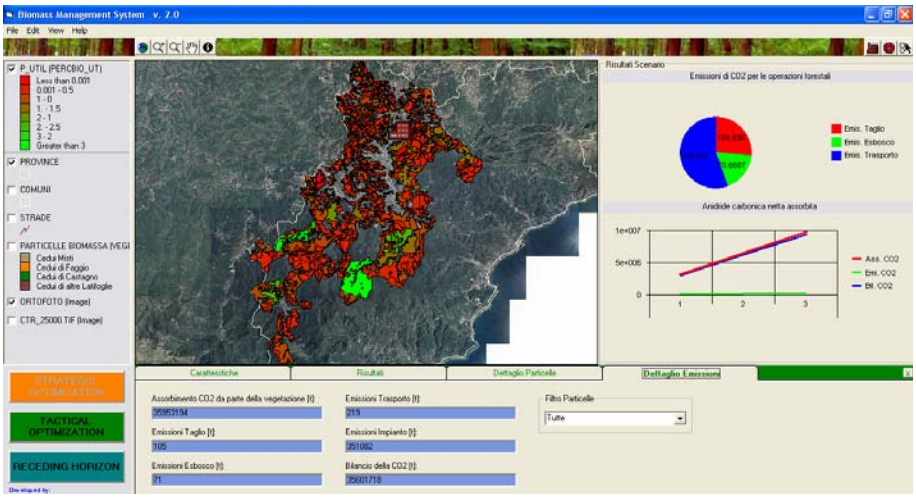


Figure 3. The optimal results for the tactical planning

5 THE REMOTE SENSING AID

5.1 The Remote Sensing Analysis

The main issues concerning forest management are depletion due to natural causes (fires and infestations) or human activity (clear-cutting, burning, land conversion), and monitoring of health and growth for effective commercial exploitation and conservation.
The analysis of the location, the extension, the health, and the productivity, the sustainability of such natural resources are critical information for natural land management. Remote sensing can reduce the cost of resource inventory and monitoring if remotely sensed data are well correlated with important field measurement, and available when needed.
The immediate advantages that Remote Sensing technology provides to a field as regards to forest management are the following: it makes possible a synoptic view of areas which otherwise should be investigated with enormous human as well as economic resource waste, it allows bypassing non-trivial obstacles, such as hard accessibility to some regions or the risk of modifying the environment by means of invasive in site inspections (Lillesand et al., 1987).
In order to manage and interpret this kind of data, it appeared to be wise to attempt to construct a data analysis system that takes advantage of keen perceptive and associative powers of humans in conjunction with the objective quantitative abilities of computers.

5.2 The Remote Sensing Module

The proposed RS module deals with the preprocessing and application of the remote sensing analysis for the exploitation of digital data in order to generate biomass information thematic maps that represents added-value layers in input to support the EDSS procedures.

With this purpose, this module will be developed to produce the following thematic layers, based on the exploitation of remote sensing data by image processing techniques, for the areas of interest:

- Archive and current satellite images for monitoring/visualization purposes;
- Vegetation species maps;
- Extension of the vegetation cover;
- Productivity Index (PI), as available wood biomass [kg]/area unit [m^2];
- Digital Elevation Model (DEM) and the related slope information.

All these kind of thematic layers have to be localized in plane geographic coordinates and updated when needed, then directly made usable in EDSS. All the layers will be stored in the database module and be available on the GIS interface (left side column, Figure 1) to be displayed on the right side, once selected.

Remote sensing, providing a detailed mapping of land at high resolution, can be an immediate decision support tool in the evaluation phase of biomass variables of the resource management and modelling, in the updating of cadastral information, in the assessment of the land, and in the evaluation of dimensions and areas of forest sites. Furthermore, the remote sensing analysis can be dynamic if multi-data images are processed with the purpose to detect natural and anthropic modifications, growth trend and so on.
Based on these considerations, the RS module aims to support both the strategic and the tactical decision approach. An example, in the strategic planning approach will be considered the possibility to "adjust" the value of a local state variable, like PI, by the correspondent RS information to obtain a *calibrated* PI value for each considered parcel by a suitable algorithm. In the tactical approach, the use of multitemporal RS data will allow the validation of the growth trend of PI in the areas of interest, for example considering at least RS data of the last 15 years.

The integration of the produced thematic layers in the EDSS is needed with the purpose to increase the reliability of the system. In this context, the use of remote sensing data, together with the extensive extraction of useful information as support tool to the management of the energy production by forest biomass, is particularly innovative.

6 THE REMOTE SENSING THEMATIC LAYERS: AN EXAMPLE

The great amount of remotely sensed data production requires to be analyzed before being available to the wide range of their customers.
With this purpose, this section aims to consider an adequate approach that makes thematic mapping by means of RS data feasible and reliable, exploiting the power of application-oriented data processing.

6.1 Application-oriented RS Data

In the case of application-oriented data processing, a brief example of RS layer generated in the RS module is showed in the following. It refers to Digital Elevation Model (DEM) obtained from radar satellite data.
Figure 4 shows a 3D-visualization of a DEM of Liguria Region, Italy, where mountainous chain is present very close to the sea. The different colours show the elevation ranges, where the red areas are the peaks.

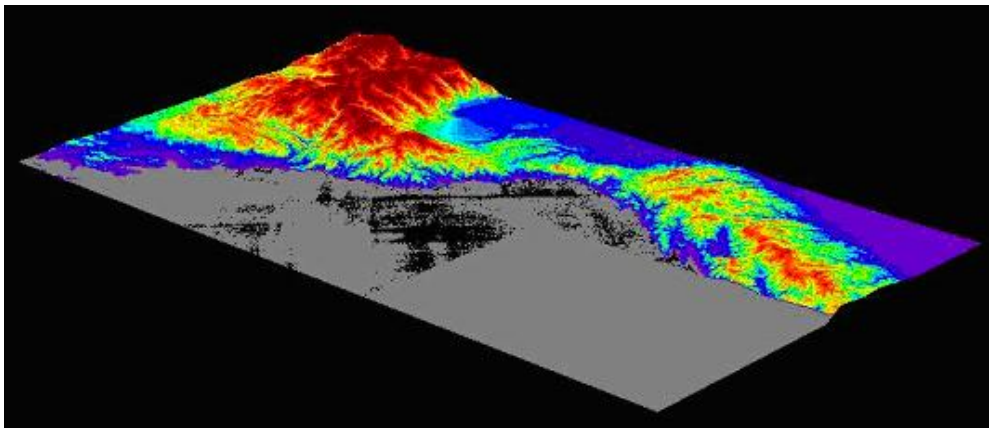


Figure 4. DEM of Liguria Region, Italy, obtained from radar satellite data.

Other information can be derived from this kind of RS product, like an elevation map where the lines at the same altitude are represented in different colour (Figure 5). From this kind of map is possible to extract the slope information.

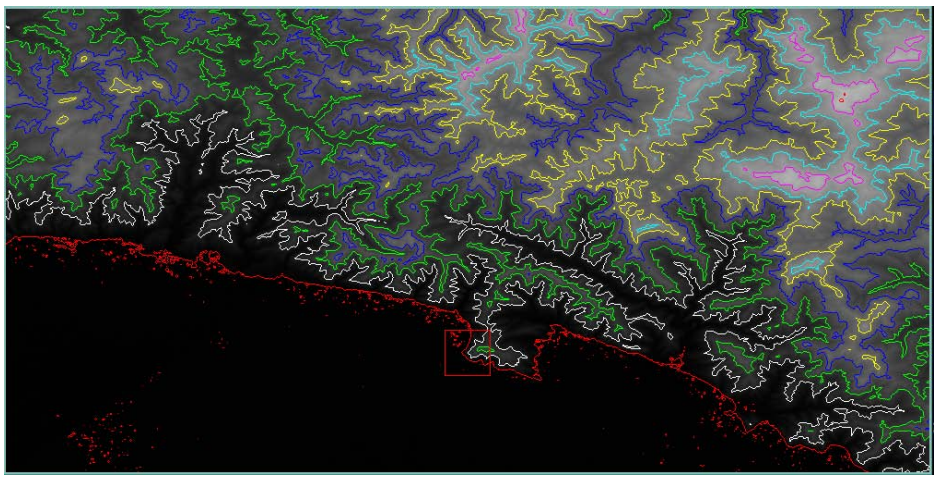


Figure 5. Altitude map obtained from radar satellite data in Fig.4.

Furthermore, it is important to note that DEM data (Figure 4) can be used in GIS interface to support the three-dimensional visualization of other available layers, like ortho-photos, vegetation and PI maps.

7 CONCLUSION

An EDSS for the monitoring and planning of forest biomass use for energy production and the features of the GIS-based interface is described in detail. The architecture includes different modules: the GIS-based interface, the database, the Remote Sensing (RS) module, and the optimization module, subdivided in strategic planning and tactical planning optimization problems. The EDSS is showed in connection with the application to a specific case study.

The main objective of the proposed work is to create a system able to transfer knowledge, data, software tools, and prototypes to the stakeholders in order to satisfy the requirements of effective monitoring, design, planning and management related to the wood biomass use for energy production.

The idea is based on the application of an innovative approach to biomass exploitation for energy production, based on a sustainable management and planning taking into account environmental, energetic, economic and social issues in order to develop a permanent market of wood production for energetic purposes that allow, on the other hand, to activate regular land maintenance to mitigate the hydrological risk and forest species degradation.

Future developments regard the implementation of the operational management optimization problem and the algorithms to process and integrate data from RS in order to

generate information thematic maps that represents added-value layers in input to the described EDSS.

REFERENCES

Balzter H., Skinner L., Luckman A., Brooke R., Estimation of tree growth in a conifer plantation over 19 years from multi-satellite L-band SAR, *Remote Sensing of Environment 84*, 184-191, 2003.

Freppaz D., Minciardi R., Robba M., Rovatti M., Sacile R., Taramasso A., Optimizing forest biomass exploitation for energy supply at a regional level, *Biomass and Bioenergy 26*, 15-25, 2004.

Lillesand,T., and R.W. Kiefer, Remote Sensing and Image Interpretation, *2nd Edition*, 1987.

Saatchi,S., Halligan K., Despain DG, Crabtree, R.L., Estimation of Forest Fuel Load From Radar Remote Sensing, *IEEE Transactions on geoscience and remote sensing*, vol. 45 (6), 1726-1740, 2007.

Voivontas, D., Assimacopoulos, D. Koukios, E.G., Assessment of biomass potential for power production: a GIS based method, *Biomass and Bioenergy, 20*, 101-112, 2001.

Zheng, D., Rademacher, J., Chen, J., Crow, T., Bresee, M., Le Moine, J., Ryu, SR., Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA, *Remote Sensing of Environment, 93*, 402-411, 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

Rapid Sizing of Renewable Energy Power Components in Hybrid Power Plants for Reverse Osmosis Desalination Process

A. Kartalidis, G. Arampatzis and D. Assimacopoulos

National Technical University of Athens, School of Chemical Engineering, Process Analysis and Plant Design Department, 9 Heroon Polytechniou, Zografou Campus, GR-15780, Athens, Greece. e-mail: assim@chemeng.ntua.gr

Abstract: The cogeneration of water and electricity through the exploitation of Renewable Energy Sources is becoming an increasingly promising option, especially for arid and remote areas, where alternative energy supply is either unavailable or too costly to develop. This paper presents a new methodological approach for the preliminary design of a Renewable Energy (RE) power plant, primarily aimed at meeting the energy requirements of a Reverse Osmosis desalination unit. The design of the power plant's components involves the calculation of the installed capacity of each RE component (Photovoltaics and Wind Energy Conversion Systems), the size of the energy storage system (Battery), and required auxiliary energy sources (Diesel Consumption). The sizing of the two RE components and the calculation of the corresponding energy production is performed using a simplified mathematical model for the diurnal variation of wind speed and solar radiation. The overall approach is applied for the rapid sizing of components and the estimation of auxiliary energy supply needed for a medium-sized desalination unit in Tunisia, and is complemented with a preliminary economic assessment of the power plant costs.

Keywords: Renewable Energy Sources; Reverse Osmosis; Wind Power; Photovoltaics; Desalination

1. INTRODUCTION

The co-generation of electrical power and desalinated water from renewable energy sources (RES) is a very promising option [Koroneos et al., 2007], especially in arid and isolated from the grid regions, where the use of conventional energy is costly or unavailable [Elhadidy et al., 2000]. The energy and water requirements in such regions can be satisfied through a combination of RE components, conventional diesel generators and energy storage systems [Houcine et al., 1999; CRES, 1998]. The design of such a hybrid power system would require an overall system engineering approach, which can be further integrated into a Decision Support System (DSS) to address the problem of decision making in a generic way. The decision-making process involves three steps. The first step comprises the sizing of the power system, based on a minimum set of data and guided by specific design goals; the output of this step is a number of alternative configurations (scenarios). The second step involves the assessment of different indicators for each scenario, based on simulation techniques using detailed data. The final step is the evaluation of the scenarios on the basis of the selected indicators, following a multi-criteria decision analysis approach.

The work presented in this paper focuses on detailing the methodology for the first step of the decision making process, where alternative component sizing configurations are defined. In previous works, the definition of the required battery size and diesel usage is performed on the basis of simulation results, using general rules of thumb. For example, a rule of thumb proposes that battery power supply for three days reduces diesel consumption by 45% [Elhadidy et al., 2000]. Such rules of thumb cannot give accurate results, because

they do not take into account the combined energy production from the two renewable energy subsystems. The proposed methodology takes into account diurnal fluctuations in wind speed and solar radiation, and transforms the corresponding mean monthly meteorological data into daily profiles of renewable power production, without the use of simulation.

Often, the design of a RES-powered desalination plant needs to address conflicting goals, rules and constraints. These can be overcome using a scenario-based approach, where a set of alternative configurations (scenarios) for the size of the different system components is defined. These scenarios can then be fed into a Decision Support System for more detailed simulation and design, and, subject to economic evaluation and performance assessment, can be used to define the final structure of the power plant.

Section 2 of this paper provides an overview of the methodology for preliminary design, including a description of the power plant and the specific sizing and modeling algorithms. Section 3 presents the application of this approach for the sizing of a RES-powered plant used to supply a brackish water reverse osmosis plant in Tunisia (capacity of 50 m^3/d). Section 4 presents the main outcomes of the analysis undertaken and outlines orientations for further research.

2. METHODOLOGY

2.1. Power Plant Description

The energy subsystem of the Renewable Osmosis (RO) power plant combines different renewable energy sources, backed up by conventional ones (e.g. Diesel Generator). A configuration representative of this subsystem is presented in Figure 1.

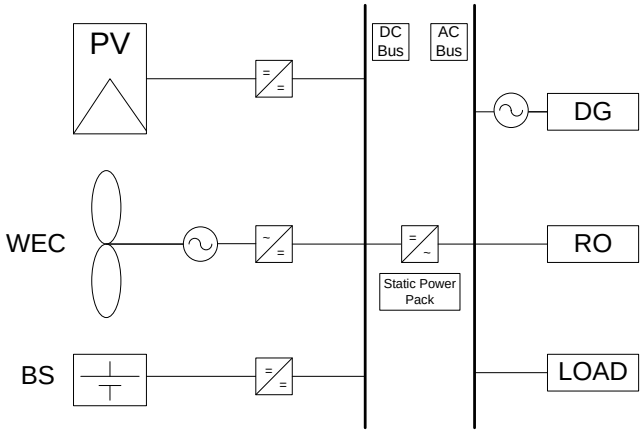


Figure 1. The hybrid energy configuration.

The main parts of the system are the: (a) Wind Energy Conversion system (WEC); (b) Photovoltaic system (PV); (c) Diesel Generator (DG); (d) Battery Storage system (BS); (e) Reverse Osmosis plant (RO); and (f) other power loads (LOAD), if the system is designed to supply energy to additional units at the installation area.

In such hybrid systems, and under adequate solar radiation and/or wind speed conditions, the WEC and PV feed the load demand (RO plant and additional power loads). The excess energy (i.e. the energy above this demand) from the WEC and PV is stored in the battery until full storage capacity is reached. If output from the WEC and PV exceeds the load demand and the battery is fully charged, then excess energy is drained away (undelivered energy) or fed back into a utility grid (if the system is connected). The diesel back-up is used to support the system in meeting the load demand should the WEC and PV systems fail to manage the load and battery storage be depleted.

2.2. Design Goals and Operational Constraints

The system design is based on the following set of goals:

- The local Wind and Solar Energy Potential are exploited to the maximum possible extent (MAX-RES goal).
- The undelivered excess energy (the energy drained away from the system) is minimal (MIN-UNDELIVER goal).
- The capital cost for infrastructure investment is minimal (MIN-CAP goal).
- The operational cost is minimal (MIN-OP goal).
- Environmental impacts (i.e. CO_2 emissions) are minimal (MIN-ENV goal).

At the operational level, there are two important constraints to guarantee the continuous and stable operation of the reverse osmosis unit. The first outlines that there should always be energy available for water production (Continuous Operation constraint) and the second that the power supplied to the RO high-pressure pumps should be stable (Stable Operation constraint).

The above goals and constraints are used in the design phase, to derive a number of rules that allow the sizing of the system (design rules). If a design parameter influences various goals in a conflicting manner (trade-offs) then it is treated as a decision variable and alternative values are examined. This corresponds to a "scenario-based" approach, defining the maximum, minimum and intermediate sizes for system components installed power.

2.3. Design Stages and Design Rules

The design process is structured in three sequential stages:

Stage 1 addresses the estimation of the energy and power requirements of the RO unit during the design period. This is a preliminary stage that addresses the energy requirements of the RO plant, which depend on water demand and the salinity of the feed water. The power demand is taken as constant and the additional electric load (as shown in Figure 1) is assumed to be zero.

Stage 2 involves the sizing of the renewable energy components (PV, WEC or both). This stage addresses the estimation of the size of the renewable energy components (PV and WEC) that will achieve, where possible, the design goals and satisfy the operational constraints of Section 2.2. The design variables are the photovoltaic installed power and the total rated power of the WEC system. All the above factors greatly influence the capital cost of the power system. The design rule that allows the estimation of the RE components size is summarized as:

- "The renewable energy delivered during the design period (year or month) is equal to the energy required by the RO plant during the same period."

This rule is in agreement with the "MAX-RES" and "MIN-UNDELIVER" design goals and meets the "Continuous Operation" constraint. However, the above design rule can lead to alternative configurations of the hybrid system, because the same amount of energy can be produced by different combinations of RE components. Therefore, an extra decision variable is required in order to define the hybrid system, which is the contribution of each one of the components (PV and WEC) to the total energy produced. The decision variable needed is the ratio of energy produced from PV to the total energy produced from renewable sources denoted by "α" (PV contribution). Figure 2 and Figure 3 present the complete algorithms for sizing the PV and WEC respectively. The calculation for the WEC components uses the well known Weibull equation.

Stage 3 comprises the sizing of auxiliary energy components (BS and DG). The design variables for the auxiliary components are the battery capacity of the battery system (in Ah) and the annual diesel usage (in kWh). The capacity of the battery system (and consequently the number of batteries for a given battery type) mainly influences the capital cost of the power system. Diesel generators are relatively inexpensive to purchase but the energy they supply determines the operational cost and environmental impacts [Patel, 1999; Elhadidy et al., 2000]. The design rules that allow the sizing of the auxiliary components are:

- The installed power of the diesel generator is equal to the power demanded by the RO plant.
- The unmet (by RE components) energy demand is fully satisfied first by the excess energy stored in the battery, and then by the energy supplied from diesel engine.

Both rules are derived based on the "Stable Operation" constraint. The first rule is required in order to ensure stable operation during the time periods where renewable energy sources are unable to provide power. The second rule is in agreement with the "MIN-UNDELIVER" goal. Here too the above design rules can lead to alternative configurations, as the unmet demand can be satisfied by different configuration of auxiliary components. A new decision variable is required to define the system, the utilization of the diesel engine. This variable influences the "MIN-CAP", "MIN-OP" and "MIN-ENV" goals in a conflicting manner and thus two extreme cases are considered, as follows:

- The "Diesel-Min" case addresses the minimization of the plant operational cost and the environmental impacts by minimizing the energy supplied from the diesel generator (maximization of energy supplied from RES);
- The "Diesel-Max" case addresses the minimization of the plant capital cost, by minimizing the required battery capacity and the installed power of RE components.

2.4. Sizing Process and Modeling

The PV and WEC component sizing is undertaken under the primary design rule described in design step 2. The output from this process is the rated installed power of the photovoltaic panels and the nominal installed power of the WEC system, and the algorithms are shown in the figures below

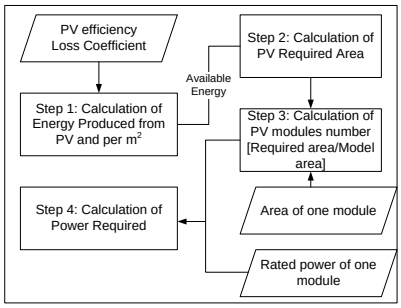


Figure 2. The PV sizing algorithm

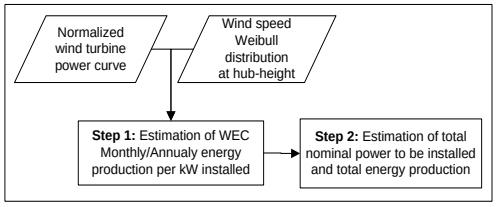


Figure 3. The WEC system algorithm

The design of the auxiliaries is based on the system performance over the whole design period, taking into account the variation of the renewable sources over a day, on the performance of the system over the whole design period. The approach is described in the following paragraph.

Firstly, a daily distribution of the renewable energy sources (solar radiation and wind speed) is assumed. This distribution is transformed into a model distribution of the renewable power produced over the day, so that the energy produced over the design period is equal to the renewable energy estimated in the design stage (for both Diesel-Min and Diesel-Max scenarios). This model distribution enables the estimation of the unmet energy demand (energy demand that cannot be directly satisfied by renewable sources) and the excess renewable energy (renewable energy produced in excess of demand for all the months of the year). The above energy quantities, combined with the design application of design rule 2, are used for sizing the battery and estimating the energy required by the diesel engine.

In general, the power distribution by photovoltaics (P_{PV}) is produced by a solar distribution in the shape of a triangle with its summit at the solar midday (12:00 PM). The power

distribution by WEC (P_{WEC}) is produced by a cosine curve that represents the diurnal variations of the wind speed [Manwell et al., 2006]. Figure 4 presents the distribution of renewable power produced (P_{RES}) over a day, and its relation to the load demand (P_d), where three different areas (which represent different energy quantities) are formed. The P_{RES} distribution is the sum (per time) of the P_{PV} and P_{WEC} distributions:

- $E_{RES,d}$ is the renewable energy produced and delivered directly to demand.
- $E_{RES,x}$ is the excess renewable energy
- $E_{d,u}$ is the unmet energy demand.

The sizing of the battery system and the diesel engine depends on the relation between the unmet energy demand and the excess renewable energy, where three cases are foreseen:

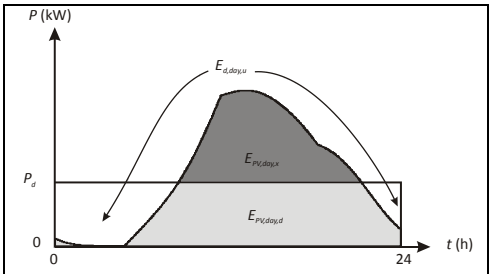


Figure 4. Distribution of renewable power and its relation to the load demand.

- $E_{RES,x} = E_{d,u}$ (the excess energy is equal to the unmet demand). All excess renewable energy will be stored in the battery system and will satisfy all unmet energy demand. The capacity of the battery system is equal to the excess renewable energy (or the unmet energy demand) and no diesel engine is required.
- $E_{RES,x} > E_{d,u}$ (the excess energy is greater than the unmet demand). Part of the excess renewable energy will be stored in the battery system and will satisfy all unmet energy demand. The capacity of the battery system is equal to the unmet energy demand and no diesel engine is required.
- $E_{RES,x} < E_{d,u}$ (the excess energy is less than the unmet demand). All excess renewable energy will be stored in the battery and will satisfy part of the unmet energy demand. The rest of the unmet demand will be satisfied by the diesel engine. The capacity of the battery system is equal to the excess renewable energy, and the energy required by the diesel engine is equal to the difference between unmet demand and excess renewable energy ($E_{d,u}$–$E_{RES,x}$) summed for all days of the period. The complete algorithm for the battery sizing and the estimation of the diesel usage is depicted in Figure 5.

The modelling approach is being applied to all examined system configurations. In order to define the two alternative extreme cases for auxiliary energy supply, design of the system is repeated for two months:

The month when the renewable source (solar radiation or wind speed) is at minimum (Worst month-based design). This case leads to the estimation of the size for the RE components that would maximize RE utilization and the energy required from auxiliary energy sources (Diesel-Min scenario). In this case, the defined battery size is larger than required, as the unmet energy requirements of the RO unit will be lower during the rest of the year, and the energy stored will never be fully supplied to the desalination plant.

The month when the renewable source is at maximum (Best month-based design). This case leads to the estimation of a minimum size for the RE components, but also to the maximization of the energy required from other auxiliary sources (Diesel-Max scenario). The estimated battery storage is again larger than that required, in the sense that the battery will never be fully charged during the rest of the year.

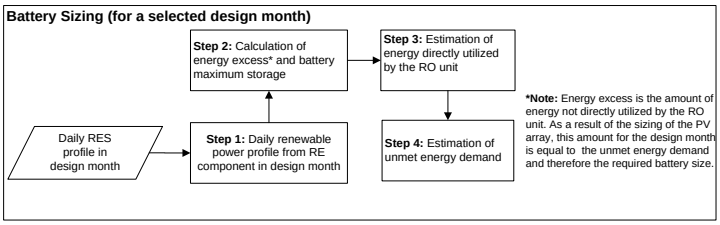


Figure 5. Algorithm for the sizing of the battery

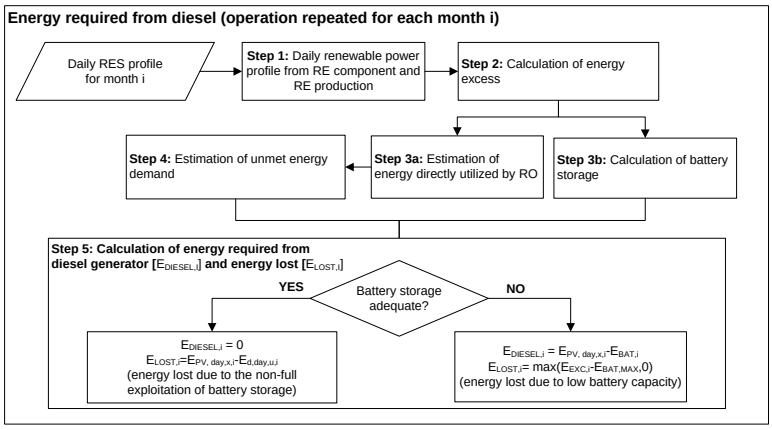


Figure 6. Algorithm for the estimation of diesel usage.

3. CASE STUDY

3.1. Implementation

The above methodology was implemented in a test site in Tunisia, where water demand is 50 m^3 per day and its specific energy for desalination is 2 kWh/m^3 (Brackish water) [CRES, 1998]. Wind and solar data can be acquired from public databases such us the NASA/RetScreen Database (eosweb.larc.nasa.gov) for wind data and METEONORM 6 demo application. The PV efficiency used for these calculations is 11.9%, with a module area of 1.26 m^2 and rated power of 0.15 kWp per module. The wind generator characteristic power curve is modeled by a (normalized) polynomial with appropriate coefficients. Additional data are the latitude (36.2^o) of the area, the time of the day that the maximum wind speed occurs (15:00 PM) and the DC bus voltage (48V). The mean monthly solar radiation per day ranges from 3.97kWh/m^2 for January to 7.32kWh/m^2 for July for the Latitude tilt. The mean wind speed ranges from 4.08m/s for July to 5.67m/s for February.

3.2. Results

The proposed installed power for both the PV and WEC system for the Diesel-Max and Diesel-Min is a linear function of the PV contribution (α). For PV contribution 50% the proposed installed power for PV is 17.5kWp for Diesel-Min and 9.5kWp for Diesel-Max and the installed power for the WEC is 21.8kW for Diesel-Min and 8.3kW for Diesel-Max. Figure 7 and Figure 8 show the battery capacity and the expected diesel consumption for the Diesel-Max case respectively.

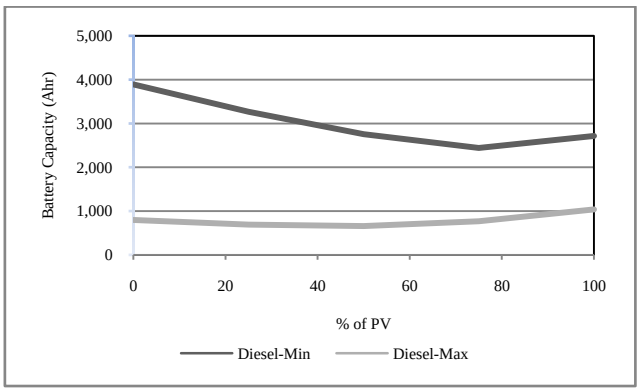


Figure 7. Battery system capacity as a function of the percentage of energy produced by PV

Figure 9 displays the basic energy indicators describing the utilization of the renewable and conventional energy for each scenario. The two extreme scenarios of Diesel-Min and

Diesel-Max indicate the minimum and maximum limits for the installed power of the power plant components, for use with reverse osmosis desalination. The acceptable diesel usage consumption, or other parameters such as environmental pollution, can be used to define the optimal size of the power components.

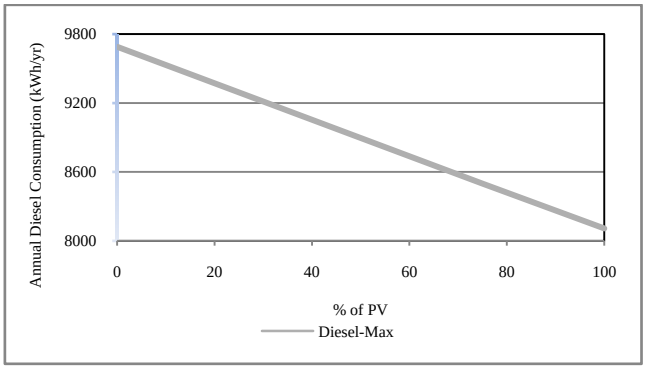


Figure 8. Annual diesel consumption as a function of the percentage of energy produced by PV

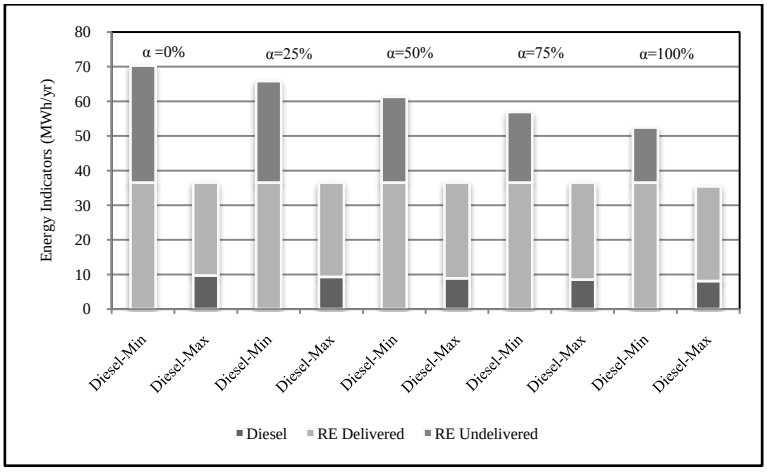


Figure 9. The main energy indicators for selected values of the percentage of energy produced by PV

Capital and operational costs are critical (but not the only) parameters for the final selection scenario (Diesel-Min and Diesel-Max). The design parameters calculated for these two scenarios are analyzed for their financial performance. The cost equations for each component can be found in Appendix A of Kaldellis et al. [2007]. The diesel price is set at 0.8 €/l, the battery life is set at 5 years and the life of the plant is set at 20 years. The results are shown in Figure 10.

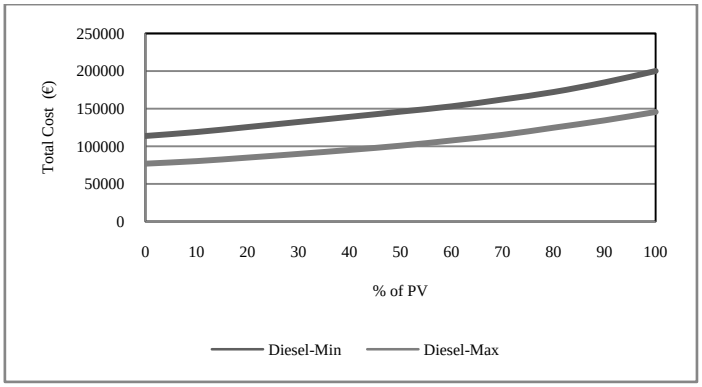


Figure 10. The total cost of the power plant for each extreme scenario and as a function of the percentage of energy produced by PV.

3.3. Discussion

It is evident from the indicators in Figures 10 and 11 that incorporating more photovoltaics into the renewable mixture (% of PV→100%) enables better exploitation of the energy sources for the area of the case study. Furthermore, the diesel usage in the Diesel-Max scenario decreases considerably. (It should be impossible to design a power plant that will operate in steady state and continuously only with the use of RES, and thus the diesel consumption should be greater than zero in the Diesel-Min scenario; however the proposed scenario is an extreme approach). When more WEC are introduced into the system (% of PV→0) the renewable energy undelivered is greater than the renewable energy delivered due to the great fluctuation of the wind speed. On the other hand, Figure 10 clearly indicates that the overall cost is less when more WEC are incorporated into the system. Figure 7 shows that the battery capacity has a minimum value when % of PV~80% at the Diesel-Min case and a minimum value when % of PV~50% at Diesel-Max case.

4. CONCLUSIONS AND SUGGESTION FOR FURTHER RESEARCH

A method for estimating the size of a renewable energy power plant was presented in this paper. The method is based on a rapid sizing algorithm which can be used to calculate alternative scenarios for the power plant of a reverse osmosis desalination unit driven by renewable energy sources. The calculations can be achieved under minimal set of monthly mean meteorological data. The effect of the variations of renewable resources on the size of the auxiliary components is estimated using proper diurnal profiles. The proposed rapid sizing algorithm was applied in a test site in Tunisia. Several alternative configuration scenarios can be designed easily showing that such rapid algorithms are important components of a decision support system due to the number of indicators that can be derived from this method.

The presented sizing method is only the first step of a decision support system for engineering design of a hybrid power system. The next steps involve the estimation of different indicators through simulation techniques and the evaluation of scenarios following a multi criteria decision analysis approach. As part of the decision support system can be also the sensitivity analysis for battery and diesel price and for battery life.

ACKNOWLEDGEMENTS

The above research was undertaken within the project "Optimal Engineering Design for Dependable Water and Power Generation in Remote Areas Using Renewable Energies and Intelligent Automation – OPEN-GAIN", supported by the European Commission through the 6th Framework Programme (Contract No: 032535).

REFERENCES

CRES. *Desalination Guide using Renewable Energy Souurces* - European Commision, Joule Thermie Programme, 1998.

Elhadidy, M.A. and Shaahid, S.M. Parametric study of hybrid (wind + solar +diesel) power generating systems. *Renewable Energy*, 21, 129-139, 2000.

Houcine, I., Benjemaa F., Chahbani, M.-H., Maalej, M. Renewable energy sources for water desalting in Tunisia, *Desalination*, 125, 123-132, 1999.

Kaldellis, J.K. and Kavadias, K.A. Cost-benefit analysis of remote hybrid wind-diesel power stations: Case study Aegean Sea islands. *Energy Policy*, 35, 1525-1538, 2007.

Koroneos, C., Dompros, A. and Roumbas, R. Renewable energy driven desalination systems modelling. *Journal of Cleaner Production*, 15, 449 – 464, 2007.

Manwell, J.F., et al. Hybrid2 - A Hybrid system simulation model, 2006.

Patel, Mukund. Wind and Solar Power Systems. CRC Press, 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

Particle Swarm Optimization for the biomass supply chain strategic planning

J. Izquierdo[1], R. Minciardi[2], I. Montalvo[1], M. Robba[2], M. Tavera[1]

[1] *Centro Multidisciplinar de Modelación de Fluidos, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain (jizquier,imontalvo@gmmf.upv.es, mata4@doctor.upv.es)*
[2] *DIST, Department of Communication, Computer and System Sciences, University of Genova, via Opera Pia 13, 16145, Genova, Italy (michela.robba@unige.it)*

Abstract: Particle Swarm Optimization is a very well established evolutionary optimization technique that has shown great potential for the solution of various optimization problems. In this paper, a real-life application, namely a biomass supply chain where optimal biomass flows from sources to energy production plants must be suitably defined, is tackled by using a variant of PSO recently introduced by the authors. In particular, the performance of the variant herein proposed is investigated by applying the model to the strategic planning of the biomass supply chain. The optimization problem is non-linear with binary and continuous decision variables and is based on a previous formalization regarding the biomass strategic planning, but with the addition of integer decision variables for plant technologies selection. The model has been applied to the mountain community of Val Bormida (Savona district, Italy).

Keywords: Biomass supply chain; Strategic planning; Optimization; Evolutionary algorithm; Particle Swarm Optimization.

1. INTRODUCTION

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. Recently, a number of algorithms that imitate certain natural principles, evolutionary algorithms like Genetic Algorithms, Ant Colony Optimization, Particle Swarm Optimization, Harmony Search, have provided with robustness to different applications.

Among them, Particle Swarm Optimization (PSO) is a very well established evolutionary optimization technique that has shown great potential and good perspective for the solution of various optimization problems [Dong et al., 2005; Izquierdo et al., 2008; Janson et al., 2008; Jin et al., 2007; Liao et al., 2007; Montalvo et al., 2008a; Pan et al., 2007]. 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 behavior and emergent intelligence that arise in socially organized populations.

Originally designed to deal with continuous variables, the PSO derivative we consider here overcomes 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 [Izquierdo et al., 2008; Montalvo et al., 2008a]. 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 [Montalvo et al., 2008b; Izquierdo et al., 2007]. Needing a low number of generations is a major advantage in real world problems, where costs and time constraints prohibit repeated runs of the algorithm and function evaluations. Finally, the cumbersome aspect, common to all metaheuristics, of choosing the right parameter values is tackled through self-adaptive dynamic parameter control [Montalvo et al., 2008].

One of its real-life applications is a biomass supply chain where optimal biomass flows from sources to energy production plants must be suitably defined. In particular, the performance of the variant herein proposed is investigated by applying the model to the strategic planning of the biomass supply chain.

The decision variables are represented by plant capacity, biomass yearly harvested in a specific forest parcel, and the presence/absence of a specific technology (pyrolysis, gasification or combustion), while the objective function is the sum of costs related to transportation, plant costs, and biomass collection costs. The localization of the conversion plant is assumed not to be a matter of decision but it is pre-specified by the user.

The optimization problem is non-linear with binary and continuous decision variables. It is based on a previous formalization regarding the biomass strategic planning approach [Freppaz et al., 2004; Frombo et al., 2006; Frombo et al., 2007], but with the addition of binary decision variables for plant technologies selection.

The model has been applied to the mountain community of Val Bormida (Savona district, Italy). In this area, the installation of a biomass-to-energy plant in the Cairo Montenotte municipality has been evaluated. The Val Bormida community is constituted by 18 municipalities for a total area of over 53000 hectares. The tree density index, about 75% (that corresponds to a forest vegetation of about 40000 hectares), is high, even compared with the Italian tree index, that is about 60%. The whole forest territory cannot be used due to legislative constraints (natural parks and protected areas), the existence of areas of bio-naturalistic relevance, and because of the occurrence of forest fires and hydro geological disasters. The whole Val Bormida forest vegetation and the area that is available for exploitation is about 27000 hectares. The available forest area corresponds to 506 forest parcels to be considered in the optimization problem. Each parcel is characterized by one main typology of biomass among four typologies (European Beech, Sweet Chestnut, Mixed and Hardwood Coppice) and by different slope classes.

The remainder of this paper is organized as follows. Next section introduces the optimization problem for the biomass supply chain at a strategic planning. Then, the fitness evaluation and the search space for the mathematical formalization are presented. Section 3 provides 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. Finally, the main results are reported. A conclusion section wraps up the paper.

2. THE OPTIMIZATION PROBLEM

The decision problem faced in this work regards the definition of the plant's size and kind for energy production from woody biomasses. Moreover, biomass collection over the territory should be defined, taking into account also slope characteristics. In the following, the decision variables, the objectives, and the constraints of the decision problem are described in detail.

2.1 Decision variables

The decision variables are those quantities that represent the decisions to be taken, and are necessary to formalize the objective function and the constraints. In particular, they are:

- $u_{i,j}$: the annual biomass quantity harvested in the *j*-th slope class, [%], $j = 1,...,J$, of the *i*-th parcel, $i = 1,...,N$, [m^3y^{-1}];
- δ_k : a binary variable that states the presence of the technology in a known location.

2.2 The objective function

The objective function takes into account the costs and the benefits of the decision problem. In particular, in order to evaluate the overall cost function to be minimized, it is necessary to consider felling and processing, primary transportation, transportation, purchasing and plant costs, as well as the benefits deriving from the sales of the products. The objective function is then composed by six terms that can be expressed as a function of the decision variables:

- C_{FP} , forest biomasses felling and processing costs. They depend on the required time for the operations (that is, of course, proportional to the harvested biomass and to the operation productivity), and on the number and types of operations executed in the forest;
- C_{FT} , forest biomass primary transportation cost. It corresponds to the transport from the felling areas to the landing points near the first available road;
- C_T , transportation cost from landing points to the plant;
- C_P , non forest biomasses (i.e., agricultural or industrial residues) purchasing cost;
- $C_{I,k}$, plant cost related to the *k*-th plant installation and management;
- B_k , benefits deriving from the products sale for the *k*-th plant.

The objective function to be minimized is then

$$C = C_{FP} + C_{FT} + C_T + C_P + \sum_{k=1}^{K} (C_{I,k}\delta_k - B_k\delta_k) \tag{1}$$

with

$$C_{FP} = \sum_{i=1}^{N}\sum_{j=1}^{5} \tilde{C}^{FP} \frac{u_{i,j}}{Pr^{FP}} \left(1 - \Delta_{Del}\delta_{Del} - \Delta_{Deb}\delta_{Deb} - \Delta_{Cc}\delta_{Cc}\right) \tag{2}$$

$$C_{FT} = \sum_{i=1}^{N}\sum_{j=1}^{J} u_{i,j} \frac{\tilde{C}_j^{FT}}{Pr_{i,j}^{FT}} \tag{3}$$

$$C_T = \sum_{i=1}^{N}\sum_{j=1}^{J} d_i^{LW}\ u_{i,j}\ VM_i\ \tilde{C}^T \tag{4}$$

$$C_P = \sum_{\substack{i=1 \\ i\in\phi_{NF}}}^{N} u_i VM_i \tilde{C}_i^P \tag{5}$$

$$C_{I,k} = \tilde{C}_k^I HVP_k\, ETA_k + \tilde{C}_k^M \sum_{i=1}^{N}\sum_{j=1}^{J} \left(u_{i,j} VM_i\right) \tag{6}$$

$$B_k = Ep_k^{Th} P^{Th} + Ep_k^{El} P^{El} + Ep_k^{Bd} P^{Bd} \tag{7}$$

$$Ep_k^{El} = \delta_{El,k}\left[\ \left(HVP_k\, ETA_k^t\, h\right) - Dt\ \right] ETA_k^e \tag{7a}$$

$$Ep_k^{Th} = \delta_{Th,k}\left[\ \left(HVP_k\, ETA_k^t\, h\right) - \left(1 - Dt\right)\ \right] \tag{7b}$$

$$Ep_k^{B} = \delta_{Bd,k}\left[\ HVP_k\, ETA_k^b\, h\ \right] \tag{7c}$$

where:

- $\tilde{C}^{FP}$ is the unit cost for the forest felling and processing phase, [€ m^{-3}];
- Pr^{FP} is productivity for the process of felling and processing;
- Δ is necessary time to execute the different operations of felling and processing (delimbing, debarking, and cross cutting);
- δ_{del}, δ_{deb}, δ_{Cc} are parameters indicating the presence/absence of the different operations of felling and processing (delimbing, debarking, and cross cutting);
- $\tilde{C}_j^{FT}$ is the unit cost for the forest primary transportation phase, [€ h^{-1}], associated to each slope class (i.e., primary transportation technique), j=1,…, J;
- $Pr_{i,j}^{FT}$ is productivity for the primary transportation phase, considering each parcel and slope class;
- $\tilde{C}^{T}$ is the unit cost for the forest transport phase, [€ kg^{-1} km^{-1}];
- d_i^{LW} is the distance from the forest parcel i, i = 1,…,N to the plant location, [km];
- $\tilde{C}_i^{P}$ is the unit cost for the purchasing cost of non forest biomass, [€ kg^{-1}];
- $\tilde{C}_k^{I}$ is the unit installation cost for the k-th technology, [€ MW^{-1} y^{-1}];
- $\tilde{C}_k^{M}$ is the unit maintenance cost for the k-th technology, [€ kg^{-1}];
- ETA_k is the overall plant efficiency for the k-th technology, [%];
- HVP_k is plant capacity, expressed in terms of the maximum developed thermal power for the k-th technology [MW];
- h is the useful working time in a year [hour];
- Ep^{El} is the production of electrical energy;
- Ep^{Th} is the production of thermal energy;
- Ep^{Bd} is the production of bio-oil;
- $\delta_{El,k}$ is a parameter that is set equal to 1 when the plant produces some amount of electrical energy, 0 otherwise;
- $\delta_{Th,k}$ is a parameter that is equal to 1 when the plant produces some amount of thermal energy, 0 otherwise;
- $\delta_{Bd,k}$ is a parameter that is equal to 1 when the plant produces some amount of bio-oil, 0 otherwise;
- ETA_k^e is the net electrical efficiency of the k-th conversion technology;
- ETA_k^t is the net thermal efficiency of the k-th conversion technology;
- ETA_k^b is the net bio-diesel efficiency of the k-th conversion technology;
- Dt is the amount of required thermal energy [MWh];
- P^{Th}, P^{El}, and P^{Bd} are prices for thermal and electric energy, and bio-oil, respectively [€/MWh], including the renewable energy certificates.

2.3 The constraints

Three different classes of constraints are included in the optimization problem: the restrictions over the forest biomass collection in order to prevent species extinction, the continuity equation at the conversion plant, and the constraint related to plant kind. Specifically,

$$u_{i,j} \le \alpha_i \cdot X_{i,j}^0 \; i = 1,...,N \,, \quad j = 1,...,J \,, \tag{8}$$

where $X_{i,j}^0$ [m^3] is the initial biomass present in each parcel i of acclivity j, and α_i [%] is an upper bound for biomass collection,

$$HVP_k = \sum_i \sum_j LHV_i \left(u_{i,j} VM_i \frac{1}{3600\, h} \right), \qquad k = 1,\ldots, K, \tag{9}$$

where 3600 is the number of seconds in one hour, VM is volumetric mass, and LHV is low heating value, and

$$\sum_{k=1}^{K} \delta_k = 1 . \tag{10}$$

3. PSO AND 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 particle, Y^* is identified. Let N be the number of particles in the swarm.

3.1 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:

$$V_i = \omega V_i + c_1\, rand(\,)\,(Y_i - X_i) + c_2\, rand(\,)\,(Y^* - X_i), \tag{11}$$

On each dimension, particle velocities are clamped to minimum and maximum velocities, which are user defined parameters,

$$V_{\min} \le V_j \le V_{\max}, \tag{12}$$

to control excessive roaming of particles outside the search space. Usually $V_{\min}$ is taken as $-V_{\max.}$

The position of each particle is also updated every generation. This is done by adding the velocity vector to the position vector,

$$X_i = X_i + V_i. \tag{13}$$

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*( ) represents a function which creates random numbers between 0 and 1; two independent random numbers enter Equation (11), which are used to maintain the diversity of the population.

3.2 Manipulation of parameters

The role of the inertia, ω, in (11), is considered critical for the PSO algorithm's convergence behavior. 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]

$$\omega = 0.5 + \frac{1}{2(\ln(k)+1)}, \tag{14}$$

where k is the iteration number. In the framework herein described this parameter is adaptively controlled by using (14).

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 (11) and (13). 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:

$$X_i = (x_{i1}, \ldots, x_{iD}, x_{iD+1}, \ldots, x_{iD+P}). \quad (15)$$

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 acceleration constants and velocity limit.

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:

$$V_i = (v_{i1}, \ldots, v_{iD}, v_{iD+1}, \ldots, v_{iD+P}), \quad (16)$$

and

$$Y_i = (y_{i1}, \ldots, y_{iD}, y_{iD+1}, \ldots, y_{iD+P}). \quad (17)$$

This way, by using equations (11) and (13), each particle will be endowed additionally with the ability to adjust 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 bring 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 velocities V_i are integer and consequently the new position vector components will also be discrete (since the initial position vectors were generated with discrete values for discrete variables). According to this idea, instead of Eq. (11), velocity updating for discrete variables turns out to be:

$$V_i = fix(\omega V_i + c_1\, rand(\,)\,(Y_i - X_i) + c_2\, rand(\,)\,(Y^* - X_i)), \quad (18)$$

where $fix(\cdot)$ implies that we only take the integer part of the result.

For another, in [Montalvo et al., 2008b], PSO was endowed with a re-generation-on-collision formulation, later generalized in [Izquierdo et al., 2007], 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.

- $k = 0$
- Generate a random population of M particles: $\{X_i(k)\}_{i=1}^{M}$, according to (15)
- Evaluate the fitness of the particles (only the first D variables enter the fitness function)
- Record the local best locations $\{Y_i(k)\}_{i=1}^{M}$; according to (17) the values of the corresponding parameters are also recorded
- Record the global best location, $Y^*(k)$, and the list of the m best particles to check collisions (including their corresponding parameters)
- While (not termination-condition) do
    - Determine the inertia parameter $\omega(k)$, according to (14)
    - Begin cycle from 1 to number of particles M
        - Start
            - Calculate new velocity, $V_i(k+1)$, for particle i according to (11), and take its integer part (for discrete optimization) for the first D variables, according to (18)
            - Update position, $X_i(k+1)$, of particle i according to (13)
            - Calculate fitness function for particle i and update Y_i
            - 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 and update the list of the m best particles
            - If particle i is not currently one of the m best particles but coincides with one of the selected m best particles, then re-generate particle i randomly (including its parameters)
        - End
    - $k = k + 1$
- 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.

4. PROBLEM SOLUTION

The main decision variables are the $u_{i,j}$, representing the annual biomass quantity harvested in the *j-th* slope class, $j = 1,...,J$, of the *i-th* parcel, $i = 1,...,N$, [m^3y^{-1}], which are continuous decision variables. For understanding the problem dimensionality it should be noticed that 506 forest parcels have been considered and each one is divided in 5 sub-parcels having different slopes. It means that there are 506×5 = 2530 decision variables, considering only $u_{i,j}$.

The presence/absence of a technology in a known location is also a decision variable. PSO algorithm did not consider a δ_k variable for every technology indicating its presence or not; instead, PSO considered a discrete decision variable δ, that takes values between 1 and the amount of technologies (5 for this problem), indicating which technology will be used. δ_k would be equal to 1 if and only if $k = \delta$. The cardinality of variable δ multiplies the solution space by the number of technologies to be taken into account. Only for the sake of illustration, if the range of every $u_{i,j}$ should to be discretized into 50 values, the amount of possible solution would be 5×50^{2530}; this number is really huge and gives an idea of the size of the solution space where PSO will be searching for an "optimal solution".

Maximum velocities were established considering variable types:

- Maximum velocity for continuous variables = 20% of variable range
- Maximum velocity for discrete variables = 50% of variable range

In any case, minimum velocity was established as:

- Minimum velocity = - Maximum velocity

The termination condition stopped the process if after 800 iterations no improvement in the solution had been obtained.

As a result, it was decided the most convenient technology to be used in the plant, which turned to be #5, corresponding to FBP (Fast Pyrolysis), and how much biomass should be harvested annually in every *j*-th slope of every *i*-th parcel (data not included in this paper). Regarding the used data for calculating, it was decided to use a plant capable of assimilating approximately 67% of total available biomass.

Taken into account the wide degree of generality used to tackle the problem, the use of the considered approach to deal with a similar problem for other regions would be straightforward, provided that the data are available in databases with the same characteristics as the ones used here. Also, adding new terms to the fitness function would represent neither a conceptual problem nor a long or difficult task in terms of programming.

5. CONCLUSION

In this paper, a PSO algorithm has been used to tackle the strategic planning of a biomass supply chain. In particular, the decision problem was to define the optimal amount of biomass to be harvested and the best technology to produce energy from forest biomass in a specific mountain community.

A simple example has been examined using just one plant, however, a software solution were made for more than one plant to be analysed in the same region.

Using the proposed algorithm, it can be considered any kind of relations between variables for evaluating the objective function. Graphical facilities should be useful to be incorporated in the near future for making a better visualization and data input-output.

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 Beca MAEC-AECI 0000202066 awarded to one of the authors by the Ministerio de Asuntos Exteriores y Cooperación of Spain.

REFERENCES

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.

Bonabeau, E., Dorigo, M., and Théraulaz, G., From Natural to Artificial Swarm Intelligence, Oxford University Press, New York, 1999.

Dong, Y., Tang, B.X.J., Wang, D. An application of swarm optimization to nonlinear programming. *Computers & Mathematics with Applications*, 49(11-12), 1655–1668, 2005.

Freppaz, D., Minciardi, R., Robba, M., Rovatti, M., Sacile, R., Taramasso, A. Optimizing forest biomass exploitation for energy supply at regional level. *Biomass and Bioenergy*, 15-25, 2004

Frombo, F., Minciardi, R., Robba M., Rosso, F., Sacile, R. Strategic Optimization of biomass flows for energy production: a logistic point of view. Proceedings of the 5th Biennial International Workshop Advanced in Energy Studies "Perspectives on Energy Future". Porto Venere (SP) Italy, September 2006. http://www.chim.unisi.it/portovenere/

Frombo, F., Minciardi, R., Robba, M., Rosso, F., Sacile, R. "Supply chain optimization of biofuels". Proceedings of the 15th European Biomass Conference and Exhibition" Berlin, Germany, 2007.

Izquierdo, J., Montalvo, I., Herrera, M., and Pérez, R., "A derivative of Particle Swarm Optimization with enriched diversity," submitted to Computational Optimization and Applications, 2007.

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, 2008.

Janson, S., Merkle, D., Middendorf, M. Molecular docking with multi-objective Particle Swarm Optimization. *Applied Soft Computing*, 8, 666–675, 2008.

Jin Y.X., Cheng, H.Z., Yan, J.I., Zhang, L. New discrete method for particle swarm optimization and its application in transmission network expansion planning. *Electric Power Systems Research*, 77(3-4), 227-233, 2007.

Kennedy, J. and Eberhart, R.C., Swarm Intelligence, Morgan Kaufmann, 2001.

Liao, C.J., Tseng, C.T., Luarn, P. A discrete version of particle swarm optimization for flowshop scheduling problems. *Computers and Operations Research*, 34(10), 3099-3111, 2007.

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.

Montalvo, I., Izquierdo, J., Pérez, R., Mora, D., Improved performance of PSO with self-adaptive parameters for computing optimal design of Water Supply Systems, submitted to *Engineering Applications of Artificial Intelligence*, 2008.

Pan, Q.K., Tasgetiren, F., 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.

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

Forest Biomass Sustainable Use for Energy Production: a dynamic optimization problem

F. Frombo, R. Minciardi, M. Robba, F. Rosso, R. Sacile

DIST – Department of Communication, Computer and System Sciences, University of Genoa, via Opera Pia 13, 16145 Genoa, Italy (michela.robba@.unige.it)

Abstract: The use of forest biomass for energy production requires a careful attention to the sustainable silvicultural practices that can guarantee the satisfaction of the environmental constraints, the control of the forest growth, the carbon stock, and the CO_2 emissions. This is a complex task because of the different environmental and economic issues (related to the characteristics of the territory, the energy demand, the forest biomass potential production, and the techniques for forest utilization) to be taken into account. Environmental Decision Support Systems (EDSS) are considered as valuable tools for the planning and management of renewable resources use for energy production. In this paper, an EDSS for the tactical planning of forest biomass use (i.e., for the planning over a medium-short term horizon, within a discrete-time setting and the assumption that the plant capacity and the sizing of all facilities are known) is proposed. In particular, attention is focused on a dynamic decision model. An optimal control problem, whose control variables are represented by the biomass quantity to be harvested, is formalized and solved through mathematical programming techniques. The novelties of the proposed EDSS regard the dynamical formalization of an optimal control problem for a sustainable use of the forest resources for energy production, the possibility of including different forest growth models (embedded as constraints in the optimization problem) and a carbon sequestration model as a function of the control and state variables, an accurate definition of forest felling and processing, primary transportation, and transportation costs, a constraint that limits the yearly harvesting to the biomass mean annual increment (calculated as function of the average age and of the control variables). The EDSS has been tested within the Val Bormida mountain community (Savona District, Liguria Region, Italy).

Keywords: Renewable Energy, Forest growth, Optimization, Carbon balance, Environmental Decision Support System

1. INTRODUCTION

The use of renewable energy resources, as promoted by European Union, may represent the main alternative to avoid a further increase of carbon dioxide concentration in the atmosphere. Within the framework of the use of renewable resources for energy production, there is a growing interest towards forest biomass. However, the use of forest biomass requires a deep analysis, especially as regards to its environmental sustainability. This is due, first of all, to the necessity of harvesting in a sustainable way according to the forest vegetation growth. Attention should be also dedicated to the CO_2 emissions of the whole process. The development of optimal strategies that can guarantee the use of wood resources, satisfying the environmental constraints and controlling the forest growth, the carbon stock, and the CO_2 emissions seem to be a hard and important task in silvicultural practices. Different forest management systems affect the capability of forests to sequester carbon. In this framework, the use of forest growth models is very important for a

successful woodland management. The last century has seen the development of the various forest models. However, despite the high potential of such models as management tools, only few of them seem to be suitable to be embedded in decision models for the forest management. Forest growth models generally applied as management can be classified, according to Gadow et al [2001] and Hasenaur [2006], within the following categories: a) highly aggregate volume-over-age models, used for regional yield forecasting; b) stand models used to predict the growth as a function of age; c) size class models used to predict the plants growth variations as regards the diameter distribution; d) individual tree models that provide information about the plant growth, on the basis of spatial relations. Forest growth models are necessary to understand carbon sequestration models. In fact, carbon assimilation and storage in the vegetal living tissues vary with the age of the plants and are strongly correlated with the increment of biomass (Masera [2001], Masera et al [2003]).
The general aim of this work is to define a decision model that can help in finding the optimal planning strategies over time for the sustainable use of forest biomass for energy production. These strategies ought to minimize costs and, at the same time, should guarantee the satisfaction of environmental constraints, the control of forest growth, the carbon stock and the CO_2 emissions. An optimal control problem has been formulated in which the control variables are represented by the quantity of forest biomass that is harvested in a specific time interval, while the state variables are the biomass quantity and the biomass average age. The decision model has been included in a GIS-based EDSS. The novelties of this paper with respect to other recent works (e.g. Freppaz et al. (2004)) regard the dynamical formalization of an optimal control problem for a sustainable use of the forest resources for energy production, the possibility of including different forest growth models (embedded as constraints in the optimization problem) and a carbon sequestration model as a function of the control and state variables, an accurate definition of forest felling and processing, primary transportation, and transportation costs, a constraint that limits the yearly harvesting to the biomass mean annual increment (calculated as function of the average age and of the control variables). Four different alternative forest growth models have been included: yield table, regional, stand, and matrix transitional models. For brevity, in the following sections, only the yield table model is described. The EDSS has been tested for the case study of Val Bormida (Savona district, Italy), for which the yield table model has been selected.

2. THE SYSTEM DESCRIPTION

The developed EDSS requires the information on the territory characteristics, the available biomass quantity, the plant size and location, the costs related to harvesting techniques associated to the different slope classes, the available roads, the forest growth models (and the related parameters) that can be used for the territory under concern. The EDSS outputs are the planning of the optimal harvesting over time, minimizing costs and guarantying good silvicultural practices. Moreover, other outputs regard the impact of the harvesting over the forest growth, and CO_2 emissions of the whole process, considering both emissions from technologies and uptake from the forest system. The main objectives of the present work is the use of forest growth models, with the aim of embedding them as constraints in the optimization problem formalization, and the quantifications of the carbon balance. In this section, the forest growth and carbon balance models are described in detail as a function of the control and state variables.

2.1 Modelling forest growth

The main features characterizing different forests are the species, the situ, the age, and the density. Forest growth is influenced by the above mentioned parameters and by their degree of homogeneity in a specific area. The spatial area for which the classification of forests is addressed varies on the basis of different factors, such as the degree of homogeneity, the available data, and the level of aggregation that can be (or has to be) pursued. A forest characterized by a high degree of homogeneity as regards the age of the plants is, in general, defined as *even-aged forest*. The whole population has the same silvicoltural phase and the same behaviour in time. On the contrary, a forest characterized

by the presence of trees at different ages on the same territory is called *uneven-aged forest*. Forests can be further classified as *monospecific* if they are represented by only one species, or *mixed* forest if more species cohabit together. Independently of the degree of variability, forests can be composed by representative units. On the basis of such units, a model may be classified as a whole stand model or a single-tree model, which differ on the detail level of the information required and provided. A stand is defined as an area relatively homogeneous in terms of vegetation structure, growth dynamics, and species composition, and contains a number of trees for which a common set of characteristics can be created. Therefore, a stand is conveniently defined as a homogeneous unit of the forest and may be characterized by even-aged or uneven-aged structure. All models are constituted by empirical equations derived by interpolation of data sets and originally developed for silvicoltural management purposes. Forest models applied in the EDSS take into account the whole stand while single tree models are not considered. Specifically, four classes of forest growth models have been considered: yield table, regional, stand, and transition matrix models. In this paper, for brevity, attention is focused on the yield table model. Yield table models consider even-aged forest structure and are monospecific. They are represented by highly aggregated yield-over-age equations and are used, specifically, for predicting the development of a forest in response to a series of periodic harvesting levels. The yield table represents a simple and widely used aid in forestry practice and is used when the tables (or analogous curves or mathematical equations) are available for the specific site. They represent the growth of the whole stand in terms of height (mean and dominant), diameter (at human breast height), number of trees for hectare, basal area, total biomass and annual increment (mean and current), in function of the stand age. In this work, the mean annual increment, I_i^t (per unit surface), for each forest parcel i, i=1,… ,N and each time interval (t, t+1), t=0,…,T-1, has been used to describe the forest growth as a function of time. The values that are found in the yield tables are fitted by a mathematical relation between I_i^t , and the forest age, Age_i^t [year]. This relation expresses the current annual increment as an empirical function of the forest age and is represented by $I_i^t = f_i(Age_i^t)$, where, in this paper, the following structure for function $f_i(\cdot)$ has been adopted:

$$I_i^t = a_i(Age_i^t)^2 + b_i(Age_i^t) + c_i \qquad i = 1,\ldots,N \qquad t = 0,\ldots,T\text{-}1 \tag{1}$$

where a_i, b_i, c_i are species specific parameters (corresponding to the dominant specie in parcel i).

The biomass at the end of the time interval depends on the annual biomass increment, and on the amount of biomass collected during the same time interval on the whole parcel i, u_i^t [m^3]. Thus, for the proposed model, the state variable is the whole quantity of biomass v_i^t [m^3] at the beginning of time interval (t, t+1), while the control variable is represented by u_i^t [m^3] and refers to the amount of biomass collected in time interval (t, t+1). Moreover, it is necessary to take into account that, every year, only a part of the whole biomass present on the parcel is collected and that the increment of the biomass is age-dependent. The new biomass grows in a different way with respect to the ones that are not collected. After thinning, the total biomass is given by the sum of the one that is present in the different sub-parcels with different rates of growth. In this paper, the forest parcel i is considered *even-aged*, and average age is calculated in each parcel after the thinning operation. Thus, the average age Age_i^t can be updated as follows

$$Age_i^{t+1} = \frac{u_i^t}{v_i^t} + \left(Age_i^t + 1\right)\cdot\frac{v_i^t - u_i^t}{v_i^t} \qquad t=0,\ldots,T\text{-}1 \qquad i=1,\ldots,N \tag{2}$$

where $\frac{u_i^t}{v_i^t}$ is the fraction of the parcel that is harvested in time interval (t, t+1), and $\frac{v_i^t - u_i^t}{v_i^t}$ is the fraction of the parcel that is not harvested in time interval (t, t+1).

The forest growth state equation is expressed by:

$$v_i^{t+1} = v_i^t - u_i^t + I_i^t \cdot S_i \qquad t=0,\ldots, T\text{-}1 \quad i=1,\ldots,N \tag{3}$$

where S_i is the forest parcel surface [ha].

2.2 Modelling carbon balance

The CO_2 balance model does not enter either the constraints or the objective of the optimization problem. However, it is important to express it as a function of the state and control variables in order to calculate the CO_2 emissions for the optimal solution. The general structure of the model follows the CO2FIX model (Masera [2001], Masera et al [2003]), as regards the quantification of carbon uptake or release. The model evaluates the net carbon release or uptake by the ecosystem with a step time of one year. The net annual carbon uptake or release, $NetCO_2^t$ [t CO_2 y^{-1}], is given by the difference between the carbon stored in wood, $CO_{2,stock}^t$ [t CO_2 y^{-1}], and the emissions due to biomass exploitation (transport, plant emissions, harvesting), $CO_{2,emis}^t$ [t CO_2 y^{-1}]. Both $CO_{2,stock}^t$ and $CO_{2,emis}^t$ are calculated as a function of the harvested biomass u_i^t. In the former case, the carbon stored in wood depends on the forest age that depends on u_i^t.
The CO_2 stored in the forest ecosystem, $CO_{2,stock,i}^t$ [t CO_2 y^{-1}], at time t in each forest parcel i, is given by the sum of carbon stored in the vegetation, Cv_i^t [t C y^{-1}] and the carbon stored in the wood products, Cp_i^t [t C y^{-1}]. To convert the amount of carbon into equivalent CO_2 absorbed, it is necessary to multiply for a conversion factor, K_{CO2}, that represents the ratio between the molecular weight of the CO_2 and of the carbon (equal to 44/12). Carbon in the living biomass, Cv^t, is determined as the sum of the carbon stored in each forest parcel i. Carbon stored in parcel i at time t, Cv_i^t, is calculated as the sum of the amount of carbon present in the parcel at the time t-1, $Cv_i^{(t-1)}$ [t C], plus the carbon stored by the biomass growth during the time interval (t-1, t), calculated as a function of the biomass annual increment I_i^t [m^3 ha^{-1} y^{-1}], minus the losses due to the harvest operations u_i^t [m^3 ha^{-1}]. The wood products represent a sink of carbon at different degree of release of CO_2 as function of the rate of degradation. Wood products kinds can be classified as long, medium and short term. Carbon in wood products is determined by the carbon content in the portion of biomass used for other destinations than the energy production, minus the share of the product that decomposes each year. The equations are omitted for brevity.

To evaluate the net amount of CO_2 released in the atmosphere from the whole biomass to energy supply chain, three major contribution should be taken into account: the carbon dioxide emissions from the conversion plant, CO_2fuel [t CO_2 y^{-1}], from the biomass transport, CO_2tr [t CO_2 y^{-1}], and from the forest operations, CO_2ha [t CO_2 y^{-1}].
The yearly emissions from the conversion plant depend on the amount of carbon in the fuel, Kc, and on the biomass quantity collected from each parcel i, u_i [m^3 y^{-1}]. These two parameters must be multiplied for a stechiometric ratio, K_{CO_2} (44/12), to obtain the amount of CO_2 released in the atmosphere, and for the parcel surface, S_i (ha). The yearly emissions from transport depend on the quantity of transported biomass. The transport emissions are divided in two phases, the emission deriving from the primary transportation, $CO2pt^t$ [t CO_2 y^{-1}], from the forest to a point near the first road (landing point), and the secondary transportation, $CO2sp^t$ (t CO_2 y^{-1}), from the landing point to the plant. Emissions during primary transportation phase are determined on the basis of the primary transportation system used, that is associated to the acclivity j (j = 1, …, J), and to the total distance from the parcel centroid to the landing points. For each type of primary transportation system, a specific efficiency is assigned on the basis of the fuel consume, $E_{j,pt}$ (km l^{-1} fuel). Moreover, for each system, the number of times that a forest vehicle employs to transport the biomass collected from the felling point to the nearest road is determined. Such parameter is determined on the basis of the amount of biomass collected from each parcel i, u_i (m^3 y^{-1}), and the forest vehicle capacity (vV_j). Emissions during secondary transportation phase are determined on the basis of the number of times employed by a vehicle, the distance to the plant, and the vehicles fuel consumption. Finally, CO_2 production from forest operations is expressed as a function of the amount of CO_2 emitted per m^3 of biomass cut by chainsaw, and of the harvested biomass. For brevity, equations are not reported here.

3. FORMALIZATION OF THE DECISION PROBLEM

3.1 The decision problem description

The main objective of this paper is to define a decision model able to determine the quantity of forest biomass to be harvested over time in a specific territory, in order to satisfy the material request from a biomass conversion plant and to guarantee a sustainable use of the forest resources. Specifically, great attention is focused on the forest system and on the definition of state equations, embedded as constraints in the optimization problem. In the following, an explicit formalization has been provided for the yield table model. The forest system is subdivided into a set of i-th parcels, $i = 1,.....,N$, defined on the territory on the basis of a homogeneous vegetation type. Each parcel i is moreover subdivided in sub-parcels in base of the specific acclivity. Five slope classes j, $j = 1,.....,J$, have been considered. The objective function includes costs related to the trees harvesting and transport. Two different forest operations, the felling and processing phase, and the primary transportation phase, according also to the area acclivity, have been considered. The felling and processing phase, FP , only regards operations executed by chainsaw, while the harvester and the processor are not taken into account. The felling process regards the biomass cutting phase, without with no other treatment or process. The processing phase (whose productivity is indicated as Pr^{FP} [m^3 h^{-1}]) regards the first phase of trees pre-treatment and it is constituted by three independent operations: the delimbing (Del), the debarking (Deb) and the cross cutting (Cc). Delimbing consists of removing the barks from the trees, debarking is the elimination of branches while cross cutting is the stem cutting in smaller parts. On the basis of many factors, like the territory configuration or the primary transportation techniques, one or more of these operations can be neglected. The absence of an operation causes the increase/decrease of the productivity and in particular of the overall time to perform the various operations. To model this possibility for each of these a reduction factor for the operation time, T [%], is found. A binary parameter δ for each possible operation allows to exclude one or more forest operations with the consequential improvement of the productivity and the reduction of the unit cost, C_U^{FP} [€ h^{-1}]. The forest primary transportation, FT , is the transport from the felling areas to the landing points near the first available road. The transportation technique mainly depends on the j-th slope class of the sub-parcel. For the sake of simplicity, the most suitable primary transportation technique is defined for a j-th class and consequently for each sub-parcel, choosing among sliding, skidding, trailer, yarder and chute. For each technique a certain productivity value, $\mathrm{Pr}_{i,j}^{FT}$ [m^3 h^{-1}], and unit cost, C_j^{FT} [€ h^{-1}] are established. The productivity value decreases with the increasing of the distance, d_i^{FL} [km], from the felling areas to the landing points, following the course of a decreasing monotonic continuous function specific for each transportation technique. The dependence of the monotonic function from the transportation technique is expressed from the numerical parameter p_j and m_j , evaluated from the fitting of bibliographical values. The forest primary transportation cost for each parcel is then obtained as the average sum of the costs of the j-th parcels. Instead, the biomass transportation includes the transport from the landing points to the plant or warehouse location. This cost is strongly dependent on the distance, d_{SP_j} , on the biomass quantity that must be transferred, and on a unit cost for transport, C_U^T [€ kg^{-1} km^{-1}].

3.2 The control and state variables

In this work, the control variables are represented by, u_i^t that correspond to the yearly amount of biomass harvested from the *i-th* parcel [m^3 y^{-1}]. The state variables, necessary to

represent the system state, are the amount of biomass, v_i^t [m^3], present in each parcel at a specific time *t*, and the average age of a forest parcel Age_i^t [y].

3.3 The objective function

The objective function is composed by three different cost; the cost of forest biomasses felling and processing, C_{FP} [€ y^{-1}], the cost of forest biomasses primary transportation C_{FT} [€ y^{-1}] and the transportation cost from the landing points to the plant, C_T [€ y^{-1}]. That is,

$$C = C_{FP} + C_{FT} + C_T \tag{4}$$

with

$$C_{FP} = \sum_{t=1}^{T}\sum_{i=1}^{N} C_U^{FP} \frac{u_i^t}{\mathrm{Pr}^{FP}}\left(1 - T_{Del}\delta_{Del} - T_{Deb}\delta_{Deb} - T_{Cc}\delta_{Cc}\right) \tag{5}$$

$$C_{FT} = \sum_{t=1}^{T-1}\sum_{i=1}^{N} u_i^t \sum_{j=1}^{J} \frac{C_j^{FT}}{\mathrm{Pr}_j^{FT}\left[1-\left(1-Exp^{\left(-m_j d_i^{FL}\right)}\right)^{p_j}\right]} \tag{6}$$

$$C_T = \sum_{t=0}^{T-1}\sum_{i=1}^{N} d_{SP,_i}\, u_i^t\, VM_i\, C_U^T \tag{7}$$

where VM_i is the volumetric mass for forest parcel *i*.

3.4 The constraints

Three classes of constraints have been formalized: the biomass growth state equation, described in section 2.1 (see (1), (2) and (3)), the restrictions over the biomass use, and the plant energy balance. As regards the constraints over biomass use, two kinds of constraints are formalized: thresholds imposed by regulation, and thresholds suggested by good silvicoltural practices. In the former case, a coefficient imposes to harvest just a fraction of the available biomass. In the latter case, it is imposed that it is not possible to harvest more than the total annual increment of biomass in each parcel. As regards regulation limits, the constraints for the generic time interval (*t*, *t*+1) are:

$$u_i^t \le \alpha_i v_i^0 \qquad i=1,\ldots,N;\ t=0,\ldots,T-1 \tag{8}$$

$$u_i^t \le \alpha_i v_i^t \qquad i=1,\ldots,N;\ t=0,\ldots,T-1 \tag{9}$$

where α_i is a species-specific parameter.

As regards silvicoltural good practices, it is possible to introduce a constraint on the amount of biomass annually harvested to determine a sustainable forest exploitation. Such constraint determines that the annual biomass collection, u_i^t, should be lower or, at maximum, equal to the mean annual increment, I_i^t, for the parcels that have average age $Age_i^t \le Age_i^{\max}$, and to the maximum increment I_i^{max}, otherwise. That is,

$$u_i^t \le \begin{cases} I_i^t & if \quad Age_i^t \le Age_i^{\max} \\ I_i^{\max} & if \quad Age_i^t > Age_i^{\max} \end{cases} \qquad i=1,\ldots,N;\ t=0,\ldots,T-1 \tag{10}$$

For a yield table forest growth model this means

$$u_i^t \le (a_i\left(\min\{\ Age_i^t, Age_i^{\max}\ \}\right)^2 + b_i\left(\min\{\ Age_i^t, Age_i^{\max}\ \}\right) + c\ )S_i$$

$$i=1,\ldots,N \quad t=0,\ldots,T-1 \tag{11}$$

The third class of constraints regards the energy balance at the conversion plant. As regards production plant constraints, the plant is supposed to operate at the maximum productivity level. The following equation states that the plant capacity *HVP* should be equal to the power developed by the conversion of the biomass entering the plant. That is,

$$HVP = \sum_i LHV_i \left(u_i^t VM_i \; f \right) \tag{12}$$

where $f = \frac{1}{3600\ h}$ is the conversion factor to transform thermal energy [MWh] in power [MW], with 3600 the number of seconds in an hour, and LHV_i is the low heating value.

4. APPLICATION TO A CASE STUDY

The EDSS has been tested within the Val Bormida mountain community (Savona District, Liguria Region, Italy). The study area includes 18 municipalities and the territory is characterized by a high forest density: it is covered by forests for the 75% of the total valley surface. Excluding the parts of the forest territory that can not be exploited for the presence of legislative constrains (natural parks and protected areas), for the occurrence of forest fires and hydro-geological disasters, for the presence of great bio-naturalistic importance areas, the total study area is about 28000 hectares. Such surface has been divided in 506 forest parcels that correspond to a specific homogeneous vegetation type as indicated by the forest map of the Liguria Region. Therefore, each parcel is characterized by one main typology of biomass out of five typologies of biomass (beech, chestnut, conifer, hardwood and softwood) located in various parts of the territory. Forest parcels have been divided into five classes of acclivity for the determination of forest operation costs. Biomass transport and energy production costs have been calculated too. The application of the EDSS in the territory of Val Bormida allowed determining the optimal harvesting management over time for a known plant capacity. A receding-horizon control scheme has been adopted. Moreover, a sensitivity analysis as regards different plant capacities (*HVP* = 3, 4, 5, 6 MWe) has been performed. For each plant capacity, the effects of different biomass collection levels on forest growth and on the potential carbon sequestration by vegetation are evaluated. In Table 1, the results of the optimization problem have been reported, for a time horizon of 15 years and for the different plant capacities.

Table 1. The results of the optimization problem

Cap [MWh]	**3**	**4**	**5**	**6**
Average harvesting biomass [m^3]	59606	79880	99345	119211
Yearly cost [€]	1.572.617	2.247.856	3.045.621	3.985.500
Yearly average cost [€ m^{-3}]	26,38	28,14	30,66	33,43
Total electrical energy [MW]	24 10^3	32 10^3	40 10^3	48 10^3
Used parcel	228	311	372	428
% on total biomass	0,58%	0,77%	0,96%	1,15%

An interesting analysis is to compare vegetation growth and CO2 balance in the optimal solution under different levels of biomass collection (that obviously correspond to the different plant capacities (3, 4, 6 MWe)). In Figure 1, the CO2 balance for the three plant sizes, is shown: the CO2 balance is never negative.

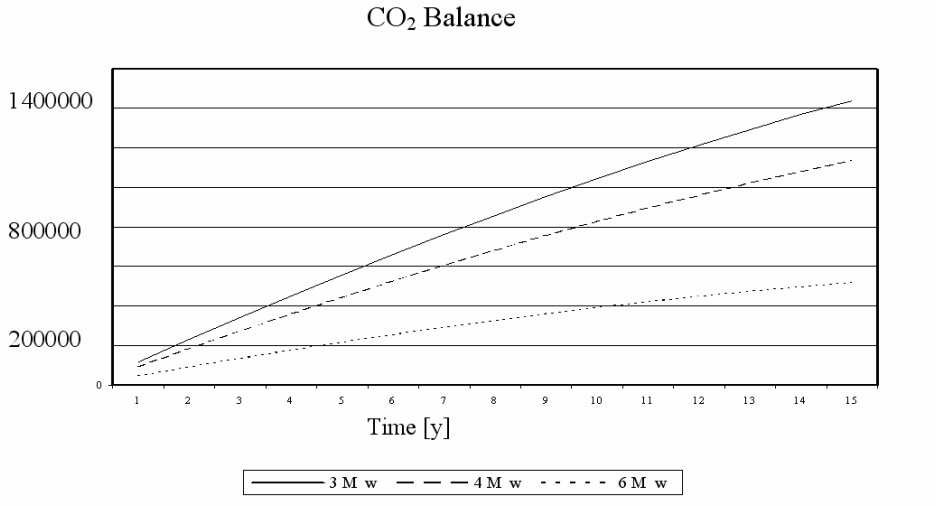


Figure 1. The carbon-dioxide balance

5. CONCLUSION

A decision model for the planning of forest biomass use for energy production has been described. Attention has been focused on the forest system in order to take into account sustainable silvicultural practices and the CO_2 balance that characterizes the biomass supply chain. The optimization problem has been solved for the case study of Val Bormida (Savona district, Italy), for which the yield table model has been selected. A sensitivity analysis as regards different plant capacities has been performed. The effects of different biomass collection levels on forest growth and on the potential carbon sequestration by vegetation are evaluated. The novelties of the proposed decision model regard the dynamical formalization of an optimal control problem for a sustainable use of the forest resources for energy production, the possibility of including different forest growth models (embedded as constraints in the optimization problem) and a carbon sequestration model as a function of the control and state variables, an accurate definition of forest felling and processing, primary transportation, and transportation costs, a constraint that limits the yearly harvesting to the biomass mean annual increment (calculated as function of the average age and of the control variables). Future developments of the proposed work regard, first of all, a more detailed characterization of the forest parcels, according to the slope and the available roads. Then, satellite data can be used in order to compare collected data for the forest system over the whole territory. As regards plant technologies, results have been found for combustion but can easily found for other plant typologies (pirolysis, gasification). Finally, a more detailed schedule of the forest primary transportation operation among parcels might be performed.

REFERENCES

Freppaz D., Minciardi R., Robba M., Rovatti M., Sacile R. and Taramasso A., 'Optimizing forest biomass exploitation for energy supply at a regional level', Biomass and Bioenergy, 26(1), pp. 15-25, 2004.

Gadow K., Hui G., Modelling Forest Development, Forestry Science, Kluwer Academic Publishers, Springer first Edition, 2001.

Hasenauer H., Sustainable Forest Management: Growth Models for Europe, Hubert Hasenauer (Ed.), Springer, 2006

Masera O.R., Carbon sequestration dynamics in forestry project: the CO2FIX V.2 model approach, Simposio internacional medicion y monitoreo de la captura de Carbono en ecosistems forestales, Valdivia Chile, 2001.

Masera O.R, Garza-Calligaris J.F., Kanninem M., Karjalainen, Liski J., Nabuurs G.J., Pussien A., de Jong B.H.J. and Mohren G.M.J, Modelling carbon sequestration in afforestation, agroforestry and forest management projects: the CO2FIX V.2 approach, Ecological modelling, vol.164, 177-199, 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

Optimizing biogas production: an application to an Italian farming district

G. Fiorese[a, b], G. Guariso[a] and A. Polimeni[a]

[a] *Dipartimento di Elettronica e Informazione, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy (fiorese@elet.polimi.it)*
[b] *Dipartimento di Scienze Ambientali, Università degli Studi di Parma, Via Usberti 33, 43100 Parma, Italy*

Abstract: Biogas can be extracted from animal manure through anaerobic digestion (AD) and can afterwards be used to produce energy. Moreover, this process reduces completely methane emissions and stabilizes the manure before its agronomic use. AD plants can be built in a wide range of capacities: as capacity increases, economies of scale in capital equipment are realized, but transportation costs increase as manure and the digested substrates must be conveyed over longer distances. It is thus a key issue to assess the tradeoffs between biomass' transportation and plants' capacity. We propose a method to evaluate the AD plants' convenience on a given territory by an economic, energy and emissive point of view. A mathematical model is formulated in order to optimize biomass use by finding the optimal AD plants' number, capacity, location, and the corresponding biomass collection basin. The method is applied to the district of Cremona, one of the most important Italian farming areas. The optimal solution is achieved by widespread AD plants over the territory in order to exploit biomass locally. Biomass transportation is minimized for its high costs are not balanced by economies of scale. AD plants in Cremona yield positive returns in economic terms, as energy produced and GHG emissions avoided (7% reduction with respect to 2003). The robustness of this result has been confirmed by sensitivity analysis of the plant and transportation costs. The final result is crucial for local planning of biomass exploitation: local governments can encourage the development of conversion plants at municipal level without the need for centralized decisions.

Keywords: anaerobic digestion; biogas; plant location; GHG emissions.

1. INTRODUCTION

Anaerobic Digestion (AD) is a biological process in which microorganisms break down biodegradable material producing biogas suitable for energy conversion, thus helping replace fossil fuels. Biogas can in fact be burned and converted into heat and power in cogeneration plants. Moreover, the nutrient-rich solids left after digestion can be used as fertiliser. In recent years, this bio-energy conversion technology has been developing as one of the most attractive renewable energy resources especially in Northern Europe (Germany, UK and Denmark) [Dagnall and Wooley, 2008]. The agricultural-zoothecnical sector represents a great source for the production of substrates for anaerobic digestion, e.g. agricultural residues, animal manure and energy crops.
In Italy, one of the main problems concerning farming areas is animal manure management, mainly because of methane and malodorous atmospheric emissions produced by manure, as well as excessive nitrogen load on soils. Currently only a storage period in tanks is imposed by law for this kind of waste before spreading on the agricultural land, but this system turns out to be ineffective with reference to the emissions issue. In this context, AD of manure represents a valid solution that offers at the same time the possibility to produce energy and to completely reduce CH_4 emissions, making the manure stable before its agronomic use.

AD plants can be built in a wide range of capacities. As capacity increases, economies of scale in capital equipment are realized, but transportations costs increase as manure and other substrates must be conveyed over longer distances to the plant site [Marrison e Larson, 1995]. As a result, estimating the convenience of biomasses transportation in realizing central plants turns out to be a key issue. This is particularly important for manure because of its low energy content. For these reasons, we propose a methodology to define the AD energy system configuration that optimizes plants' size and location, accounting for economic, energy and emissive performances. The optimization problem formulated for this purpose derives from the standard approach described in the literature [e.g., Drezner and Hamacher, 2001]. As a case study, the method is applied to the district of Cremona in Northern Italy.

2. BIOGAS ENERGY CONVERSION TECNOLOGY

Many different feedstocks can be fed to AD plants: agricultural residues, dedicated energy crops, animal manure [e.g., Borjesson and Berglund, 2007]. We assume in the following to feed only bovine and pig manure to AD plants. In order to reach a uniform organic matter content, row materials are mixed together before being fed to the digesters; these are continuously stirred tank reactors and operate with a maximum temperature of 55°C (thermophilic conditions). The digested effluent is stocked in one or more storage tanks. The gas produced in the digester is flushed through a condenser and a sulphide scrubber and collected in a gas storage tank. We assume that the biogas losses in the process are negligible.
The biogas produced is then used in a co-generation system to produce heat, with a constant efficiency (η_{heat}), and electricity, with an efficiency (η_{el}) function of the nominal plant capacity (Pe) as in Figure 1. The generator is connected to the plant circuits consisting of a mixer, pumps and gas blower and supplies the needed electricity. The heating of the digester and of the manure before being added to the digester is achieved by the hot water circuits. Capital costs for the construction of this type of AD plants are function of the plant nominal capacity, according to a moderate scale economy as pictured in Figure 2.

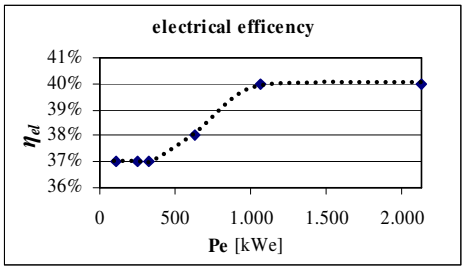


Figure 1. Electrical efficiency (η_{el}) as a function of AD plant nominal capacity [UTS, 2007].

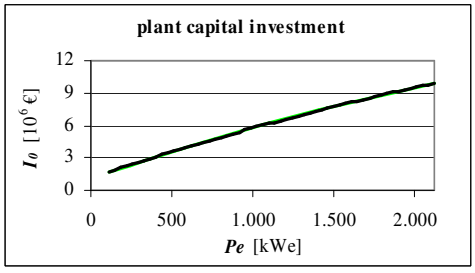


Figure 2: Capital investment (I_0) as a function of AD plant nominal capacity [UTS, 2007].

3. THE DECISION PROBLEM

A decision problem is formulated in order to optimize the biomass energy use by finding the optimal AD plants number, location, capacity, and the corresponding collection basin on a given territory. Similar models have been used to design the optimal use of ligno-cellulosic biomass in co-generation plants [Fiorese *et al.*, 2006; Freppaz *et al.*, 2004]. The studied area is divided into N parcels, each with its own biomass availability that is assumed to be concentrated in the centre of the parcel. Though not strictly necessary, we assume in the following that only one AD plant can be assigned to each parcel and the all biomass must be treated. As shown in Figure 3, the model comprises the following terms:

- The cost of bovine and pig manure transport from the i-th origin parcel to the j-th destination plant in special trucks with 30 ton capacity [Ghafoori *et al.*, 2007];

- The cost of digested biomass transport from the *j*-th plant back to its *i*-th parcel of origin where it is finally used as a fertiliser;
- The cost of building and operating each *j*-th AD plant;
- The revenues from the heat and the electricity sold, accounting for auto-consumption, and from the national incentives for renewable energy.

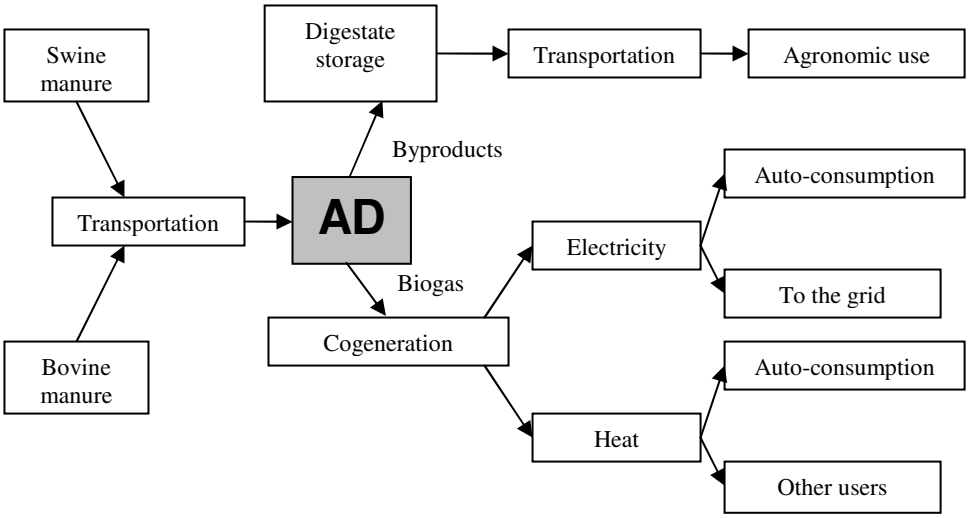


Figure 3. Bio-energy supply chain.

3.1 Decision variables

The decision variables (x_{ij}) are the fractions of biomass in the *i*-th parcel conferred to the *j*-th plant. Two auxiliary variables can be derived once the optimal values are determined. The first is the nominal plant capacity, Pe_j, of the *j*-th plant which is a linear function of biomass supply calibrated from experimental data of cogenerative plants currently acquirable in the area (capacities in the range 100 to 2100 kW):

$$Pe_j = -26{,}31 + 0{,}0003 \cdot \sum_{i=1}^{N} \sum_{s} (a_{i,s} \cdot f_{b,s} \cdot b_s) \cdot x_{ij} \quad (1)$$

The second, y_i, simply accounts for the presence or absence of a plant in the *j*-th parcel:

$$y_j = 1 \ if \ \sum_{i=1}^{N} x_{ij} > 0, \qquad y_j = 0 \ otherwise \quad (2)$$

3.2 Objective function

The decision variables value is assigned by maximizing the economic return from energy production that represents the objective function (J_{EC}) of the decision problem. The economic return is formulated through the net present value (*NPV*) method:

$$J_{EC} = \sum_{j=1}^{N} NPV_j = \sum_{j=1}^{N} \left(-\frac{I_{0,j}}{(1+r)^{-1}} + \sum_{t=0}^{19} \frac{F_{j,t}}{(1+r)^t} \right)_j \quad (3)$$

where *NPV* is calculated considering the discount rate *r*, assumed constant, the cash flow ($F_{j,t}$) for the *j*-th plant in year *t* and the capital investment ($I_{0,}$ Figure 2); a 20 years life period is assumed for AD plants, plus one year for plant construction. The annual cash flow is given by:

$$F_{j,t} = EC_{energy\ j,t} - \left(EC_{main,j} + EC_{oper,j} + EC_{transp,j} \right) \quad (4)$$

where EC_{energy} represents the expected benefits from energy production, EC_{main} and EC_{oper} represent plant maintenance and operation costs and EC_{transp} represents biomass transportation costs.
All the parameters used in the model are listed in Tables 1 and 2 (for details refer to Polimeni, 2007).

Table 1. Parcel dependent parameters in the objective function.

Symbol	Description	Unit of measure
d_{ij}	Distance between the *i*-th parcel and the *j*-th plant	km
$a_{i,s}$	Biomass available in the *i*-th parcel, *s*=1 for bovine manure and *s*=2 pig manure	t
$f_{b,\,s}$	Organic fraction in the *s*-th biomass	-

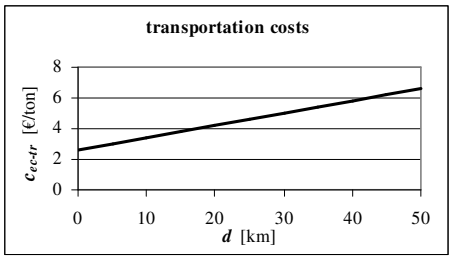


Figure 4. Transportation costs as a function of distance.

Table 2. Parameters used in the objective function and in the calculation of the emissive and energy indicators.

Symbol	Description	Value	Unit of measure
b_s	Biogas yield for biomass *s*	420 – 460 [(a)]	m^3_{biogas} /t_{om}
LHV_{biogas}	Biogas low heating value	19,4 [(a)]	MJ/m^3_{biogas}
fd_s	Digestate fraction for biomass *s*	0,96 – 0,98 [(b)]	-
OH	Hours of operation	7468 [(a)]	hours/year
η_{heat}	Plant heat efficiency	0,44 [(a)]	-
f_{el}	Electrical auto-consumption fraction	0,095 [(a)]	-
f_{heat}	Heat auto-consumption fraction	0,85 [(a)]	-
$fc_{ec\text{-}tr}$	Fixed economic transportation cost	2,6 [(c)]	€/t
$vc_{ec\text{-}tr}$	Variable transportation cost	0,08 [(c)]	€/km/t
$c_{en\text{-}tr}$	Energy for transportation	0,984 [(d)]	MJ/km/t
$c_{em\text{-}tr}$	Transportation emission	0,076 [(d)]	kg $CO_{2\ eq}$/km/t
ef_{cogen}	Cogeneration emission factor	0,147 [(e)]	kg $CO_{2\ eq}$/m^3_{biogas}
$ef_{ng,el}$	Natural gas electrical emission factor	0,098 [(d)]	kg $CO_{2\ eq}$/MJ
$ef_{gn,term}$	Natural gas heat emission factor	0,07 [(d)]	kg $CO_{2\ eq}$/MJ
p_{el}	Electrical energy price	0,074 [(f)]	€/kWh
p_{gc}	Green certificates price	0,125 [(f)]	€/kWh
p_{heat}	Heat price	0,08 [(g)]	€/kWh
c_{main}	Plant maintenance costs	0,03 [(a)]	€/kWh
T	Plant lifetime	20	years
t_{gc}	Green certificate duration	12	years
r	Annual discount rate	3	%
C	Maximum plant nominal capacity	2127	kWe

[(a)] UTS, 2007; [(b)] Provolo, 2004; [(c)] adapted from Ghafoori et al., 2007; [(d)] Spiellman et al., 2004; [(e)] NERI, 2001; [(f)] GRTN, 2007.

3.3 Economic benefits and costs calculation

Economic benefits from energy production are given by revenues from selling the heat and the electricity produced (excluding the fraction, f_{el}, needed for auto-consumption), and the associated green certificates, which are an Italian incentive for the development of renewable energy resources:

$$EC_{energy,j} = (p_{el} + p_{gc}) \cdot EN_{out-el,j} + p_{heat} \cdot EN_{out-heat,j} \quad (5)$$

The electricity $EN_{out\text{-}el}$ and the heat $EN_{out\text{-}heat}$ produced are calculated as follows:

$$EN_{out-el,\,j} = \frac{1}{3{,}6} \cdot (1 - f_{el}) \cdot \eta_{el,\,j} \cdot LHV_{biogas} \cdot \sum_{i=1}^{N} \sum_{s} (a_{i,s} \cdot f_{b,s} \cdot b_s) \cdot x_{ij} \tag{6}$$

$$EN_{out-heat,\,j} = \frac{1}{3{,}6} \cdot (1 - f_{heat}) \cdot \eta_{heat} \cdot LHV_{biogas} \cdot \sum_{i=1}^{N} \sum_{s} (a_{i,s} \cdot f_{b,s} \cdot b_s) \cdot x_{ij} \tag{7}$$

where $\eta_{el,j}$ is a function of *j*-th plant capacity as in Figure 1. Plant maintenance costs (EC_{main}) are calculated as a fraction c_{main} of gross energy output, whilst operation costs (EC_{oper}) are calculated as a function of the plant nominal capacity.
Finally, transportation costs (EC_{transp}) are the sum of manure ($EC_{tr\text{-}manure}$) and digestate ($EC_{tr\text{-}digested}$) round-trip transportation costs. They comprise both fixed costs ($fc_{ec\text{-}tr}$) representing loading and unloading operations, and variable costs ($vc_{ec\text{-}tr}$), function of the distance as shown in Figure 4 (adapted from Ghafoori *et al.*, 2007 to the current regional situation):

$$EC_{tr-manure,\,j} = \sum_{i=1}^{N} \sum_{s} \left[\left(vc_{ec-tr} \cdot d_{ij} + fc_{ec-tr} \right) \cdot a_{i,s} \cdot x_{ij} \right] \tag{8}$$

$$EC_{tr-digestate,\,j} = \sum_{i=1}^{N} \sum_{s} \left[\left(vc_{ec-tr} \cdot d_{ij} + fc_{ec-tr} \right) \cdot a_{i,s} \cdot fd_s \cdot x_{ij} \right] \tag{9}$$

3.4 Constraints

Two types of constraints have been defined. The first imposes that, in each *i*-th parcel, all the available biomass is carried to AD plants, because of the necessity of stabilizing manure:

$$\sum_{j=1}^{N} x_{ij} = 1 \qquad \forall i \tag{10}$$

and the second imposes that the nominal capacity of each *j*-th plants is limited to a maximum value *C* (due to the technology of the plant itself):

$$Pe_j \le C \qquad \forall j \tag{11}$$

To avoid very small plant sizes, a penalty has been established on low capacities.

4. THE ENERGY AND ENVIROMENTAL INDICATORS

An energy indicator (I_{EN}) and an emissive indicator (I_{EM}) are defined to evaluate the solution of the optimization problem. I_{EN} estimates the system net energy production (MJ/yr) from the heat and electricity (EN_{out}) minus the energy needed for the transportation of biomass, both of manure from the parcel to the plant and of the digestate the way back (EN_{transp}):

$$I_{EN} = \sum_{j=1}^{N} \left(EN_{out,\,j} - EN_{transp,\,j} \right) \tag{12}$$

I_{EM} assesses the system GHG mitigation potential (t $CO_{2,eq}$/yr): it comprises the avoided emissions (methane emissions from traditional storage, $EM_{storage}$; emissions from fossil fuels combustion for an equivalent amount of energy, EM_{fossil}) and the emissions produced (cogeneration emissions, EM_{cogen}; emissions for the transportation of both manure and digested substrates, EM_{transp}):

$$I_{EM} = \sum_{j=1}^{N} \left(EM_{storage,\,j} + EM_{fossil,\,j} - EM_{cogen,\,j} - EM_{transp,\,j} \right) \tag{13}$$

5. DESCRIPTION OF THE CASE OF STUDY

The optimization problem presented above has been applied to the district of Cremona, one of the most important Italian farming areas in Northern Italy. The district is located in the plain of the Po valley and it encompasses an area of 1.770 km^2 divided in 115 municipalities. The agro-industrial sector is of major importance in the economy of the district. Particularly, the area is characterized by a large number of feedlots: official data of year 2000 indicated an animal density of 157,2 bovines and 363,5 swine per km^2, whereas these figures are respectively 20,7 and 28,7 for Italy [ISTAT, 2000].
The amount of animal manure available in the district of Cremona is estimated from the number of bovines and pigs in the area [ISTAT, 2000]. The parameters used to derive the potential biomass supply and its organic content (the fraction important for the anaerobic digestion process) from the animal number are listed in Table 3, while the numbers of animals and potential biomass supply for the district are listed in Table 4. We estimate a total production of 190 million m^3 of biogas per year.
Because biomass is largely distributed over the area of study, we perform the analysis at the finer spatial scale allowed by the available data. This means we have used the municipalities as parcels, except for those whose biomass could supply more than one plant, which were split into two parts. Therefore, in the optimization model the number of parcels considered is equal to 124 (average size of 14 km^2).

Table 3. Parameter used to assess the availability of biomass.

Parameter	Bovines	Swine
Manure production (t/unit/day) [(a)]	0,05	0,01
Dry matter content (%) [(b)]	8	4,5
Organic matter content in dry matter (%) [(b)]	86	90
Organic matter content in manure (%) [(b)]	6,9	4,1

[(a)] Ab Energy [2007]; [(b)] UTS [2007].

Table 4. Number of animals in the district of Cremona, relative biomass supply and potential biogas production.

	Bovines	Swine
Animal units [(a)]	278.270	643.656
Manure production (10^3 t/year)	5.078	2.349
Organic matter production (10^3 t /year)	349	95
Biogas production (10^6 m^3/year)	147	43

[(a)] ISTAT [2000].

6. RESULTS

The overall non linear optimization problem can be solved by a commercial software package without significant computational problems. Its solution entails AD plants distributed over the territory in order to exploit locally produced biomass. For the whole district, the optimal solution foresees the realization of 105 plants of small-medium capacities (Figure 5). Only a few parcels (in white in Figure 5) do not have enough biomass on their territory and, therefore, do not host any plant. Three plants have a capacity over 1 MWe (the two largest plants have a capacity of 1,7 MWe), all other plants are smaller. This means that biomass transportation is minimized. Its high costs are not balanced by savings achievable exploiting the economy of scale of centralized plants and thus the upper capacity constraint is not active in the optimal solution.
The solution shows that biomass exploitation in AD plants yields positive returns in economic terms (the NPV exceeds 300 M€, but is slightly overestimated because the collection costs within each municipality have been disregarded), energy produced (1.500 TJ/yr) and $CO_{2,eq}$ emissions avoided (300 Gg $CO_{2,eq}$/yr). Economic return is strongly influenced by the presence of public incentives for power production from renewable sources: when the green certificates price (p_{gc}) is set to zero, the objective function may

become negative. However, in a carbon constrained world with a carbon market, the CO_{2eq} avoided emissions may also be evaluated in economic terms. Assuming a value of 19 €/t $CO_{2,eq}$ (ExternE, 2007), the objective increases of 26% and the overall plan may still be balanced even without incentives.

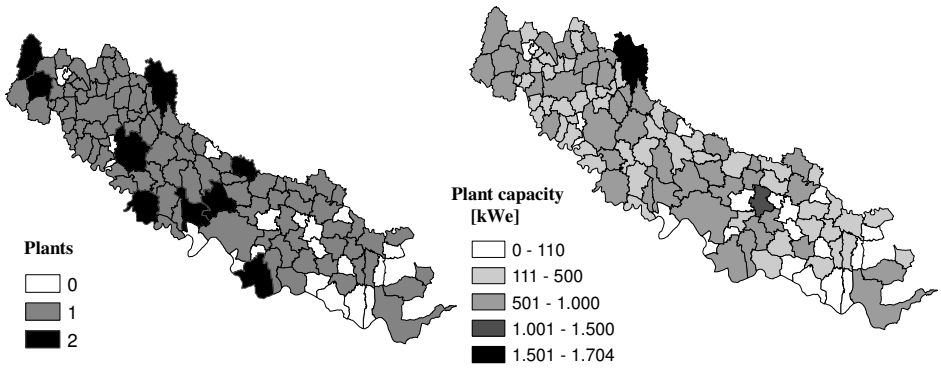


Figure 5. Optimal plant distribution in terms of number (left) and size (right).

An extensive sensitivity analysis was carried out to check the robustness of results obtained. Varying the plant cost curve shows that economies of scale cannot compensate transportation costs over a wide range of plant capacities. The optimal solution changes only when transportation costs are drastically reduced: if fixed transportation costs are set to zero, the optimal solution is characterized by 27 centralized plants. Moreover, even annulling fixed transportation costs, if the variable transportation costs increase, as it is quite probable in the near future, the optimal solution rapidly goes back to a more decentralized one, with a decrease of the economic performances of about 20% for a tenfold transportation cost, as shown in Figure 6 and Figure 7.

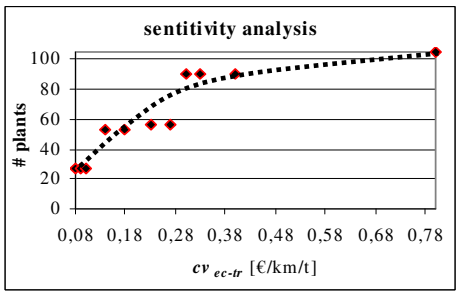


Figure 6. Optimal number of plants as a function of variable transport costs

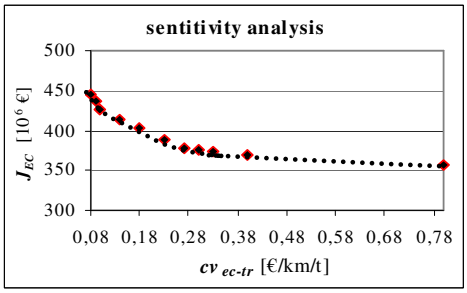


Figure 7. Optimal economic objective as a function of variable transport costs

7. CONCLUSIONS

The implementation of the proposed methodology led to an optimal solution in which the low energy content and the high manure transportation costs suggest the construction of small-medium AD plants distributed all over the studied area. Sensitivity analysis confirms that this solution is robust for a wide variation of plant investment and manure transportation costs. Thus local governments can encourage the construction of conversion plants at municipal level without the need for central planning, exploiting locally produced biomass. To implement this plan, the model solution should be integrated with GIS tools to define actual plant location taking into account the existing infrastructures, protected areas and other local constraints, as suggested by Ma *et al.* [2005].
The distributed solution in the district of Cremona leads to positive returns in economic terms (4 years for investment payback), energy produced (13% of the district power

consumption in 2005) and, most of all, avoided CO_{2eq} emissions (7% reduction with respect to 2003). The proposed plan allows the district of Cremona to effectively contribute to national CO_2 balance: the Kyoto Protocol set the Italian reduction target to 6,5% with respect to 1990.

ACKNOWLEDGEMENTS

The authors would like to thank the Provincia di Cremona for funding this research project within the development of the Energy and Environmental Program of the district.

REFERENCES

AB Energy, personal communications, Italy, 2007.

Borjesson P., Berglund M., Environmental systems analysis of biogas systems--Part II: The environmental impact of replacing various reference systems, *Biomass and Bioenergy,* 31 (Issue 5), 326-344, 2007.

Dagnall S., Wooley B., Overview of centralised and on-farm anaerobic digestion plants in Europe, http://www.ADNETT.org, 2008.

Drezner, Z., Hamacher, H.W. (Eds.), *Facility Location: Applications and Theory*. Springer Verlag, Berlin, 2001.

ExternE, http://www.externe.info/, 2007.

Fiorese G., Gatto M., Guariso G., Environmental and economic evaluation of biomass as an energy source in a farming area, Biomass & Bioenergy (submitted), 2006.

Freppaz D., Minciardi R., Robba M., Rovatti M., Sacile R., Taramasso A., Optimizing forest biomass exploitation for energy supply at a regional level, *Biomass and Bioenergy*, 26, 15-25, 2004.

Ghafoori E., Flynn P.C. , Feddes J.J., Pipeline vs. truck transport of beef cattle manure, *Biomass and Bioenergy*, 31 (Issues 2-3), 168-17, 2007.

GRTN, Gestore della Rete di Trasmissione Nazionale, www.grtn.it, Italy, 2007.

ISTAT, V Censimento Generale dell'Agricoltura, Roma, Italy, 2000.

Ma J., Scott N.R., DeGloria S.D., Lembo A.J., Siting analysis of farm-based centralized anaerobic digester systems for distributed generation using GIS, *Biomass and Bioenergy*, 28, (Issue 6), 591-600, 2005.

Marrison C.I., Larson E.D., Cost versus scale for advanced plantation-based biomass energy systems in the U.S., Proceedings: The 1995 Symposium on Greenhouse Gas Emissions and Mitigation Research, EPA/600/R-96/072, US Environmental Protection Agency, Washington, DC, Sec. 4:26, 49, June 1996.

NERI (Nat. Environment Res. Inst.), Emission factors for CHP plants < 25 MWe, 2001.

Parawira W., Read J.S., Mattiasson B., Bjornsson L., Energy production from agricultural residues: High methane yields in pilot-scale two-stage anaerobic digestion, *Biomass and Bioenergy*, 32 (Issue 1), Pages 44-50, 2008.

Polimeni A., Ottimizzazione dell'utilizzo del biogas: un'applicazione alla provincia di Cremona, Tesi di Laurea, Politecnico di Milano, Italy, 2007.

Provolo G., Utilizzazione agronomica del digestato alla luce della recente normativa per le zone vulnerabili – Energie rinnovabili in agricoltura: il biogas, 2007.

Spielmann M., Kägi, Stadler P., Tietje O., Life Cycle Iventories of Trasport Services, Swiss Centre for Life Cycle Inventories, 2004.

UTS Italia Tecnologie ambientali, Technical specifications of 330 kWe, 625 kWe, 1025 kWe and 2127 kWe biogas plants, 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

Decision support models for scheduling hydro and thermal power plants in the liberalized electric energy market

M.T. Vespucci [a] **and M. Innorta**[a]

[a]*Department of Information Technology and Mathematical Methods, University of Bergamo, viale Marconi 5, 24044 Dalmine (BG), Italy*
(maria-teresa.vespucci@unibg.it, mario.innorta@fastwebnet.it)

Abstract: We consider the problem of an electric energy producer operating in the liberalized electric energy market, who has to schedule his own production plants aiming at maximizing profit. The power system owned by the producer is assumed to consist of hydro power plants and thermal power plants. Two decision support procedures are developed for the short-term scheduling problem of a price taker power producer and a price maker power producer respectively. The procedures are based on mixed integer linear programming models, where the constraints describe the hydro system, the thermal system and the market, while the objective function represents profit to be maximized. The thermal system is modelled in great detail as it allows start-up and shut-down manouvres in every hour of the planning period, taking into account minimum up-time and down-time constraints as well as ramping constraints. By these procedures the optimal unit commitment of thermal plants and the optimal dispatch of hydro and thermal plants are determined, given either a forecast of the energy price, in the price taker case, or an estimate of the competitors' supply curves, in the price maker case. Numerical results are reported and discussed.

Keywords: electricity market; decision support models; renewable sources.

1 INTRODUCTION

In the last decade the electric power industry has undergone a fundamental transformation from one dominated by a regulated vertically integrated monopoly to an industry where electricity is produced and traded as a commodity. In the liberalized electricity market each power producer, in competition with other producers, aims at maximizing his own profit; production is sold by power producers either directly to consumers, on the basis of bilateral contracts, or by presenting sell bids to the Market Operator for each hour of the following day. Analogously, electricity is purchased by consumers either on the basis of bilateral contracts or by presenting purchase bids to the Market Operator for each hour of the following day. On the basis of the aggregated supply and demand curves, the Market Operator determines the equilibrium point of supply and demand, taking into account transmission system constraints defined by the Transmission System Operator. The electricity price is then a "market clearing price", resulting from the interactions among all market participants. In the previous monopolistic context, production resource scheduling aimed at minimising production costs while satisfying given security standards. In liberalized markets each power producer aims at maximizing his own profit: if the producer is a price-taker, resource scheduling must take into account the price at which electricity may be sold; if the producer is a price maker, resource scheduling decisions must take into account both other producers' decisions and the Market Operator's rules for determining the electricity price. Different time horizons are considered in the production resource scheduling problem. A time horizon of at least one year (medium term) is considered when determining the optimal maintenance plans of hydro and thermal plants and the optimal weekly discharge of seasonal basins. A time horizon of a week or

ten days (short term) is considered for determining the unit commitment of thermal plants, i.e. the start-up and shut-down manouvres of the available (not in maintenance) thermal plants, as well as the production levels of the committed thermal plants and of the available hydro plants in each hour. However, the unit commitment of thermal plants introduces various elements of complexity in the short-term scheduling problem (see Scheble and Fahd [1994] and Sen and Kothari [1998]): thermal generation costs are nonlinear functions of the production level; binary variables need to be used for modeling the state ON/OFF of thermal plants as well as for determining the scheduling of start-up and shut-down manouvres; a large number of constraints is necessary for describing the technical characteristics of thermal plants. Solution procedures based on dynamic programming (Martini et al. [2001]) have been introduced to deal with the high dimensionality of the solution space in hydro-thermal coordination problems. These procedures, however, may not guarantee an optimal solution and may just provide a suitably defined local optimum. In recent years very powerful solvers for Mixed Integer Linear Programming problems have become available, that can compute the optimal solution of instances of very large dimension. Mathematical programming models and methods have proven to be efficient tools for analysing and solving operation scheduling problems (see Read [1996]). This has opened the way to the development of resource scheduling models containing very detailed descriptions of both the generation technologies owned by the producer and the market in which the producer operates (see the review by Ventosa et al. [2005]).

Aim of the paper is the development of decision support procedures, based on Mixed Integer Linear Programming models, for the short-term resource scheduling problem of a profit maximizing power producer. We assume that the power production system consists of hydro plants and thermal plants, but the models can be easily generalized to deal with other generation technologies. Two cases are considered: a price taker producer, who cannot influence the market price and therefore determines his optimal schedule on the basis of price forecasts, and a price maker producer, whose production decisions can influence the market price, which is then a model endogenous variable. Both models determine the unit commitment of thermal plants and the production levels of committed thermal plants and available hydro plants in each hour so as to maximize profits, while satisfying constraints describing the hydro system, the thermal system and the market. The annual scheduling decisions (optimal maintenance plans of hydro and thermal plants, optimal weekly discharge of seasonal basins) are given, as well as forecasts on basin natural inflows. The short-term planning horizon is discretized in hours: T denotes the number of hours considered and t, $0 \leq t \leq T$, is the hour index, with $t = 0$ denoting the last hour of the planning horizon immediately preceeding the one in consideration. The paper is organized as follows. In sections 2 and 3 the model constraints describing the hydroelectric system and the thermal system, respectively, are discussed; in section 4 a linearized model for thermal production costs is briefly addressed; in sections 5 and 6 market constraints and model objective functions are presented for the price taker and the price maker model respectively. In section 7 some numerical results are presented and in section 8 lines for further research are discussed.

2 MODEL OF THE HYDROELECTRIC SYSTEM

The hydroelectric system consists of a number of sets, called *cascades*, of hydraulically interconnected hydro plants, pumped-storage hydro plants and basins. A cascade is mathematically represented by a directed multi-graph: every node represents a basin, with a given storage capacity, and every arc represents either a hydro plant (power generation), or a hydro pump (power storage), or a basin spillage (water flow to a downstream basin for keeping water storage within the storage capacity limit). Let J denote the set of nodes and I denote the set of arcs. The following data are relevant for the description of the hydroelectric system: for $i \in I$ and $j \in J$

- $A_{i,j}$: (i,j)-entry of network arc-node incidence matrix ($A_{i,j} = -1$: arc i leaves node j; $A_{i,j} = 1$: arc i enters node j; $A_{i,j} = 0$: arc i neither leaves nor enters node j)
- k_i $[MWh/10^3m^3]$: energy coefficient ($k_i > 0$, if arc i represents generation; $k_i < 0$, if arc i represents pumping; $k_i = 0$, if arc i represents spillage)

- $\overline{q}_i$ $[10^3 m^3/h]$: maximum water flow in arc i
- ρ_i $[h]$: delay on arc i
- $F_{j,t}$ $[10^3 m^3/h]$: natural inflow in basin j in hour t
- $\overline{v}_j$ $[10^3 m^3]$: maximum storage volume in basin j
- $v_{j,0}$ $[10^3 m^3]$: initial storage volume in basin j
- $\underline{v}_{j,T}$ $[10^3 m^3]$: minimum storage volume in basin j at the end of hour t, determined by the medium-term resource scheduling.

The power producer must schedule the hourly production of each hydro plant, which is expressed as the product of the hydro plant energy coefficient times the turbined volume in hour t, as well as the hourly pumped and spilled volumes. Decision variable $q_{i,t}$ $\left[10^3 m^3/h\right]$ represents the water flow on arc i in hour t (turbined volume, if arc i represents generation, pumped volume, if arc i represents pumping, and spilled volume, if arc i represents spillage), while decision variable $v_{j,t}$ $\left[10^3 m^3\right]$ represents the storage volume in basin j at the end of hour t. The values assigned to the decision variables must satisfy the following constraints that describe the hydroelectric system:

- flow on arc i in hour t is nonnegative and bounded above by the maximum volume that can be either turbined, or pumped, or spilled

$$0 \leq q_{i,t} \leq \overline{q}_i \qquad i \in I, \quad 1 \leq t \leq T \tag{1}$$

- the storage volume in basin j at the end of hour t is nonnegative and bounded above by the maximum storage volume

$$0 \leq v_{j,t} \leq \overline{v}_j \qquad j \in J, \quad 1 \leq t \leq T \tag{2}$$

- at the end of hour T the storage volume in basin j is bounded below by the minimum storage volume required at the end of the current planning period, so as to provide the required initial storage volume at the beginning of the following planning period

$$\underline{v}_{j,T} \leq v_{j,t} \qquad j \in J \tag{3}$$

- the storage volume in basin j at the end of hour t must be equal to the basin storage volume at the end of hour $t-1$ plus the sum of basin inflows in hour $t-\rho_i$ minus the sum of basin outflows in hour $t-\rho_i$

$$v_{j,t} = v_{j,t-1} + F_{j,t} + \sum_{i \in I} A_{i,j} \cdot q_{i,t-\rho_i} \qquad j \in J, \quad 1 \leq t \leq T \tag{4}$$

where $v_{j,0}$ is a data representing the initial storage volume in basin j. Basin inflows are natural inflows, turbine discharge from upstream hydro plants, pumped volumes from downstream hydro plants, spilled volumes from upstream basins. Basin outflows are turbine discharge to downstream hydro plants, pumped volumes to upstream hydro plants and spilled volumes to downstream basins.

3 MODEL OF THE THERMAL SYSTEM

Let K denote the set of thermal plants owned by the power producer. For every plant $k \in K$ the producer must decide the state (ON-OFF) in hour t and whether or not a manouvre (start-up or shut-down) has to take place in hour t, taking into account minimum up-time and minimum

down-time constraints. This problem is known as *unit commitment* problem. The producer must also decide the hourly production of each committed unit, taking into account lower and upper bounds on production levels and ramping constraints. The corresponding decision variables in the model are the real variables $p_{k,t}$ $[MWh]$, that represent the production level of plant k in hour t, and the binary variables, $\alpha_{k,t}$ ($\alpha_{k,t} = 1$ [0]: start-up [no start-up] of plant k in hour t), $\beta_{k,t}$ ($\beta_{k,t} = 1$ [0]: shut-down [no shut-down] of plant k in hour t) and $\gamma_{k,t}$ ($\gamma_{k,t} = 1$ [0]: plant k is ON [OFF] in hour t). The following relations among manouvres in t, state in t and state in $t-1$ must hold:

$$\gamma_{k,t-1} + \alpha_{k,t} = \gamma_{k,t} + \beta_{k,t} \qquad k \in K, \quad 1 \leq t \leq T \tag{5}$$

When a thermal plant is started-up, it must be ON for at least ta_k hours; analogously, when a thermal plant is shut-down, it must be OFF for at least ts_k hours: these restrictions are called *minimum up-time* and *minimum down-time* constraints. Given $\gamma_{k,0}$, the state of plant k at the beginning of the planning period, and nh_k, the number of hours in which plant k has been in state $\gamma_{k,0}$ after the last manouvre in the previous planning period, the minimum up-time constraints are expressed as

$$\sum_{\tau=t+1}^{\min(t+ta_k-1,T)} \gamma_{k,\tau} \geq \alpha_{k,t} \cdot \min\left(ta_k - 1, T - t\right) \qquad k \in K, \quad 1 \leq t \leq T \tag{6}$$

i.e. if a start-up manouvre takes place in hour t, then plant k must be ON either for $ta_k - 1$ subsequent hours, if $ta_k - 1 \leq T - t$, or for $T - t$ subsequent hours, otherwise. Moreover, if plant k was ON in the last hour of the previous planning period, it must be ON for at least the first $ta_k - nh_k$ hours of the current planning period: therefore the value assignment $\gamma_{k,t} = 1$ is made, for $1 \leq t \leq ta_k - nh_k$, if $\gamma_{k,0} = 1$.

The minimum down-time constraints are analogously expressed as

$$\sum_{\tau=t+1}^{\min(t+ts_k-1,T)} \gamma_{k,\tau} \leq (1 - \beta_{k,t}) \cdot \min\left(ts_k - 1, T - t\right) \qquad k \in K, \quad 1 \leq t \leq T \tag{7}$$

i.e. if a shut-down manouvre takes place in hour t, then plant k must be OFF either for $ts_k - 1$ subsequent hours, if $ts_k - 1 \leq T - t$, or for $T - t$ subsequent hours, otherwise. Moreover, if plant k was OFF in the last hour of the previous planning period, it must be OFF for at least the first $ts_k - nh_k$ hours of the current planning period: therefore the value assignment $\gamma_{k,t} = 0$ is made, for $1 \leq t \leq ts_k - nh_k$, if $\gamma_{k,0} = 0$.

The hourly production levels $p_{k,t}$ are subject to the following constraints:

- if plant k is started-up in hour t, the hourly production $p_{k,t}$ cannot be greater than vsu_k $[MWh]$, the maximum production level at start-up; moreover, if the production levels in two subsequent hours $t-1$ and t are such that $p_{k,t-1} \leq p_{k,t}$, the production variation is bounded above by δu_k $[MWh]$, the maximum production increase per hour of plant k (*ramp-up* constraint); these restrictions are imposed by the constraints

  $$p_{k,t} - p_{k,t-1} \leq \delta u_k + \alpha_{k,t} \cdot (vsu_k - \delta u_k) \qquad k \in K, \quad 1 \leq t \leq T \tag{8}$$

  where $p_{k,0}$ $[MWh]$ is the production level of plant k in the last hour of the previous planning period;

- if plant k is shut-down in hour t, the hourly production $p_{k,t}$ cannot be greater than vsd_k $[MWh]$, the maximum production level at shut-down; moreover, if the production levels

in two subsequent hours $t-1$ and t are such that $p_{k,t-1} \geq p_{k,t}$, the production variation is bounded above by δd_k $[MWh]$, the maximum production decrease per hour of plant k (*ramp-down* constraint); these restrictions are imposed by the constraints

$$p_{k,t} - p_{k,t-1} \geq -\delta d_k + \beta_{k,t} \cdot (-vsd_k - \delta d_k) \qquad k \in K, \quad 1 \leq t \leq T \tag{9}$$

- if plant k is ON in hour t, the hourly production $p_{k,t}$ cannot be less than the minimum level $\underline{p}_k$ $[MWh]$ and cannot be greater than the maximum level $\overline{p}_k$ $[MWh]$; if plant k is OFF in hour t, the hourly production must be zero: these restrictions are imposed by the constraints

$$\gamma_{k,t} \cdot \underline{p}_k \leq p_{k,t} \leq \gamma_{k,t} \cdot \overline{p}_k \qquad k \in K, \quad 1 \leq t \leq T. \tag{10}$$

4 THERMAL PRODUCTION COSTS

Two types of costs are associated to thermal production: costs of manouvres and generation costs. For every plant k costs csu_k $[Euro]$ and csd_k $[Euro]$ are associated to every start-up and shut-down manouvre respectively. The thermal generation cost $G_{k,t}$ of plant k in hour t is assumed to be a convex quadratic function of the production level $p_{k,t}$

$$G_{k,t}\left(p_{k,t}\right) = g_{2,k} \cdot p_{k,t}^2 + g_{1,k} \cdot p_{k,t} + g_{0,k} \tag{11}$$

where $g_{2,k}$ $[Euro/MWh^2]$, $g_{1,k}$ $[Euro/MWh]$ and $g_{0,k}$ $[Euro]$ are the quadratic generation cost coefficients for plant k. Because of the unit commitment binary decision variables, the model we develop is of Mixed Integer type: the generation cost functions are therefore linearized, so as to obtain a Mixed Integer Linear Programming model. For every plant k the interval $\left[\underline{p}_k, \overline{p}_k\right]$ is divided in H subintervals of width $\overline{pl}_{k,h}$, $1 \leq h \leq H$. Let $p_{k,t,h-1}$ and $p_{k,t,h}$ denote the extreme points of subinterval h, let $cl_{k,h}$ denote the slope of the straight line segment passing through points $\left(p_{k,t,h-1}, G_{k,t}\left(p_{k,t,h-1}\right)\right)$ and $\left(p_{k,t,h}, G_{k,t}\left(p_{k,t,h}\right)\right)$ and let $pl_{k,t,h}$ denote the real variable associated to subinterval h. The linearized generation costs of thermal plant k in hour t are then given by

$$G_{k,t}^{lin}\left(p_{k,t}\right) = \left(g_{2,k} \cdot \underline{p}_k^2 + g_{1,k} \cdot \underline{p}_k + g_{0,k}\right) \cdot \gamma_{k,t} + \sum_{h=1}^{H} cl_{k,h} \cdot pl_{k,t,h} \tag{12}$$

where variables $pl_{k,t,h}$ are subject to the constraints

$$p_{k,t} = \underline{p}_k \cdot \gamma_{k,t} + \sum_{h=1}^{H} pl_{k,t,h} \qquad k \in K, \quad 1 \leq t \leq T \tag{13}$$

$$0 \leq pl_{k,t,h} \leq \overline{pl}_{k,h} \qquad k \in K, \quad 1 \leq t \leq T, \quad 1 \leq h \leq H \tag{14}$$

See Vespucci et al. [2007] for a detailed description of the linearization procedure. Summarizing, the total costs of thermal plant k in hour t are

$$C_{k,t}\left(p_{k,t}\right) = csu_k \cdot \alpha_{k,t} + csd_k \cdot \beta_{k,t} + G_{k,t}^{lin}\left(p_{k,t}\right) \tag{15}$$

with $G_{k,t}^{lin}\left(p_{k,t}\right)$ given by (12).

5 MARKET CONSTRAINTS AND OBJECTIVE FUNCTION OF THE PRICE TAKER MODEL

In this section we introduce the market constraints and the objective function for a producer who cannot influence the market price. It is assumed that in every hour t of the planning period the Price Taker must satisfy the load car_t that derives from his bilateral contracts. If his total production exceeds the load from bilateral contracts, the Price Taker sells the excess quantity, $sell_t$, on the spot market; if his total production is less than the load from bilateral contracts, the Price Taker must buy on the market the amount of energy, buy_t, necessary to meet the load car_t. Therefore the market constraints are

$$\sum_{i \in I} k_i \cdot q_{i,t} + \sum_{k \in K} p_{k,t} + buy_t - sell_t = car_t \quad 1 \leq t \leq T \tag{16}$$

The objective function representing the Price Taker profits is

$$\sum_{t=1}^{T} \left[\lambda_t \cdot sell_t - \mu_t \cdot buy_t - \sum_{k \in K} C_{k,t}\left(p_{k,t}\right) \right] \tag{17}$$

where λ_t is the market sell price in hour t, μ_t is the market purchase price in hour t and $C_{k,t}\left(p_{k,t}\right)$ are the linearized total thermal costs given by (15). It is supposed that $\mu_t > \lambda_t$, as it also includes transaction costs.

6 MARKET CONSTRAINTS AND OBJECTIVE FUNCTION OF THE PRICE MAKER MODEL

In this section we introduce the market constraints and the objective function for a producer who can influence the market price. It is assumed that in every hour t the Price Maker can estimate the load CAR_t required by the system and therefore, depending on his total production, he can determine the residual demand Π_t, to be satisfied by his competitors. The market constraints are

$$\sum_{i \in I} (k_i \cdot q_{i,t}) + \sum_{k \in K} p_{k,t} + \Pi_t = CAR_t \quad 1 \leq t \leq T \tag{18}$$

It is also assumed that the Price Maker can estimate the sell bids presented in hour t by his competitors. On the basis of the sell bids, the competitors' aggregated offer function is constructed and included in the model for determining the hourly energy price, which is therefore an endogenous variable. Sell bids are pairs $(QV_{s,t}, PV_{s,t})$, $1 \leq s \leq S_t$, where S_t denotes the number of sell bids presented by the competitors in hour t, $QV_{s,t}$ denotes the maximum energy quantity offered and $PV_{s,t}$ denotes the minimum price at which the producer is willing to sell quantity $QV_{s,t}$. Sell bids have been reordered in merit order, i.e. $PV_{s,t} \leq PV_{s+1,t}, 1 \leq s \leq S_t - 1$. The competitors' aggregated offer function in hour t is a step-wise linear function, with S_t horizontal segments and $S_t - 1$ vertical segments, where the vertical segment s has zero length if $PV_{s,t} = PV_{s+1,t}$. The mathematical representation of the competitors' aggregated offer function in hour t is obtained by defining the points $(aqv_{p,t}, apv_{p,t})$, with $1 \leq p \leq P_t = 2 \cdot S_t$, in the plane (Q, P): for $1 \leq s \leq S_t$

if $p = 2s - 1$ $\quad apv_{p,t} = PV_{s,t}$ and $aqv_{p,t} = \sum_{\sigma=1}^{s-1} QV_{\sigma,t}$ (with $aqv_{1,t} = 0$)

if $p = 2s$ $\quad apv_{p,t} = PV_{s,t}$ and $aqv_{p,t} = \sum_{\sigma=1}^{s} QV_{\sigma,t}$.

Therefore any point $(\Pi_t, Price_t)$ of the competitors' aggregated offer function can be represented as the convex linear combination of the points $(aqv_{p,t}, apv_{p,t})$, $1 \leq p \leq P_t$, where the combination coefficients are the decision variables $\eta_{p,t}$, subject to the following constraints:

- the set of variables $\eta_{p,t}$, $1 \leq p \leq P_t$, is a special ordered set of type 2 (SOS2), see Beale and Tomlin [1970], i.e. an ordered set within which at most two adjacent variables can be nonzero
- variables $\eta_{p,t}$ are subject to the convexity constraints

$$0 \leq \eta_{p,t} \leq 1 \qquad 1 \leq p \leq P_t, \quad 1 \leq t \leq T \tag{19}$$

$$\sum_{p=1}^{P_t} \eta_{p,t} = 1 \qquad 1 \leq t \leq T. \tag{20}$$

The Price Maker revenues in hour t are then modelled as a function of the endogenously determined hourly price $Price_t$, i.e. the P-coordinate of the intersection point between the competitors' aggregated offer function and the vertical line of Q-coordinate Π_t, that represents the competitors' residual demand function. Indeed this will be the position of the market equilibrium point if the Price Maker offers his own production at a price less than the optimal $Price_t$ determined by the model. It is easily shown that the Price Maker revenues in hour t are

$$Price_t \cdot (CAR_t - \Pi_t) = \sum_{p=1}^{P_t} apv_{p,t} \cdot (CAR_t - aqv_{p,t}) \cdot \eta_{p,t} \tag{21}$$

Indeed let $\eta_{\hat{p},t}$ and $\eta_{\hat{p}+1,t}$, $1 \leq \hat{p} \leq P_t - 1$, be the nonzero variables in the SOS2 set of variables $\eta_{p,t}$, $1 \leq p \leq P_t$, at the optimal solution:

- if $\hat{p}$ is odd, the intersection point belongs to the $\hat{p}$-th segment of the competitors' aggregated offer function, which is horizontal: relation (21) follows from the equalities $Price_t = apv_{\hat{p},t} = apv_{\hat{p}+1,t}$ and $CAR_t - \Pi_t = (CAR_t - aqv_{\hat{p},t}) \cdot \eta_{\hat{p},t} + (CAR_t - aqv_{\hat{p}+1,t}) \cdot \eta_{\hat{p}+1,t}$, where the last equality holds since $\Pi_t = aqv_{\hat{p},t} \cdot \eta_{\hat{p},t} + aqv_{\hat{p}+1,t} \cdot \eta_{\hat{p}+1,t}$ and $\eta_{\hat{p},t} + \eta_{\hat{p}+1,t} = 1$;
- if $\hat{p}$ is even, the intersection point belongs to the $\hat{p}$-th segment of the competitors' aggregated offer function, which is vertical: relation (21) follows from the equalities $CAR_t - \Pi_t = CAR_t - aqv_{\hat{p},t} = CAR_t - aqv_{\hat{p}+1,t}$ and $Price_t = apv_{\hat{p},t} \cdot \eta_{\hat{p},t} + apv_{\hat{p}+1,t} \cdot \eta_{\hat{p}+1,t}$.

Then the objective function representing the Price Maker profits is

$$\sum_{t=1}^{T} \left[\sum_{p=1}^{P_t} apv_{p,t} \cdot (CAR_t - aqv_{p,t}) \cdot \eta_{p,t} - \sum_{k \in K} C_{k,t}\left(p_{k,t}\right) \right]. \tag{22}$$

7 MODEL VALIDATION

The two models have been validated on sets of data relative to Italian power producers. We report some examples of the numerical results obtained. In Figure 1 the graph on the left shows the hourly price forecast on the basis of which a Price Taker producer, with thermal capacity of 22945 MWh and hydro capacity of 11937 MWh, has to schedule his own plants. The profit maximizing solution is then computed and the total hydro and thermal productions in each hour are represented in the graph on the right. It can be observed that both the hydro and the thermal production curves tend to follow the price curve: in periods of high hourly price the thermal production reaches a constant value and only the hydro production curve follows the price curve; in periods of low hourly price, negative values of the hydro production indicates that in the optimal resource scheduling computed by the model some hydro plants are used for pumping.

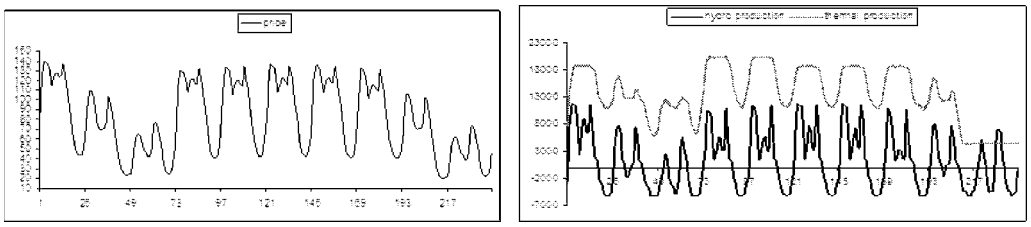


Figure 1: hourly energy prices (left) and hydro-thermal productions (right) along 240 hours

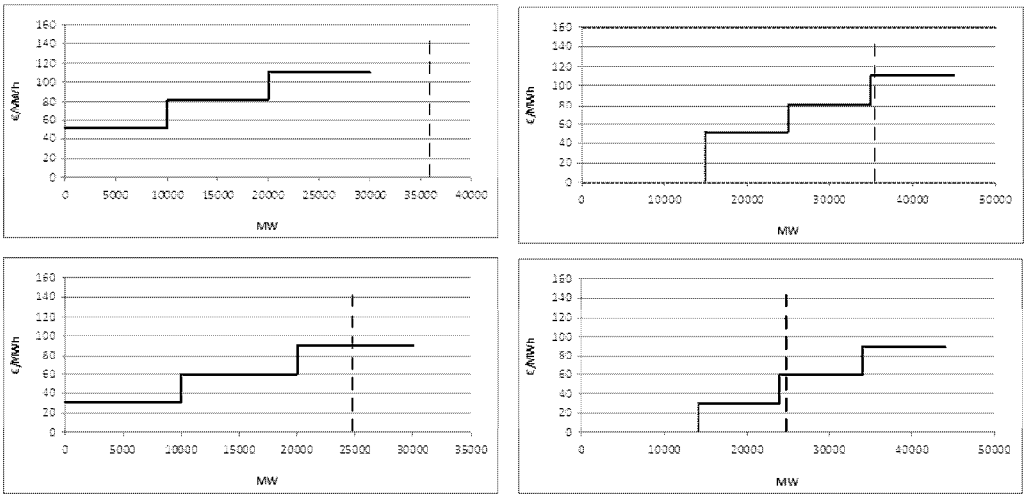


Figure 2: data (top left) and optimal solution (top right) for Price Maker indispensable in hour t, data (bottom left) and optimal solution (bottom right) for Price Maker not indispensable in hour t

In figure 2 the optimal results obtained by the short-term Price Maker model are discussed by focusing on two typical hours: a hour in which the Price Maker producer is indispensable to satisfy the system load (top graphs) and a hour in which the Price Maker producer is not indispensable to satisfy the system load (bottom graphs). In the top left graph the system load, represented by the vertical dashed line, is supposed to be rigid and to amount to 35970 MWh, while the competitors' aggregated offer function is supposed to consist of three steps. The top right graph shows the details, for the hour in consideration, of the optimal solution computed by the model: in order to maximize his own profit, the Price Maker must produce 15960 MWh, so that the shifted competitors' offer function intersects the load curve just at the beginning of the third step: the Price Maker production share determined for this hour by the optimization model is the maximum compatible with forcing the hourly market price to be equal to the highest bid price. This solution gives the Price Maker a profit maximizing option alternative to bidding at price cap, which would be only exceptionally accepted by the system. In a hour in which the system load is supposed to amount to 24740 MWh (bottom left graph), the Price Maker optimal production is 14000 MWh (bottom right graph), so that the shifted competitors' offer function intersects the load curve at the beginning of the second step: for the Price Maker it is more profitable maximizing the production share rather than forcing the highest hourly price.

8 CONCLUSIONS AND FUTURE WORK

Power production from reneawble sources has been recently becoming more and more relevant and it is important to study the effects of integrating conventional and renewable power production technologies. For instance, wind power generation is partially unpredictable and the amount of wind power generated greatly depends on the inherent variability of wind: by integrating wind

generation and hydropower with pumped storage, intermittency in wind power generation can be efficiently managed, as wind power generation exceeding demand can be used for pumping. The authors are currently developing scheduling models for a production system including hydro with pumped storage, thermal and wind power plants: a stochastic medium-term model schedules hydro with pumped storage, taking into account uncertainty of natural inflows in reservoirs and representing wind power by hourly average values; a deterministic short-term model determines the optimal unit commitment of thermal plants, with wind power generation represented by hourly average values and optimal weekly discharge defined by the medium-term model; a stochastic daily scheduling model determines the hourly production of hydro and thermal plants, given the unit commitment of thermal plants and taking into account wind stochasticity.

ACKNOWLEDGMENTS

This work was supported by MIUR PRIN 2005 "*Modelli per la gestione del mercato elettrico italiano e per il processo di bidding*".

REFERENCES

Beale, E. and J. Tomlin. Special facilities in a general mathematical programming system for nonconvex problems using ordered sets of variables. In *Proceedings of the Fifth International Conference on Operation Research*, pages 447–454, 1970.

Martini, A., P. Pelacchi, L. Pellegrini, M. Cazzol, A. Garzillo, and M. Innorta. A simulation tool for short term electricity markets. In *Proceedings of the 22nd IEEE Power Engineering Society International Conference on Power Industry Computer Applications 2001*, pages 112–117, 2001.

Read, E. OR modelling for a deregulated electricity sector. *International Transactions in Operational Research*, 3(2):129–137, 1996.

Scheble, G. and G. Fahd. Unit commitment literature synopsis. *IEEE Transactions on Power Systems*, 9(1):128–135, 1994.

Sen, S. and D. Kothari. Optimal thermal generating unit commitment: a review. *International Journal of Electrical Power and Energy Systems*, 20(7):443–451, 1998.

Ventosa, M., A. Baillo, A. Ramos, and M. Rivier. Electricity market modeling trends. *Energy Policy*, 33(7):897–913, 2005.

Vespucci, M., M. Innorta, E. Allevi, and M. Bertocchi. Models for short-term scheduling of hydro-thermal resources in a liberalized electric energy market. Technical Report 3, DMSIA University of Bergamo, 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

GIS based Model to optimize possible self sustaining regions in the context of a renewable energy supply

M. Biberacher[a], **S. Gadocha**[a], **D. Zocher**[a]

[a] *Austrian Research Centers GmbH – Research Studios, 5020 Salzburg, AUSTRIA*
(markus.biberacher@researchstudio.at)

Abstract:

During the last years an increasing energy demand, rising prices for fossil fuels, the challenge of meeting the objectives of the Kyoto protocol as well as a certain uncertainty of energy supply resulted in the two main aspects arising within the design of our future energy system, which are sustainability and security of supply.

To meet these challenges, the paper presents a modelling approach that handles information on geographically disaggregated data of renewable energy potentials as well as geographically disaggregated information on energy demand structures. The comparison of the identified energy potentials of the modelling process to the relative energy consumption structure results in a "balance grid" that represents the energy excess or shortage in every cell of the grid. The balance grid is the basis for modelling self-sustaining regions and allows a differentiated geographical consideration of energy production and consumption potentials.

Processing this information the model approach identifies optimized energy flows to balance all energy demand hot spots. This is applied for a special region of interest with the objective of finding one optimized setup for the whole prospected area. The final outcome of the model shows an ideally balanced energy flow structure for the whole examined region. In its simplest realization the energy flows only consider balanced flows for a full year timescale. Nevertheless these flows could also be treated on an arbitrary different timescale.

Based on these outcomes a possible sub-regionalisation in terms of energetic independency within the considered region of interest can be identified. This is reflected by clustering the region of interest into single self sustaining sub regions.

The model itself is a linear optimization model realised in the modelling language GAMS. There is an interface implemented to connect the model to common GIS software. In the current model all input and result data are administrated and visualised in ArcGIS.

Keywords: Renewable energy, regional energy systems, modelling

1. INTRODUCTION

A secure, efficient and environmentally conscious energy-supply is essential for a sustainable provision of goods and services. In the context of its international complexity the global energy industry has to cope with great challenges these days. This refers to aspects like the constantly increasing energy demand, insufficient energy conversion and transport capacities or geopolitical risks alongside others. Furthermore the possible long-time effects of CO_2 emissions in relation to global warming and the challenge to meet the

obligations of the Kyoto protocol have lead to an enhanced problem awareness regarding energy supply systems. Hence there are two keywords rising in decision making processes regarding our future energy system: Sustainability and Security of Supply.

Both aspects are of major interest and need to be treated carefully. Since an improvement of the current energy system towards sustainability and security of supply is also particularly determined by spatial questions, attention has to be paid to this spatial aspect when a modelling process of a possible future energy system is carried out. This namely refers to the spatial distribution of renewable energy carriers and their possible utilization in the energy system. The problem faced in this context is the generally low energy density of renewable energy carriers. Therefore it is of major interest –especially in terms of "security of supply" – to pay attention to the geographical deviation of renewable energy supply and energy demand. To reduce the risks of energy dependency from politically instable countries it is necessary to meet the demand at the smallest possible scale with local and regional energy sources, which brings geographical information in the focus [Tegou et al., 2007].

2. SPATIALLY DISAGGREGATED ENERGY BALANCE

The estimation of energy resource potentials and energy demand in a spatially high resolution, which are based on geographical methods and data, allow a discrete valorisation in the modelling process. Especially the treatment of renewable energy carriers with their relatively low energy density and small-scale variance in supply require a spatially high disaggregated modelling of the energy flows. Until now energy potentials and demand were mostly included into modelling schemes in a cumulative way.

2. 1 Energy resources modelling

An adequate method for modelling renewable energy source (RES) potentials is presented by a top-down approach. GIS is especially useful in the RES modeling, which is also determined by the special geographical qualities of RES [Dominguez et al., 2007; Voivontas, 1998]. In a first step universally valid fundamentals are used to calculate the theoretical potentials. The estimated theoretical potentials are then reduced to a technical potential by including technical limitations taking into account the state-of-the-art as well as factors concerning natural space. By using rather soft factors which may be modified over time and may vary regionally the potential can be further reduced to a realisable one (Figure 1).

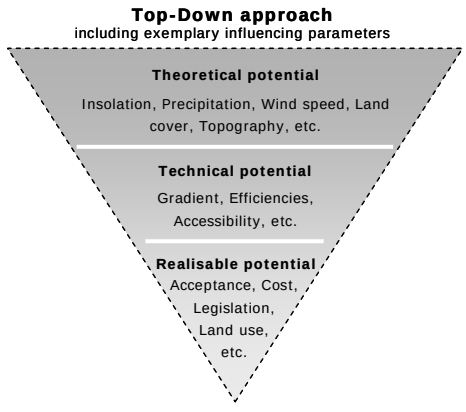


Figure 1: Top-down approach for the estimation of renewable energy potentials

The estimation of the theoretical solar potential is based on topography and global radiation data for the region of interest [Ramachandra and Shruthi, 2007]. The available solar irradiation was calibrated by "solar radiation models" including the following assumptions for solar collector orientation: South 180°, gradient 39°, latitude of the study area. Additionally it was assumed that all roofs can be used for the installation of solar collectors. For the estimation of the technical potential an average efficiency factor for solar collectors is used.

The biomass potential, including woody (forestry plantations, natural forests and scrap wood) , non woody (grassland, cropland), liquid manure and organic waste, was estimated based mainly on land cover and land use data as well as annual

biomass growth rates for the respective biomass types and official statistics for organic waste. The average annual growth factors are based on Mittlböck [2006]. As not the total amount of annual increment can be used for energetic purposes due to topography, competitive use, sustainability and ecological factors, the theoretical potential has to be reduced [Smeets et al., 2007] by a predefined factor . It is assumed that 32% of the total forest potential can be used for energetic purposes. For agricultural biomass it is assumed that 25% of cropland can be used for the production of energy. Regarding grassland the part of fodder has to be excluded from energetic use, but a part is returned in form of liquid manure for biogas production.

For the modelling process of renewable energy potentials comparable spatial units have to be defined, which are the reference objects for the potential analysis. The modelling process is started by using rather small units, e.g. a geographical raster of a few hundred square meters. In a further step the results may be summed up to any larger spatial unit, to administrative units for instance [Mittlböck, 2006].

2.2 Energy consumption modelling

Besides energy resources also energy demand is generally assigned to specific locations. Seeing that energy flows, especially in the case of low dense energy carriers play an important role not only on a global scale but also on a regional scale distances between supply and demand are of major interest. Hence it is essential to model the energy consumption on the same geographical resolution as the energy potentials to ensure their comparability. For the estimation of the heat and electricity demand characteristic values of demand structures are used, which are then assigned and located to the requested spatial resolution. The data sources for the calculation of the energy demand are generally provided by public authorities or estimations are made in existing literature. In many cases they are only made available for larger administration units. If smaller units have been chosen for the modelling process the information on the energy demand has to be broken down into the appropriate spatial units. This process is called disaggregation. To accomplish the disaggregation of the information, established factors such as settlement areas or buildings are used. The resulting allocation is based on probabilities not on exact census information which is a matter of privacy protection. .Specific data on energy demand [Schlomann et al., 2004] and statistic data of households in the specific region are used for the estimation of energy demand. By joining these data the spatial distribution of the energy demand can be identified.

2.3 Energy balance

As the renewable energy potentials as well as the energy demand have been computed referring to the same spatial units they can be compared in a balancing way. Thereby it can be shown for each defined spatial unit if and to what extent the energy demand can be satisfied by the renewable energy potential of the spatial unit [Mittlböck, 2006].

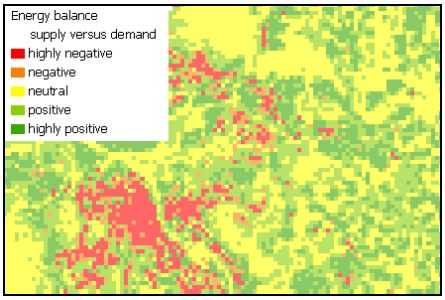
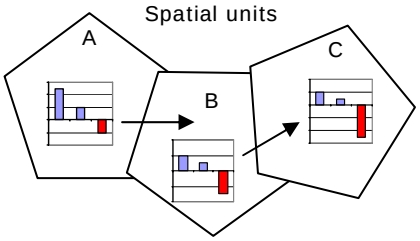


Figure 2: Energy balance – potential renewable energy sources versus given energy demand

The resulting energy balance does not state the realistic utilisable energy potential to satisfy the energy demand of a respective spatial unit. Nevertheless it provides the possibility to identify regions with a surplus energy potential or a demand excess (see Figure 2). This highly disaggregated energy balance is subject to a dynamic modelling approach. Different levels of valorisation for different energy potentials can be investigated and the resulting energy balances can be outlined and discussed.

3. MODELLING OF ENERGY FLOWS

As energy resources as well as the energy demand are commonly not distributed spatially equal, energy flows are necessary to satisfy all demands in a defined region. Of course these flows are dependent on the considered energy carriers but in general a flow can be determined by a starting unit A and a neighbouring unit B. The considered approach deals with arbitrary chosen spatial units for which energy potentials and demands are evaluated. In a real solution all demands need to be satisfied – either by energy potentials nearby or by energy flows from bordering spatial units, which is shown exemplarily in figure 3 [Mittlböck, 2006, Biberacher, 2007]. In this study only balancing energy flows are considered, energy flows based on a process chain are not included in the modelling approach developed so far.



	Before			After			
Spatial Unit	Potential	Demand	Balance	Potential	Demand	Energy Flow	Balance
A	S: 5 B: 2	2	+5	S: 5 B: 2	2	-2	+3
B	S: 2 B: 1	3	0	S: 5 B: 2	3	0	0
C	S: 2 B: 1	5	-3	S: 5 B: 2	5	+2	0

Figure 3: Scheme of Solar (S) & Biomass (B) potentials, demand and energy balance of spatial units with resulting balancing energy flows in arbitrary units.

The evaluation of balancing energy flows is influenced by numerous aspects. First and foremost it is influenced by the energy carrier itself. While electricity can be transported over large distances without suffering high losses, the technical efficiency and economic viability limits are reached quite fast in the case of heat transport [Söderman and Petterson, 2006]. Besides the energy carrier itself also natural conditions (topography, climate, etc.) and existing infrastructure as well as political aspects have an influence – either positive or negative - on the possible exchange of energy between spatial units.

Based on these aspects, energy flows between spatial units can be weighted. This weighting process is implemented as follows:

A norm weight is assigned to each single energy flow between neighbouring spatial units. This weight can be increased or decreased by the aspects already mentioned above. While natural barriers as well as political borders could have a negative influence on the energy transport, existing infrastructure like roads and electrical grids would support the transport of energy.

Figure 4 outlines how these aspects could look like for a certain region of interest. For this region the single spatial units (e.g. raster cells) are analysed whether they are connected to a certain infrastructure (e.g. roads), represent a certain borderline or represent a region with certain natural conditions.

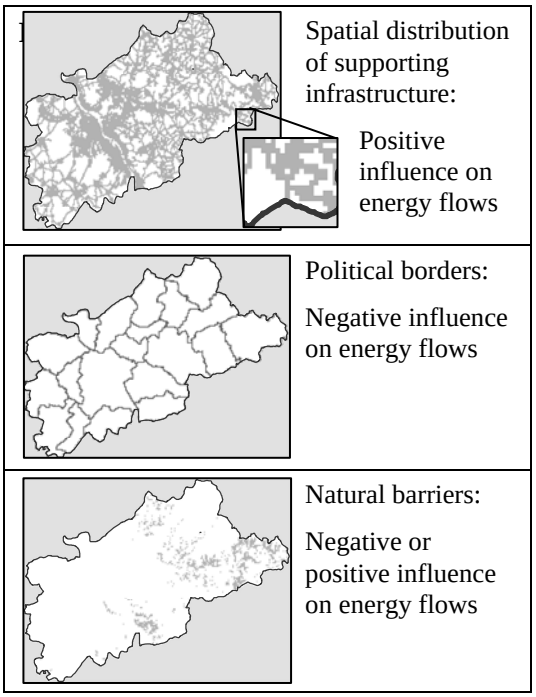


Figure 4: Influences on energy flow related costs depending on spatial factors.

This listing does not claim to be sufficient in terms of describing reality. But it can represent the range of differing aspects that have to be considered in terms of weighting energy flows between single spatial units.

However this example should provide a suitable insight in the idea of weighting energy flows.

After identifying different weighting categories these categories have to be quantified due to their influence on the energy transport. One possibility could be the interpretation of the single weighting categories as fictive costs. Based on this interpretation the single flows become quantifiable regarding their suitability (see equation 2 in section 4).

4. SPATIAL REGIONALISATION

The interpretation of energy flows by assigned fictive costs enables the following modelling approach:

Assumption 1: The energy demand (d_i) in each single spatial unit has to be satisfied by the sum of incoming energy streams (in_x) from neighbouring spatial units (n_i), added to the already available energy potential within this spatial unit (p_i), minus the sum of outgoing energy streams (out_x) to neighbouring spatial units (n_i).

$$d_i \leq p_i + \sum_{x=0}^{n_i} in_x - \sum_{x=0}^{n_i} out_x \quad \text{for all } i = 0 \,..\, u; \text{ while } u = \text{number of units;} \tag{1}$$

Assumption 2: The arising cumulated fictive costs (C) from all single energy flows ($flow_i$) between spatial units (each flow is identified once as *in*-flow and once as *out*-flow) should be minimal for the considered region. Each single energy flow participates with its individual specific cost assignment (k_i).

Assumption 3: The individual specific costs (k_i) are identified by a fixed basic cost (c) plus the sum of distinguished specific influences factors (related to figure 4, number of factors is determined by r) interpreted as increasing or decreasing cost assignments (a_{ix}).

$$\text{Minimise} \left(C = \sum_{i=0}^{f} (k_i \cdot flow_i) \right) \quad \text{while } f = \text{number of individual flows;} \tag{2}$$

$$k_i = c + \sum_{x=0}^{r} a_{ix} \quad \text{for all } i = 0 \,..\, u; \text{ while } u = \text{number of units;} \tag{3}$$

Therefore the whole model can be described as a linear optimisation problem. This optimisation problem was formulated in the algebraic modelling language *GAMS* – General Algebraic Modelling Systems [GAMS, 2007] and solved with the *cplex* linear optimisation solver developed by ILOG [ILOG, 2007].

The output of an optimisation process shows the arising energy flows and their size. In case of single spatial raster units figure 5 shows a possible solution.

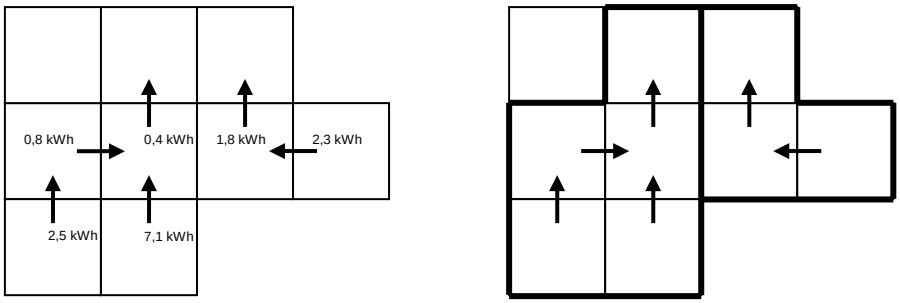


Figure 5: Calculated energy flows between neighbouring spatial units.

The solution shown in figure 5 outlines which energy flows are necessary at the minimum (in terms of the cumulated amount of the renewable energy potentials) to satisfy all arising energy needs of the considered region regarding balancing energy flows. Based on this flow chart, borders can be identified, that are not part of an energy exchange flow. That finally leads to enclosed regions, which are - in terms of energy flows - decoupled islands within the whole considered region.

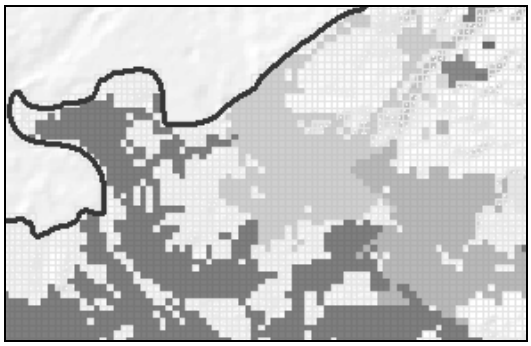


Figure 6: Optimised self sustaining energy regions (individual grey scales).

Ultimately this approach leads to a clustering of single spatial units as shown in figure 6. These clusters represent spatial areas, which can be identified as regions with a positive energy balance - related to all cumulated potentials and needs within the cluster (for clustering techniques see also Murray and Estivill-Castro, 1998). Spatial units that are not connected to a cluster of at least two single units, represent a positive energy balance by their own that is neither influenced by in-flows from neighbouring units nor by out-flows to neighbouring units.

5. SENSITIVITY ANALYSIS

The resulting clusters could be interpreted as possible self sustaining regions in terms of renewable energy supply. The distinctiveness of this approach to calculate possible self sustaining regions is that single clusters are not calculated individually. That means that the surplus of energy is not optimized for each single cluster but for the whole region of interest. Especially in the case of concentrated demand hot spots (like cities) within the examined region a huge neighbourhood will be affected by providing the counterbalance to

satisfy all needs. That also influences areas which would be self sustaining if they are considered individually.

The forming of energy clusters by the indicated model is shown in figure 7. The three scenarios are based on the different valorisation of renewable energy carriers to satisfy the heat demand of each spatial unit. Although the energy balance grid does not seem to be very different in these scenarios there is a noticeable effect on the cluster building process resulting from varying energy flows between neighbouring units

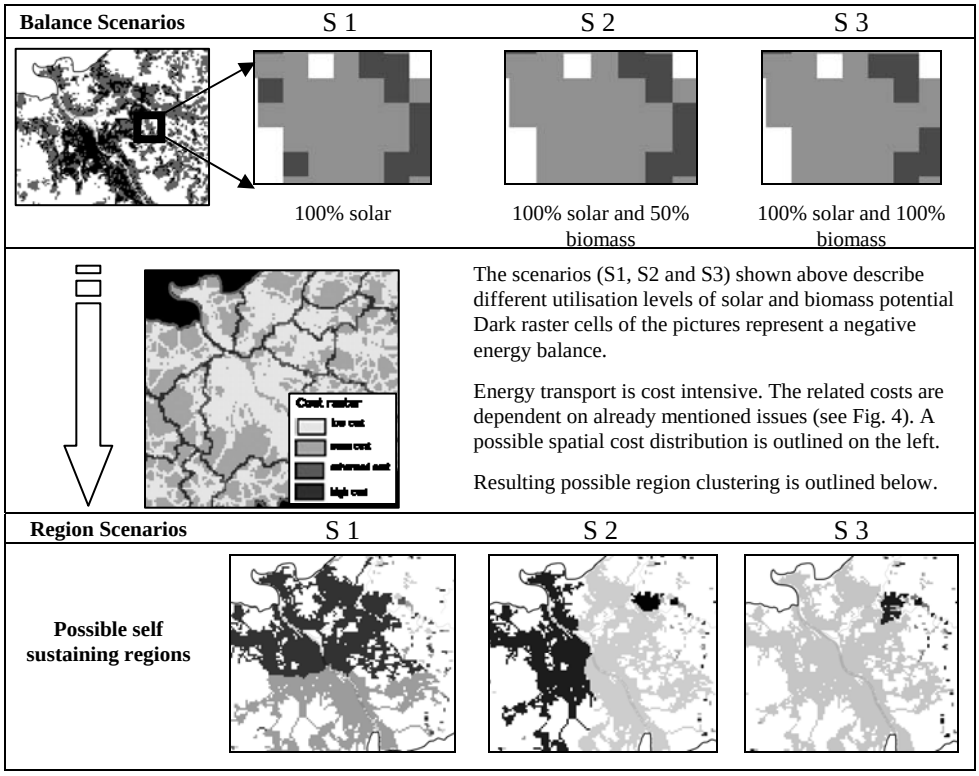


Figure 7: Different scenarios regarding the valorisation of renewable energy carriers to satisfy the heat demand lead to different optimised self sustaining energy regions (identified by individual grey scales).

In general this approach provides the possibility of sensitivity case studies regarding different input settings concerning energy needs, suitable potentials, infrastructure, etc. by this individual single location based studies as well as the "big picture" overview of an entire region of interest are enabled.

6. CONCLUSION

The outlined approach combines a spatially high disaggregated insight in energy resource-, supply- and demand structures, while other current studies work with the cumulative inclusion of energy potentials and demand.

As the approach is not fixed on one spatial scale global as well as regional or sub-regional questions on possible energy flows could be quantified. Spatial units can be chosen individually and influencing parameters on energy flows can be treated.

Hence the modelling approach supports a decision making process on different spatial scales with quantifiable numbers and system setups. Furthermore the integrated model workflow enables a sensitivity analysis. Variations in the input assumptions and their consequences on the results can be treated in a quite flexible way.

By visualising energy flows and the resulting possible energy clusters, cooperation projects between affected administrative units on the use of renewable energies could be motivated. To give a full picture of the possibilities regarding the use of renewable energy to satisfy

the heat and electricity demand also a detailed registration of already utilised renewable energy resources, especially regarding energy flows between neighbouring cells of the balance grid is necessary.

Within the discussed energy flow model a quite simple approach regarding the timescale of the balanced energy flows between the spatial units of the region of interest was used. As the energy demand as well as the supply with renewable energy carriers is not temporarily constant e.g. solar insolation the energy flows should also be treated in a flexible timescale. That will be part of future investigations.

Within the part of energy flow modelling future investigations will also regard the inclusion of more detailed factors of cost estimation, as well as for the different renewable energy carriers. Until now only balancing energy flows to reach a balanced energy balance for a whole region of interest has been included in the considerations. Future developments will also include energy flows based on process chains.

ACKNOWLEDGEMENT

The authors would like to thank the Austrian Ministry of Innovation and Technology for the financial support of the project 808565.

REFERENCES

Biberacher, M., Modelling and optimisation of future energy systems using spatial and temporal methods, VDM Verlag, ISBN 978-3836424356, Germany, 2007

Dominguez, J., Amador, J., Geographical information systems applied in the field of renewable energy resources, *Computers and Industrial Engineering*, 52(3), 322 – 326, 2007

Estivill-Castro, V. and Murray, A.T., Cluster discovery techniques for exploratory spatial data analysis; *International Journal of Geographical Information Science*, 12(5), 431 - 443, 1998

GAMS, General algebraic modeling system, http://www.gams.com/, (online 18.5.2007)

ILOG, http://www.ilog.com/products/cplex/, (online 2.3.2007)

Mittlböck, M. et. al., Virtuelle Kraftwerke für Autarke Regionen, BMvit, Hrsg., Berichte aus Energie- und Umweltforschung 58/2006, Austria, 2007.

Ramachandra, T.V., Shruthi, B.V., Spatial mapping of renewable energy potential; *Renewable and Sustainable Energy Reviews*, Volume 11(7), 1460–1480, 2007

Schlomann, B.; Gruber, E.; Eichhammer, W.; Diekmann, J.; Ziesing, H.-J.; Rieke, H.; Wittke, F.; Herzog, T; Barbosa, M.; Lutz, S.; Broeske, U.; Merten, D.; Falkenberg, D.; Nill, M.; Kaltschmitt, M.; Geiger, B.; Kleeberger, H.; Eckl, R., Energieverbrauch der privaten Haushalte und des Sektors Gewerbe, Handel, Dienstleistungen (GHD), Abschlussbericht and das Bundesministerium für Wirtschaft und Arbeit, Karlsruhe: Fraunhofer ISI, 2004

Smeets, E., Faaij, A., Lewandowski, I., Turkenburg, W., A bottom-up assessment and review of global bio-energy potentials to 2050, *Progress in Energy and Combutsion Science*, 33(1), 56–106, 2007

Söderman, J., Pettersoon, F., Structural and operational optimisation of distributed energy systems, *Applied Thermal Engineering* 26(13) 1400-1408, 2006

Tegou, Leda-Ioanna, Polatidis, Heracles, Haralambopoulos, Dias A., Distributed Generation with Renewable Energy Systems: The spatial dimension for an autonomous Grid, Paper presented at the 47th conference of the European Regional Science Association "Local governance and sustainable development, Thematic stream M: Environment, natural resources and sustainability", ERSA 2007, Paris, France, August 29th – September 2nd 2007

Voivontas, D., Assimacopoulos, D., Mourelatos, A. and Corominas, J.; Evaluation of Renewable Energy potential using a GIS decision support system; *Renewable Energy*, 13(3), 333-344, 1998

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 Benchmark Simulation Models – A Valuable Collection of Modelling Tools

John B. Copp[1], Ulf Jeppsson[2] and Peter A. Vanrolleghem[3]
[1]Primodal, Inc., 122 Leland Street, Hamilton, Ontario L8S 3A4, Canada (copp@primodal.com)
[2]IEA, Lund University, Box 118, SE-221 00 Lund, Sweden (ulf.jeppsson@iea.lth.se)
[3] modelEAU, Département génie civil, Université Laval, Québec, QC, G1K 7P4, Canada (peter.vanrolleghem@gci.ulaval.ca)

Abstract: Over a decade ago, the concept of a tool that could be used to objectively evaluate the performance of control strategies through simulation using a standard model implementation was introduced for activated sludge wastewater treatment plants. That concept resulted in the development of the Benchmark Simulation Model No 1 (BSM1), the subsequent BSM1_LT and most recently BSM2. Debate about the need and application of these models has dogged the development effort since it first began with practitioners suggesting that these models are only academically applicable, have been conceived of for publication generation purposes and provide limited benefit to the applied modelling community. The authors of this paper, as contributing members to the development, beg to differ with those detractors. The focus of this submission is the BSM models from the perspective of a modelling toolbox, and a platform, on which modelling issues have been debated, experimented upon, tested and developed to further the field of wastewater treatment modelling in general.

***Keywords**:* benchmark, BSM, modelling, activated sludge, anaerobic digestion

BACKGROUND

The Benchmark Simulation Models (BSMs) have been under development for many years through a cooperative effort involving research and corporate entities from around the globe. The initial reasoning for the development of these models was to create an unbiased tool that could be used to evaluate wastewater treatment control strategies (Spanjers et al., 1998). At that time, the literature contained many published control concepts, but the methodology used to test or examine the strategy impact in each case was specific to that control strategy. That is, these publications tended to focus on the specific advantages of the particular strategy in question without necessarily highlighting some of the adverse or spin-off effects. Because the strategy impacts were not fully reported, the comparison of different published strategies was almost impossible. It was believed at that time that a simulation-based tool would provide a means to evaluate the relative merits of all kinds of dissimilar control ideas taking into account all the effects that the strategy might have on the treatment process. The development effort has been on-going ever since.

Numerous papers have been published on the various complete benchmark models and these have been presented elsewhere (Copp, 2002; Copp *et al.*, 2002; Rosen *et al.*, 2004; Jeppsson *et al.*, 2006). As there is insufficient space in this paper to fully explore all of the tools that will be highlighted here, the reader is referred to these publications for more details. Publications by researchers, operators and consultants have all illustrated the use of the benchmark systems for the assessment of process performance and control system evaluation. However, the benchmark effort is not without its critics. The unit process sizes and model choices, influent characterisations, model transformations and evaluation criteria have all been criticised. Some of the criticism is justified as the defined BSMs have not

always taken into account the most recent advancements or accepted theory, but the critics fail to fully appreciate the benefits generated by the effort, the debates, the compromises and the solutions that have gotten the BSMs to this point. The BSMs are not all-encompassing tools to be used only as fully defined nor are they 'best-practice' tools to be interpreted as showcasing the best models for specific unit processes. To limit the BSM application this way would be a shame. Rather, these models should be considered as collections of modelling tools that address various aspects of whole-plant wastewater treatment modelling. The value of these modelling tools is much greater than the value of the BSMs as fully defined and the modular nature of the tools means that they can be used in isolation if the need arises.

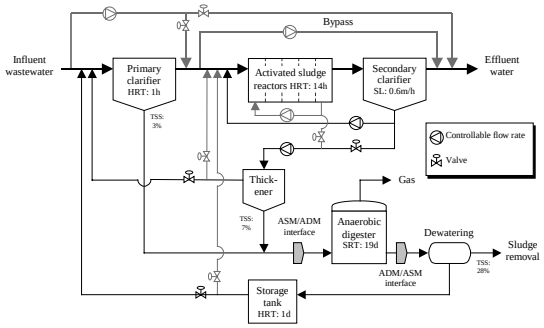


Figure 1: Plant layout for BSM2.

The BSMs contain a whole series of these modelling tools including but not limited to: an influent wastewater generating model, a temperature model, standard and ring-tested implementations of activated sludge model #1 (ASM1), anaerobic digestion model #1 (ADM1), the Takács double exponential settling model and the Otterpohl/Freund primary clarification model, anaerobic digestion/activated sludge model interfaces, empirical solids/liquid separation models, performance indices, operational cost indices as well as models for sensors and actuators and for energy consumption by aeration and pumping equipment. These modelling tools have all come as a direct result of the BSM development. Without the BSM platform, a collaborative development effort on these tools might not have occurred. This submission focuses on these tools with an aim to demonstrate the value of the BSMs as a comprehensive modelling toolbox.

BENCHMARK SIMULATION MODEL #2

The Benchmark Simulation Model No 2 consists of a model representing a general WWTP, an associated control system, a benchmarking procedure and a set of evaluation criteria. The main components of BSM2 (see also Figure 1) are: primary clarification (based on Otterpohl and Freund, 1992) and Otterpohl *et al.*, 1994)); a five-reactor nitrogen removal activated sludge system (based on Henze *et al.*, 1987); secondary clarification (based on Takács *et al.*, 1991); gravity thickening; anaerobic digestion (based on Batstone *et al.*, 2002); dewatering; AD/AS model interfaces (based on Nopens *et al.*, 2008); a storage tank; and an influent wastewater generator model (based on Gernaey *et al.*, 2005; 2006).

BSM MODELLING TOOLS

Influent Wastewater Generating Model

The evaluation of control strategies in BSM1 is done based on three different 1-week long 'weather files', corresponding to dry, storm and rain weather disturbance scenarios (Copp, 2002). However, at the outset of the BSM2 development, there was a general consensus that a 1-week evaluation period was insufficient to evaluate WWTP controller performance, especially when 'slow' actuators, such as the waste sludge flow rate, are manipulated (Gernaey *et al.*, 2006). Within the context of the BSM2 development, several options were discussed including simply repeating the weekly disturbance scenarios from BSM1, collecting 'real' data from an operating facility or creating a mathematical tool that could be used to generate a user-defined influent. The latter approach was chosen for several reasons, but the main reason was that those involved in the development felt that the model approach would solve several key problems including: 1) it would give sufficient

flexibility to manipulate the influent to suit the BSM2 requirements; 2) it would not be skewed by a 'real' event that may or may not have occurred in a 'real' plant; and, 3) it would be modular, which would allow this tool to be used in isolation outside of the BSM2 context. Figure 2 shows a schematic representation of the model developed by Gernaey *et al.* (2005) and Gernaey *et al.* (2006).

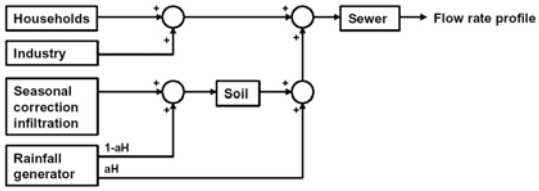


Figure 2: Schematic representation of influent generator model (Gernaey *et al.*, 2006).

The model (Figure 2) contains contributions from households, industry, rainfall and groundwater infiltration. Sub-models that include things like diurnal pollutant fluxes in the case of households, and weekend and holiday effects in the industrial model generate each of these contributing streams. These influent disturbance models allow the creation of influent dynamics that can include diurnal, weekend, seasonal and holiday effects, as well as rainfall. Being able to simulate these effects is important for control strategy evaluation, but this influent generator has far-reaching possibilities for the wider modelling community.

Temperature Model

The evolution of the BSM models has been an interesting study in modelling complexity as with each new addition has come several new challenges. The simulation procedure defined with BSM1 assumed a constant temperature of 15^{o}C. By extending the simulation period in BSM2, it was necessary to include changing temperatures in each of the streams and unit processes. Temperature has a recognised impact on the biological activity in the ASPs and digester and on the oxygen mass transfer in the ASPs. This, combined with a year-round warm return stream from the digester required that a temperature model be developed that might be used to estimate changing temperatures in each stream.

For points in the model where streams combine, several proposals were discussed from complicated heat balance models, to simple mass flow heat blending models. However, based on the fact that the BSM model is assumed to have a slow temperature dynamic, it was deemed reasonable to adopt the simple heat blending methodology. In this method, at points in the model where streams meet, heat mass flows are calculated with the total outgoing mass flow simply divided by the outgoing flow to give an estimate of the outgoing temperature. The fact that the more complicated models were rejected is no reflection on their application, but simply a further contribution of the BSM work in that both simple and more complicated solutions were discussed, debated and documented.

In addition to estimating the temperatures in the various streams, it was of interest to model the impact of temperature on the modelled kinetics and process parameters like oxygen saturation, oxygen transfer rates and by extension, energy for aeration. Typically the commercially available simulation packages have temperature models incorporated, but the BSM debate has highlighted that different relationships exist and in a general sense given them another option for these relationships.

Model Implementation Ring-Testing

One of the first highly regarded outcomes from the BSM development work was the ring-testing of several implementations of activated sludge model #1 (ASM1) and the Takács secondary clarification model. Computer simulation of wastewater treatment systems is a powerful tool, however, critical to the BSM concept is that any simulations carried out, any where in the world, using any simulator must be directly comparable to results generated everywhere else (Figure 3). This, therefore, required that the model implementations in each platform be exactly the same.

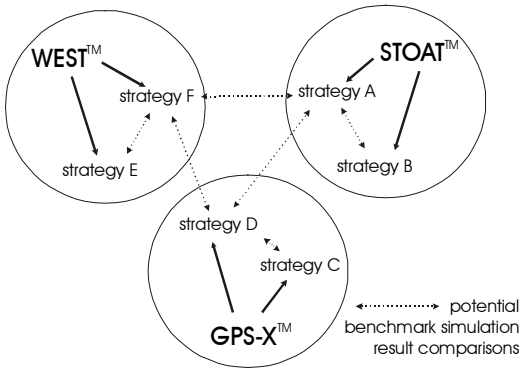


Figure 3: Illustrative representation of the comparison concept showing that in addition to comparisons being made between results generated with the same simulator, results from different simulators can also be compared.

In this case, ASM1 and the settling model were implemented into 5 commercial simulation packages (WEST, STOAT, Simba, GPS-X, BioWin) and 2 open code platforms (Matlab/Simulink, Fortran). Each of the simulator platforms had different specific features that made getting the simulators to produce exactly the same results difficult. Issues that were discovered included things like aeration model differences, simulator-specific model alterations, and also errors in the model code. Each of these differences was investigated and 'corrected' so that each simulator eventually produced the same steady state result. This steady state investigation was followed by a dynamic simulation test and here again, simulator-specific issues resulted in different dynamic results. The most prevalent problem identified at this stage was the implementation of the settler model as each package seemed to handle the clarifier's soluble components a little bit differently. Nevertheless, after nearly 2 years of work, the BSM co-operative effort had ring-tested 7 ASM1 implementations, 'corrected' any differences and achieved the same results (to several decimal places) in all platforms (Copp, 2002) proving that it was possible to achieve the same results in all platforms, but exceptional care must be taken to do it. The importance of this aspect relates to the goal of the simulation benchmark development; namely, the development of a platform independent standardised evaluation protocol, but in the larger context, this work has resulted in debugged ASM1 model code. The commercial simulators now include this ASM1 implementation in their packages so users can be assured that when using ASM1 (or the Takács settling model) in one of these packages they are using a fully tested and verified version of the model.

A similar exercise was carried out for ADM1. ADM1 (Batstone *et al.*, 2002) was implemented into the simulation packages and tested in the same way as previously described for ASM1. Similar results were found in that each simulation package required special considerations, but after these simulator-specific issues were identified, each package gave the same results (again to several decimal points). In the case of ADM1, the model had to be modified for BSM2 to optimise the simulation performance. An important difference between the ADM1 of Batstone *et al.* (2002) and the ADM1 for BSM2 is the introduction of continuous inhibition functions for pH to avoid simulation problems related to discontinuities. In Batstone *et al.* (2002), it is suggested that ADM1 be implemented as a differential algebraic system, with algebraic equations for the acid-base equilibrium (although differential equations are also given in the report). This is, however, not sufficient to remove the stiffness of the system as it was discovered that the hydrogen state is much faster than the remaining states. Therefore, an algebraic solution for the hydrogen state was implemented for BSM2. This is an important finding as the *error* introduced by this change is insignificant yet this change is critical for some simulation platforms that need to use non-stiff solvers to handle the noise and discrete events that have been introduced for realism in BSM2. Detailed descriptions of the BSM2 implementation of ADM1 are given in Rosen *et al.* (2006) and Rosen and Jeppsson (2006). As with ASM1,

the BSM work has generated a standardised implementation of ADM1 and resulted in a speed enhancement that makes this model more accessible and usable to the general modelling public.

The primary clarifier model (Otterpohl and Freund, 1992; Otterpohl *et al.*, 1994) chosen for BSM2 was also ring-tested. For the most part this was a new model to most of the packages so a standard implementation was easier in this case. As a contribution to the general modelling community, because it was a new primary model in most cases, the BSM effort has increased the choice of models for primary clarification where limited options were previously available.

Activated Sludge / Anaerobic Digestion Model Interfacing

Unfortunately (for wastewater treatment modellers) not all wastewater treatment unit process models have a common set of state variables which means that if two dissimilar unit processes are linked in reality and are to be simulated together in one model, a methodology for transforming the one set of states to the other must be developed (Figure 4).

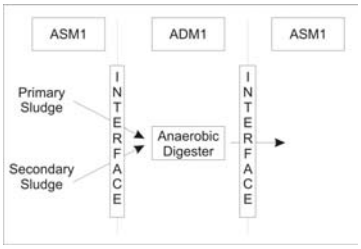


Figure 4: A conceptual representation of the interfaces needed to join dissimilar models.

The first benchmark model (BSM1) was comprised of the liquid treatment stream only, but BSM2 was expanded to include the sludge train which introduced a number of complicating issues; one of them being the coupling of ASM1 and ADM1. As this was crucial for BSM2 development, again the BSM team co-operated, debated and compromised to arrive at a reasonable solution for these transformations. An initial attempt at a transformation was made by Copp *et al.* (2003). The strengths and weaknesses of that approach were subsequently debated and a more advanced method specifically designed to account for differences in primary and secondary sludges in the digester was developed by Nopens *et al.* (2008) (Figure 5).

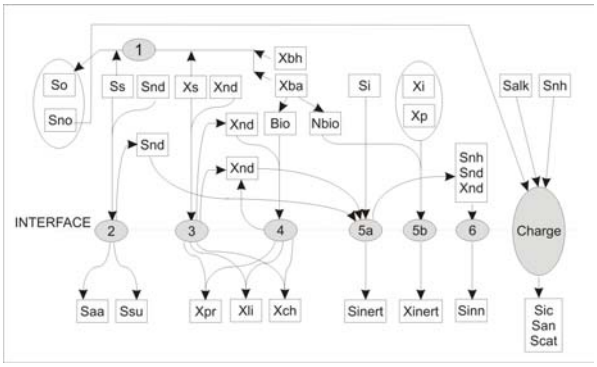


Figure 5: The BSM2 ASM1 (top) to ADM1 (bottom) interface.

However, more importantly to the general modelling community, this effort identified several deficiencies and spurred on further developments as several BSM contributors have since developed more generally applicable methodologies for interfacing all kinds of different models (Vanrolleghem *et al.*, 2005; Volcke *et al.*, 2006a).

Effluent, Cost & Risk Indices

Because simulations can generate an enormous volume of output data, comparison of that output data is difficult without some level of post-processing. The co-operative effort involved in the BSM development debated the merits of several options and settled on a performance assessment largely based on measures of general interest including:

- effluent quality
- operational costs
- risk

Effluent quality is considered through an effluent quality index (EQI), which is meant to quantify into a single term, the effluent pollution load to a receiving water body (Eq. 1) and operational costs are considered through an operational cost index (OCI) that includes seven terms.

$$EQI = \frac{1}{1000\,T}\int_{t_0}^{t_i}\left[PU_{TSS}(t) + PU_{COD}(t) + PU_{BOD}(t) + PU_{TKN}(t) + PU_{NO}(t)\right]Q_e(t)\,dt \qquad (1)$$

$$OCI = AE + PE + 3\cdot SP + 3\cdot EC + ME - 6\cdot MP + \max(0, HE^{net}) \qquad (2)$$

where PU_{xxx} is calculated as the product of β_{xxx} and the concentration of XXX at time *(t)*. The β_{xxx} factors were determined based, in part, on empirical effluent component weightings from a paper by Vanrolleghem *et al.* (1996) which cited a Flanders effluent quality formula for calculating fines. AE represents aeration energy (kWh/d), PE is pumping energy (kWh/d), SP is sludge production for disposal (average kg TSS/d), EC is external carbon addition (average kg COD/d), ME is mixing energy (kWh/d), MP represents methane production (average kg CH_4/d) and HE^{net} is the net heating energy needed to heat the sludge in the anaerobic digester. In the OCI case, the AE, PE and ME are in turn calculated based on more specific sub-models.

The pumping energy model has been modified several times over the years as new information and issues were addressed again reflecting the co-operative effort and compromises that have been adopted during the development. Initially the pumping energy was simply calculated as a constant number of kWh per m^3 pumped (the same for all streams), but this has since evolved into various ratios depending on the liquid being pumped (RAS vs WAS vs primary sludge vs….). As energy consumption becomes more and more important to the operation of treatment systems, these models could form the basis on which to estimate energy consumption and costs in any model.

To further enhance the objective evaluation of the BSMs a third type of performance index has also been developed: the risk index. This index adds a qualitative dimension to the otherwise only quantitative results from benchmark simulations. Based on a knowledge data base and fuzzy logic, a risk assessment of the simulated system is made, which estimates the risk of activated sludge system settling problems, e.g. filamentous bulking, foaming and rising sludge (Comas et al., 2006; Comas et al., 2008). The index is used to demonstrate that some control strategies, although performing better with regard to operating costs and effluent quality, induce a higher risk for solids separation problems. This is another module that can be used outside the scope of the BSMs.

Sensors and Actuators

In order to model any control strategy, sensors and actuators have to be modelled. For most modelling exercises, ideal (no noise, measurement error or time delay) sensors and actuators are used and in most cases this is sufficient. However, the reality of the situation is that sensors and actuators are not ideal. They are subject to errors and signal processing delays and possess particular dynamics due to the measuring principles (e.g. chemical reactions that must be completed within on-line analyzers). Models to describe these sensor and actuator behaviours have been developed within the BSM community and are now available for much wider use (Rosen et al., 2008).

To account for this, sensors in BSM2 can be ideal, but they can also be modelled based on

the principles of Rieger *et al.* (2003). A number of sensor classes have been defined from which a benchmark user selects the ones most appropriate. Each sensor class includes characteristics such as noise level, time response, delay time, signal saturation levels and sampling time. All actuators are considered ideal except the aeration system, which is described using a simple model creating a delay in the K_La inputs and the reject water storage tank, which requires a somewhat more complex model.

Here again, modular sub-models within the BSM context have been created. As with all the other pieces included in the BSMs, these can be used outside the BSM context as required by the larger modeling community, thus giving that modeling community the option to include or not include these real-world issues in their sensor and actuator simulations.

LIMITATIONS

The BSM models have provided a basis for many debates regarding whole-plant modelling and have resulted in some very well-tested compromises, but as with any tool of this sort, there are limitations. The published models chosen for each unit process might not be the best ones in existence, the sizes of tanks and the influent used, might cause peculiar behaviour that would not otherwise happen in reality and lastly although these tools provide the best option for objectively evaluating all kinds of control strategies simulated anywhere by anyone, there still exists the possibility that strategies will not be comparable because of the various options available to the user. Care has been taken to eliminate as much of this as possible, but this possibility still exists.

CONCLUSION

The use of this modelling toolbox provides an excellent starting point for modelling and evaluating many systems. Examples of such applications have been recently presented by Volcke *et al.* (2006b) and Benedetti *et al.* (2006). These 'extensions' are an unquantifiable benefit of the BSM work and show that the BSM influence has not been restricted to control strategy evaluation alone but rather emphasises the importance of the effort to modelling in general. The inclusion of primary treatment as well as sludge treatment in BSM2 increases the complexity of the system but more importantly allows for the study of unit process interaction and has forced the benchmark team to develop and consider a new set of modelling tools that have far reaching implications and uses. The toolbox has been freely distributed to modelling groups on all continents to provide a structured, documented and validated starting point for their future work. Seeing beyond the narrow application of the BSMs as fully defined, should silence the critics and highlight the value of the modelling toolbox that has been created through a world-wide cooperative effort.

ACKNOWLEDGEMENTS

Many excellent researchers and close friends have contributed to the development of the BSMs over the years and the authors wish to express their sincere gratitude to all of them. The authors are also grateful for the support by IWA when establishing the BSM Task Group. Peter Vanrolleghem holds the Canada Research Chair in Water Quality Modelling.

REFERENCES

Batstone D.J., Keller J., Angelidaki I., Kalyuzhnyi S.V., Pavlostathis S.G., Rozzi A., Sanders W.T.M., Siegrist H. and Vavilin V.A. Anaerobic digestion model No. 1. ISBN 19-002-2278-7, IWA STR No 13, IWA Publishing, London, UK (80 pages). 2002

Benedetti L., Bixio D. and Vanrolleghem P.A. Benchmarking of WWTP design by assessing costs, effluent quality and process variability. *Water Science & Technology*, 54(10), 95-102 2006

Copp J.B. (ed.). The COST simulation benchmark — description and simulator manual. ISBN 92-894-1658-0, Office for Official Publications of the European Communities, Luxembourg. 2002

Copp J.B., Spanjers H. and Vanrolleghem P.A. (eds) Respirometry in control of the activated sludge process: benchmarking control strategies. ISBN 19-002-2251-5, IWA STR No 11, IWA Publishing, London, UK (192 pages). 2002

Copp J.B., Jeppsson U. and Rosen C. Towards an ASM1 - ADM1 state variable interface for plant-wide wastewater treatment modeling. Proc. 76th Annual WEF Conference and Exposition (WEFTEC), Los Angeles, USA, Oct. 11-15. 2003

Comas J., Rodríguez-Roda I., Gernaey K.V., Rosen C., Jeppsson, U. and Poch, M. Risk assessment modelling of microbiology-related solids separation problems in activated sludge systems". Environmental Modelling and Software (submitted). 2008

Comas J., Rodríguez-Roda I., Poch M., Gernaey K.V., Rosen C. and Jeppsson, U. Extension of the IWA/COST simulation benchmark to include expert reasoning for system performance evaluation. *Water Science & Technology*, 53(4-5), 331-339. 2006

Gernaey, K.V., Rosen, C. and Jeppsson, U. WWTP dynamic disturbance modelling–an essential module for long-term benchmarking development. *Water Science & Technology*, 53 (4–5), 225–234. 2006

Gernaey, K.V., Rosen, C., Benedetti L and Jeppsson, U. Phenomenological modeling of wastewater treatment plant influent disturbance scenarios. *Preprints of the 10th International Conference on Urban Drainage*, Copenhagen, Denmark, 21–26 Aug. 2005

Henze M., Grady Jr C.P.L., Gujer W., Marais G.v.R. and Matsuo T. Activated Sludge Model No.1. IWA STR No 1, IWA Publishing, London, UK. 1987

Jeppsson U., Rosen C., Alex J., Copp J.B., Gernaey K.V., Pons M.-N. and Vanrolleghem P.A. Towards a benchmark simulation model for plant-wide control strategy performance evaluation of WWTPs. *Water Science & Technology*, 53(1), 287-295. 2006

Nopens I., Batstone D., Copp J.B., Jeppsson U., Volcke E.I.P., Alex J. and Vanrolleghem, P.A. A practical ASM/ADM model interface for enhanced dynamic plant-wide simulation. Manuscript in preparation. 2008

Otterpohl, R.; Freund, M.: Dynamic Models for clarifiers of activated sludge plants with dry and wet weather flows. *Water Science & Technology*, 26 (5-6), 1391-1400. 1992

Otterpohl, R., Raak, M., and Rolfs, T., A mathematical model for the efficiency of the primary clarification. Proceedings of the 17th IAWQ Biennial Conference on Water Quality, Budapest Hungary, July.1994

Rieger L., Alex J., Winkler S., Boehler M., Thomann M. and Siegrist, H. Progress in sensor technology - progress in process control? Part I: Sensor property investigation and classification. *Water Science & Technology*, 47(2), 103-112. 2003

Rosen C., Jeppsson U. and Vanrolleghem, P.A. Towards a common benchmark for long term process control and monitoring performance evaluation, *Water Science & Technology*, 50(11), 41-49. 2004

Rosen C. and Jeppsson U. Aspects on ADM1 implementation within the BSM2 Framework. Tech. Report no. LUTEDX/(TEIE-7224)/1-35/(2006). Dept of Industrial Electrical Engineering and Automation, Lund University, (available at www.benchmarkwwtp.org). 2006

Rosen C., Vrecko D., Gernaey, K.V., Pons M.-N. and Jeppsson U. Implementing ADM1 for plant-wide benchmark simulations in Matlab/Simulink. *Water Science & Technology*, 54(4), 11-19. 2006

Rosen C., Rieger L., Jeppsson U. and Vanrolleghem P.A. Adding realism to simulated sensors and actuators. *Water Science & Technology*, (in press). 2008

Spanjers H., Vanrolleghem P.A., Nguyen K., Vanhooren H. and Patry G.G. Towards a simulation-benchmark for evaluating respirometry-based control strategies. *Water Science & Technology*, 37(12), 219-226. 1998

Takács I., Patry G.G. and Nolasco D. A dynamic model of the clarification thickening process. *Water Research*, 25(10), 1263-1271. 1991

Vanrolleghem P.A., Jeppsson U., Carstensen J., Carlsson B. and Olsson G. Integration of WWT plant design and operation - A systematic approach using cost functions. *Water Science & Technology*, 34(3-4), 159-171. 1996

Vanrolleghem P.A., Rosen C., Zaher U., Copp J.B., Benedetti L., Ayesa E. and Jeppsson U. Continuity-based interfacing of models for wastewater systems described by Petersen matrices. *Water Science & Technology*, 51(1), 493-500. 2005

Volcke E.I.P., van Loosdrecht M.C.M. and Vanrolleghem P.A. Continuity-based model interfacing for plant-wide simulation : A general approach. *Water Research*, 40, 2817-2828. 2006a

Volcke E.I.P., Gernaey K.V., Vrecko D., Jeppsson U., van Loosdrecht M.C.M. and Vanrolleghem, P.A. Plant-wide (BSM2) evaluation of the effect of reject water treatment with a SHARON-Anammox process. *Water Science & Technology*, 54(8), 93-100. 2006b

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

Global sensitivity analysis of biochemical, design and operational parameters of the Benchmark Simulation Model no. 2

Lorenzo Benedetti[a,c], Damien J. Batstone[b], Bernard De Baets[c], Ingmar Nopens[a] and Peter A. Vanrolleghem[d]

[a] BIOMATH, Ghent University, Coupure links 653, B-9000 Ghent, Belgium, Lorenzo.Benedetti@UGent.be, Ingmar.Nopens@UGent.be
[b] AWMC, The University of Queensland, 4067 Australia, d.batstone1@uq.edu.au
[c] KERMIT, Ghent University, Coupure links 653, B-9000 Ghent, Belgium, Bernard.DeBaets@UGent.be
[d] modelEAU, Université Laval, Pavillon Pouliot, 1065 av. de la Médecine, Québec QC, G1V 0A6, Canada, Peter.Vanrolleghem@gci.ulaval.ca

Abstract: Wastewater treatment plant control and monitoring can help to achieve good effluent quality, in a complex, highly non-linear process. The Benchmark Simulation Model no. 2 (BSM2) is a useful tool to competitively evaluate plant-wide control on a long-term basis. A key component to characterise the system for control is output-parameter sensitivity. This paper brings the results of a global sensitivity analysis performed on the BSM2 model in its open loop version, by means of Monte Carlo (MC) experiments and linear regression. This study presents methods that were applied to make computationally demanding MC experiments on such a complex model feasible, by reducing the computation time for a single simulation and by setting low but sufficient number of runs for the MC experiments; it was found that 50 times the number of uncertain parameters was necessary. The most sensitive parameters turned out to be the design and operation parameters, followed by the wastewater treatment model parameters, while the adopted BSM2 evaluation criteria are rather insensitive to variations in sludge treatment models parameters. The results are verified on a closed loop version of BSM2, and allow future uncertainty analysis studies on BSM2 to be conducted on a smaller set of parameters and to focus the attention on the most critical parameters.

***Keywords**:* activated sludge; anaerobic digestion; BSM2; mathematical modelling; numerical methods.

1. INTRODUCTION

The biological, physical and chemical phenomena taking place in activated sludge systems are complex, interrelated and highly non-linear. Moreover, the operation of these systems should continuously meet effluent requirements, preferably at the lowest possible operational cost. In order to achieve this, monitoring and control of such plants can be very helpful but, given the complexity, this is not an easy task. Operators are often reluctant to test new control strategies on the real plant because of their possibly unexpected behaviour.

Originated in the 90's, the Benchmark Simulation Model no. 1 (BSM1) was proposed as a tool to foster the dissemination of control and monitoring strategies [Copp, 2002]. This benchmark is a simulation environment defining a plant layout, simulation models for all process units, influent loads, test procedures and evaluation criteria. For each of these items, compromises were made to match model simplicity with realism and accepted standards. Once the user has verified the simulation code, any control strategy can be applied and the performance can be evaluated according a well defined set of criteria.

Recently, the BSM2 [Jeppsson et al., 2007] was developed for plant-wide WWTP control strategy evaluation on a long-term basis, with a much more complex plant model. It consists of a pre-treatment process, an activated sludge process and sludge treatment processes.

This paper shows the results of a global sensitivity analysis (SA) performed on the BSM2 model in its open loop (without control) version, by means of Monte Carlo (MC) experiments and linear regression of the MC results [Saltelli et al., 2000]. The parameters for which the sensitivity is computed belong to the biochemical and physical models and to the design and operation of the plant. The study discusses the methods applied to reduce the computational efforts required by such a complex model, by testing possibilities to reduce the computation time of a single simulation, and by looking for a number of simulation runs for the MC experiments sufficient to accept the results of the sensitivity analysis.

2. METHODS

2. 1 The Model

The Benchmark Simulation Model no. 2 protocol [Jeppsson et al., 2007] consists of a plant wide (including wastewater and sludge treatment) model representing a general WWTP, a benchmarking procedure and a set of evaluation criteria. The three evaluation criteria used in this work are: (1) the Effluent Quality Index (EQI), a weighted sum of effluent pollutant loads with weight values set to 2 for BOD, 1 for COD, 2 for TSS, 30 for NH_4 and 10 for NO_3; (2) the Operating Cost Index (OCI) which takes into account energy consumption (aeration, pumping, mixing), external carbon addition, waste sludge production, heating of the digester and energy recovery from methane production; (3) the fraction of time in which the effluent exceeds the limit of 4mgNH_4/l, expressed as percentage of the whole evaluation period (one year, the last 365 of the 609 simulated days).

2. 2 Solver Optimisation

The BSM2 contains 265 differential equations and requires a simulation time of 609 days in very dynamic conditions (the evaluation is based on the last 365 days). In order to perform a global sensitivity analysis of such a complex model, potentially involving a very large number of MC simulations, careful selection of numerical settings is needed to minimise the time required to run a single simulation.

The modelling and simulation software used in this work was WEST (MOSTforWATER, Kortrijk, Belgium) with its new numerical engine Tornado [Claeys et al., 2006a]. The starting point was the Runge-Kutta 4th order adaptive step-size (RK4ASC) solver [Forsythe et al., 1977] with accuracy, initial and minimum step size set to 10^{-6} – which are the solver settings normally used with this type of models to provide very accurate results at reasonable computation cost.

Advanced solvers such as CVODE [Hindmarsh et al., 2005] often show a better performance, and an approach based on scenario analysis was applied to find the best solver settings (see also Claeys et al. [2006b]), which provided as optimum: IterationMethod: Newton; LinearMultistepMethod: Adams; LinearSolver: SPGMR. Using those settings, results are shown in Table 1 with regard to computation time and difference from the reference (RK4ASC) for EQI, OCI and ammonium exceedance periods. The best compromise between solution difference and calculation time was found for a solver accuracy of 10^{-3}.

Another aspect evaluated to reduce computation time and storage requirement was the reduction of output frequency. The standard for BSM2 is 15 minutes, and output frequencies of 30, 45 and 60 minutes were tested (see Table 1). The frequency of 30 minutes was chosen since it still provided acceptable results – leaving EQI and OCI practically unchanged and with NH_4 exceedance 3% different – in shorter time and with

half the output file size, which is an important factor for storage and post-processing of files. In other types of studies lower frequencies can be accepted [Ráduly et al., 2007]. The selected settings allow therefore, compared to the reference settings, a reduction to almost 1/5 of the computation time and to 1/2 of the output file size.

Table 1. Simulation performance for different solver settings and output frequencies; in dark grey the reference simulation settings, in light grey the best settings.

Solver	Accuracy	Output freq. [min]	File size [MB]	Computation time [s]	ΔEQI [%]	ΔOCI [%]	ΔNH_4 [%]
RK4ASC	10^{-6}	15	13.4	571	0.0000	0.0000	0.0000
CVODE	10^{-5}	15	13.4	249	0.0000	0.0001	-0.1391
CVODE	10^{-4}	15	13.4	158	-0.0140	-0.0132	0.0585
CVODE	10^{-3}	15	13.4	131	-0.0170	-0.0127	-0.0804
CVODE	10^{-2}	15	13.4	133	0.1102	-0.0056	1.6840
CVODE	10^{-3}	30	6.7	121	-0.0223	-0.0003	-3.4363
CVODE	10^{-3}	45	5.0	119	-0.0589	-0.0091	-10.7498
CVODE	10^{-3}	60	3.3	118	-0.0528	0.0198	-20.2360

2. 3 Method for Sensitivity Analysis

The sensitivity of the three BSM2 evaluation criteria towards model parameters was assessed by means of MC experiments – which consist of performing multiple simulations with parameter values sampled from Probability Density Functions (PDFs) – and linear regression to calculate the Standardised Regression Coefficients (SRCs) and the Partial Correlation Coefficients (PCCs) of the parameters considered uncertain [Saltelli et al., 2000]. The SRCs represent the change in an output variable that results from a change of one standard deviation in a parameter, while the PCCs are the measure of linear dependence between an output variable and a parameter in the case where the influence of the other parameters is eliminated. A number N of simulations was run for each MC experiment, sampling from the PDFs of the parameters with Latin Hypercube Sampling (LHS) [Benedetti et al., 2008]. To evaluate the quality of the linear regression, the coefficient of determination R^2, i.e. the fraction of the input variance reproduced by the regression model, was calculated; the regression is considered of good quality when R^2>0.7. The calculation of the t-statistic on the SRCs and PCCs [Morrison, 1984] allowed classifying the parameters as significant at the 5% level with a t-statistic larger than 1.96.

The number N is equal to n times the number of uncertain parameters, and n was determined as follows. Running the MC experiments with uncertain design and operational parameters (19 parameters), n was set to 4/3, 3, 12 and 20, i.e. N was 26, 57, 228 and 380. Since the ranking of the parameter sensitivities made on the basis of the SRCs and PCCs was different in all MC experiments (including three different MC experiments with n=20), it was assumed that n=20 was not sufficient, in disagreement with Manache and Melching [2008], where n=3 was sufficient for a model with similar structure, but probably with lower complexity. Three more MC experiments were performed with n set to 50 (N=950), and in this case the differences were less pronounced, allowing to select n=50 for the rest of the MC experiments as a compromise between accuracy of results and feasibility of computation.

The parameters were divided into three groups (see Table 2 for details): (1) design and operational (DO) parameters, including volumes, recirculation rates, etc.; (2) wastewater treatment (WT) parameters, including some parameters of the ASM1 and of the primary and secondary settler models; (3) sludge treatment (ST) parameters, including some parameters of the ADM1 and interface parameters. Model parameters selected for testing were based on operational knowledge, previous studies, and our own sensitivity screening. Of course, a different choice for the PDFs might lead to different results [Benedetti et al., 2008].

The PDFs of the parameters regarding design and operation of the plant were defined as uniform with their mean set to the default value for BSM2 and boundaries set as +/-20% of the mean. The PDFs of the ASM1 parameters were taken from Rousseau et al. [2001], while for all the other parameters the PDFs were assumed to be triangular with median

equal to the BSM2 default and boundaries at +/-20% of the median. The PDFs of the ADM1 parameters were mainly taken from Appendix A in Batstone et al. [2002], with additional information from Batstone et al. [2003; 2004] and Siegrist et al. [2002], while for the AD/AS model interfaces parameters they were assumed to be triangular with median equal to the BSM2 default and boundaries +/-20% of the median.

Table 2. PDFs of parameters; LB=lower bound, UB=upper bound, DO=design and operation, WT=wastewater treatment, ST=sludge treatment, T=triangular, U=uniform.

Parameter	Description or reference	Group	PDF	Median	LB	UB
AD.V_gas	Volume of gas in AD tank, in m^3	DO	U	-	240	360
AD.V_liq	Volume of liquid in AD tank, in m^3	DO	U	-	2720	4080
ASU3.Kla	kLa in AS reactor no.3, in d^{-1}	DO	U	-	96	144
ASU4.Kla	kLa in AS reactor no.4, in d^{-1}	DO	U	-	96	144
ASU5.Kla	kLa in AS reactor no.5, in d^{-1}	DO	U	-	48	72
C_source	C-source with COD=400000g/m^3, in m^3/d	DO	U	-	1.6	2.4
dewatering.rem_perc	TSS removal fraction in dewatering	DO	U	-	0.96	1
dewatering.X_under	TSS underflow concentration, as fraction	DO	U	-	0.224	0.336
internal_rec	Internal mixed liquor recirculation, in m^3/d	DO	U	-	49555.2	74332.8
PC.f_PS	Primary settler underflow as ratio on inflow	DO	U	-	0.0056	0.0084
PC.Vol	Primary settler volume, in m^3	DO	U	-	800	1200
SC.A	Surface area of secondary settler, in m^2	DO	U	-	1200	1800
SC.H	Height of secondary settler, in m	DO	U	-	3.2	4.8
SC.Q_Under	Underflow of secondary settler, in m^3/d	DO	U	-	16518.4	24777.6
sec_sludge_to_AD	Secondary sludge to AD, in m^3/d	DO	U	-	240	360
thickener.rem_perc	TSS removal fraction in thickener	DO	U	-	0.96	1
thickener.X_under	TSS underflow concentration, as fraction	DO	U	-	0.056	0.084
Vol_aer	Volume of each aerated tank, in m^3	DO	U	-	2400	3600
Vol_anox	Volume of each anoxic tank, in m^3	DO	U	-	1200	1800
f_P	Henze et al. [1987]	WT	T	0.08	0.076	0.084
F_TSS_COD	TSS/COD ratio	WT	T	0.75	0.7125	0.7875
i_X_B	Henze et al. [1987]	WT	T	0.08	0.076	0.084
i_X_P	Henze et al. [1987]	WT	T	0.06	0.057	0.063
k_a	Henze et al. [1987]	WT	T	0.05	0.025	0.075
k_h	Henze et al. [1987]	WT	T	3	1.5	4.5
K_NH	Henze et al. [1987]	WT	T	1	0.5	1.5
K_NO	Henze et al. [1987]	WT	T	0.5	0.25	0.75
K_OA	Henze et al. [1987]	WT	T	0.4	0.2	0.6
K_OH	Henze et al. [1987]	WT	T	0.2	0.1	0.3
K_S	Henze et al. [1987]	WT	T	10	5	15
K_X	Henze et al. [1987]	WT	T	0.1	0.05	0.15
mu_A	Henze et al. [1987]	WT	T	0.5	0.4	0.6
mu_A_b_A	mu_A/b_A ratio, for correlation	WT	U	-	9.5	10.5
mu_H	Henze et al. [1987]	WT	T	4	3.2	4.8
mu_H_b_H	mu_H/b_H ratio, for correlation	WT	U	-	12.66	13.99
n_g	Henze et al. [1987]	WT	T	0.8	0.64	0.96
n_h	Henze et al. [1987]	WT	T	0.8	0.64	0.96
PC.f_X	Otterpohl et al. [1994]	WT	T	0.86	0.765	0.935
SC.f_ns	Takács et al. [1991]	WT	T	0.0023	0.0018	0.0027
SC.r_H	Takács et al. [1991]	WT	T	0.0006	0.0005	0.0007
SC.r_P	Takács et al. [1991]	WT	T	0.00286	0.00228	0.00343
SC.v0	Takács et al. [1991]	WT	T	474	379.2	568.8
SC.v00	Takács et al. [1991]	WT	T	250	200	300
SC.X_Lim	Takács et al. [1991]	WT	T	900	720	1080
SC.X_T	Takács et al. [1991]	WT	T	3000	2400	3600
Y_A	Henze et al. [1987]	WT	T	0.67	0.6365	0.7035
Y_H	Henze et al. [1987]	WT	T	0.24	0.228	0.252
AD.kdis	Batstone et al. [2002]	ST	T	0.7	0.5	1
AD.khyd_ch	Batstone et al. [2002]	ST	T	0.8	0.5	1
AD.khyd_li	Batstone et al. [2002]	ST	T	1.1	0.7	1.5
AD.khyd_pr	Batstone et al. [2002]	ST	T	1.1	0.7	1.5
AD.KI_nh3_ac_km_ac	KI_nh3_ac/km_ac ratio, for correlation	ST	U	-	0.00013	0.00015
AD.kla	Batstone et al. [2002]	ST	T	150	50	200
AD.km_ac	Batstone et al. [2002]	ST	T	10	8	12
AD.km_c4	Batstone et al. [2002]	ST	T	15	10	20
AD.km_fa	Batstone et al. [2002]	ST	T	15	10	20
AD.km_pro	Batstone et al. [2002]	ST	T	10	8	12
AD.Ks_ac_km_ac	Ks_ac/km_ac ratio, for correlation	ST	U	-	0.025	0.083
AD.Ks_c4_km_pro	Ks_c4/km_pro ratio, for correlation	ST	U	-	0.025	0.1
AD.Ks_fa_km_pro	Ks_fa/km_pro ratio, for correlation	ST	U	-	0.025	0.1
AD.Ks_pro_km_pro	Ks_pro/km_pro ratio, for correlation	ST	U	-	0.025	0.083
ADM2ASM.frxs_AS	Nopens et al. [2008]	ST	T	0.7505	0.711	0.79
ASM2ADM.frlixb	Nopens et al. [2008]	ST	T	0.4	0.38	0.42
ASM2ADM.frlixs	Nopens et al. [2008]	ST	T	0.7	0.665	0.735
ASM2ADM.frxs	Nopens et al. [2008]	ST	T	0.646	0.612	0.68

3. RESULTS

Four different MC experiments were performed to conduct the SA on: (1) design and operational parameters, (2) wastewater treatment parameters, (3) sludge treatment parameters and (4) all parameters together.

Performing the SA on the design and operational parameters, no less than 17 out of 19 parameters are significant for all three criteria based on the SRCs and 12 based on the PCCs. PCCs are indeed known to produce a smaller number of significant parameters [Manache and Melching, 2008]. As expected, the aerated volume (Vol_aer) is in general the most important parameter, followed by the air supply (Kla) in the three aerated tanks and by the external carbon dosage (C_source). Also relevant is the highest importance of the primary clarifier underflow (PC.f_PS) for the OCI, given the fact that primary sludge is very well suited for methane production. The surface of the secondary clarifier and the anoxic volume are very important for the EQI.

From the analysis on the wastewater treatment parameters, only 4 out of the 28 parameters were judged as not significant for the SRCs and 9 for the PCCs, in this case because of the very different importance of the parameters towards environmental and economic performance. The only ones that strongly influence both EQI (but not NH_4) and OCI are Y_H of ASM1 and r_P and v0 of the secondary clarifier model. Very important for EQI and NH_4 are both K_OA and K_OH.

For the sludge treatment parameters, only one parameter out of 18 can be considered as not significant for all three criteria based on the SRCs, and 7 based on the PCCs. Clearly the most significant are khyd_pr of ADM1 and frxs of the AS/AD interface.

From Table 3, which shows the results for the SA on all parameters together, the three BSM2 evaluation criteria are mostly sensitive to design and operational parameters, and largely not to sludge treatment parameters. Ten out of 65 parameters were identified as not sensitive based on their *t*-statistic for SRC. With the significance tested on the *t*-statistic for the PCCs, only 25 of the original 65 parameters are classified as significant, with most of the AD parameters being not significant. An R^2>0.7 indicates a good quality of the linear regression.

Figure 1 shows the variability of the three evaluation criteria for the three parameter categories separately and altogether. It is clear that most of the output variability is due to the design and operational parameters, as suggested by the figures in Table 3.

The sludge treatment parameters only contribute to the OCI variability, because of the importance of methane production for cost recovery. The AD is largely dimensioned and is very stable in open loop. The complexity of ADM1 might be required in closed loop configurations which alter the AD influent and/or operation, pushing it towards instability.

Performing the uncertainty analysis on the BSM2 with the 25 most significant parameters only, the overall uncertainty in model output is practically unchanged, as can be seen in Figure 1. This means that sensitivity and uncertainty analyses on BSM2 can be performed by only assuming that reduced parameter set to be uncertain. Such reduced analysis will not lead to a loss of significant information and will be significantly faster to conduct.

To verify the transferability of these results to different configurations of the BSM2 (e.g. a control strategy), a SA was conducted for the open loop configuration on the 38 wastewater and sludge treatment parameters, which are the parameters to be considered for SA in case a specific design and operation configuration has to be evaluated. Based on the significance for the PCCs (see Table 3), a reduced set of 24 parameters can be accepted. Figure 2 shows the variability of the three evaluation criteria with the full and the reduced parameter sets for the open loop and a basic closed loop, consisting of a simple dissolved oxygen controller on the three aerated tanks, which strongly reduces the NH_4 exceedance period. It is evident that the changes in output variability from the full to the reduced parameter set are practically negligible in both BSM2 configurations.

Table 3. PCCs and ranking of all the parameters; in dark grey the parameters not significant for all three criteria based on SRC and PCC; in light grey significant for SRC but not for PCC; without shading significant for both SRC and PCC; in bold face not significant for SRC and PCC from the SA on WT and ST parameters (fixed DO parameters).

		EQI R^2=0.71		NH_4 R^2=0.97		OCI R^2=0.99	
Parameter	Group	PCC	rank	PCC	rank	PCC	rank
AD.V_gas	DO	0.01731	27	0.00757	32	-0.00040	58
AD.V_liq	DO	0.00956	35	-0.00630	34	-0.07242	10
ASU3.Kla	DO	-0.08355	10	-0.18763	2	0.15281	5
ASU4.Kla	DO	-0.10341	7	-0.17658	3	0.15493	4
ASU5.Kla	DO	-0.05268	15	-0.11993	4	0.07767	9
C_source	DO	-0.02537	22	0.02833	13	0.24269	3
dewatering.rem_perc	DO	0.00081	61	0.01247	23	0.01504	23
dewatering.X_under	DO	0.00487	45	-0.00791	30	0.00023	62
internal_rec	DO	-0.04679	17	-0.01410	20	0.08078	8
PC.f_PS	DO	-0.01618	28	-0.01307	21	0.39139	1
PC.Vol	DO	-0.02056	23	-0.04589	7	-0.08159	7
SC.A	DO	-0.20355	2	-0.00406	46	0.05186	13
SC.H	DO	-0.01974	24	0.00036	61	0.00369	41
SC.Q_Under	DO	-0.01106	33	-0.00099	57	-0.01100	27
sec_sludge_to_AD	DO	-0.00290	53	0.00358	51	0.02354	20
thickener.rem_perc	DO	0.00886	37	0.00762	31	0.00679	32
thickener.X_under	DO	0.00879	38	0.00396	47	-0.08870	6
Vol_aer	DO	-0.41771	1	-0.50165	1	0.35651	2
Vol_anox	DO	-0.09818	9	0.00343	52	0.02873	18
f_P	WT	0.00571	43	-0.00442	44	0.01851	21
F_TSS_COD	WT	0.00002	65	0.00859	27	0.01617	22
i_X_B	WT	-0.00213	55	-0.01245	24	0.00064	54
i_X_P	WT	0.00279	54	-0.00097	58	0.00050	57
k_a	WT	-0.06743	12	0.01114	25	-0.00038	59
k_h	WT	-0.02652	21	0.03058	10	0.00836	29
K_NH	WT	0.07427	11	0.03390	9	0.00114	50
K_NO	WT	0.03002	19	0.00165	55	-0.00022	63
K_OA	WT	0.15490	3	0.08899	5	0.00052	56
K_OH	WT	-0.14524	4	-0.03852	8	-0.00098	52
K_S	WT	0.00911	36	0.00521	39	-0.00008	65
K_X	WT	0.01856	25	-0.01730	16	-0.00389	38
mu_A	WT	-0.05464	14	-0.03004	11	-0.00015	64
mu_A_b_A	WT	-0.00649	42	-0.00034	62	0.00097	53
mu_H	WT	0.00072	63	0.02927	12	-0.01045	28
mu_H_b_H	WT	-0.01421	31	-0.02033	14	0.00220	44
n_g	WT	-0.06272	13	-0.01773	15	0.00123	47
n_h	WT	-0.02662	20	0.01285	22	0.00117	49
PC.f_X	WT	-0.00212	57	-0.00491	42	-0.00423	35
SC.f_ns	WT	0.04716	16	0.00390	48	-0.01104	26
SC.r_H	WT	0.11410	6	0.00537	38	-0.03566	15
SC.r_P	WT	-0.10327	8	0.00817	28	0.03038	17
SC.v0	WT	-0.11726	5	0.00581	37	0.03212	16
SC.v00	WT	-0.00212	56	0.01428	19	0.00372	40
SC.X_Lim	WT	-0.00957	34	-0.00017	64	0.00024	61
SC.X_T	WT	-0.00106	59	-0.00050	60	0.00032	60
Y_A	WT	-0.01787	26	-0.01546	18	0.00118	48
Y_H	WT	0.04541	18	-0.08133	6	0.07165	11
AD.kdis	ST	0.00071	64	0.00165	54	-0.00384	39
AD.khyd_ch	ST	-0.00331	51	-0.00185	53	-0.01456	24
AD.khyd_li	ST	0.00150	58	0.01009	26	-0.02715	19
AD.khyd_pr	ST	0.01463	30	0.00588	36	-0.04839	14
AD.KI_nh3_ac_km_ac	ST	0.00667	41	0.00811	29	-0.00100	51
AD.kla	ST	-0.00352	50	0.00373	50	0.00132	46
AD.km_ac	ST	-0.00439	48	-0.00617	35	-0.00409	37
AD.km_c4	ST	0.00105	60	0.00030	63	-0.00245	42
AD.km_fa	ST	-0.00479	46	0.00141	56	-0.00226	43
AD.km_pro	ST	-0.00851	39	-0.00433	45	0.00421	36
AD.Ks_ac_km_ac	ST	-0.00331	52	0.00494	41	0.01417	25
AD.Ks_c4_km_pro	ST	-0.00466	47	0.00733	33	0.00811	30
AD.Ks_fa_km_pro	ST	-0.00383	49	-0.00002	65	0.00425	34
AD.Ks_pro_km_pro	ST	-0.01571	29	-0.01715	17	0.00729	31
ADM2ASM.frxs_AS	ST	0.00074	62	0.00088	59	-0.00056	55
ASM2ADM.frlixb	ST	-0.00560	44	-0.00519	40	0.00147	45
ASM2ADM.frlixs	ST	0.00774	40	0.00482	43	-0.00560	33
ASM2ADM.frxs	ST	0.01408	32	0.00386	49	-0.06466	12

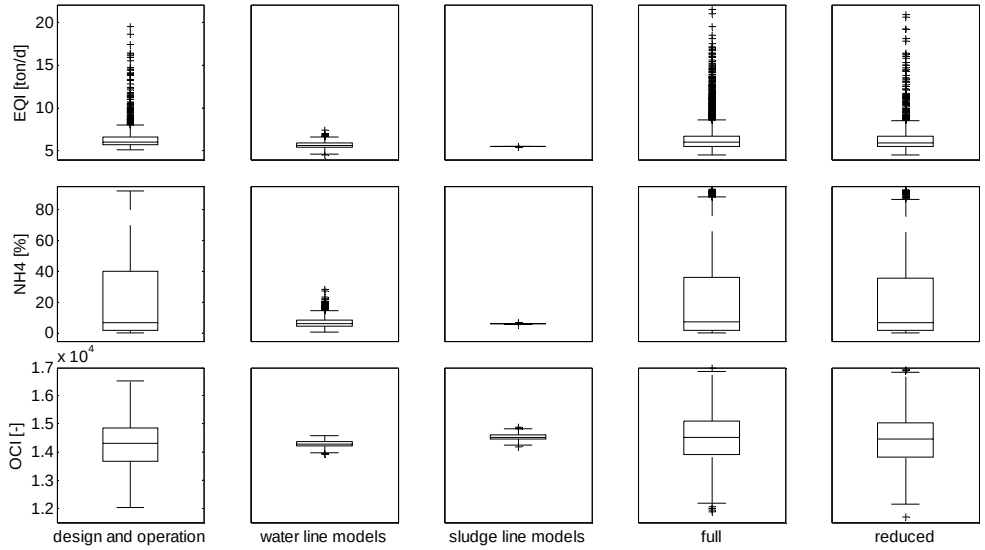


Figure 1. Variability box plots of the three BSM2 evaluation criteria for the three parameter categories separately, altogether ("full") and with the reduced set of uncertain parameters ("reduced").

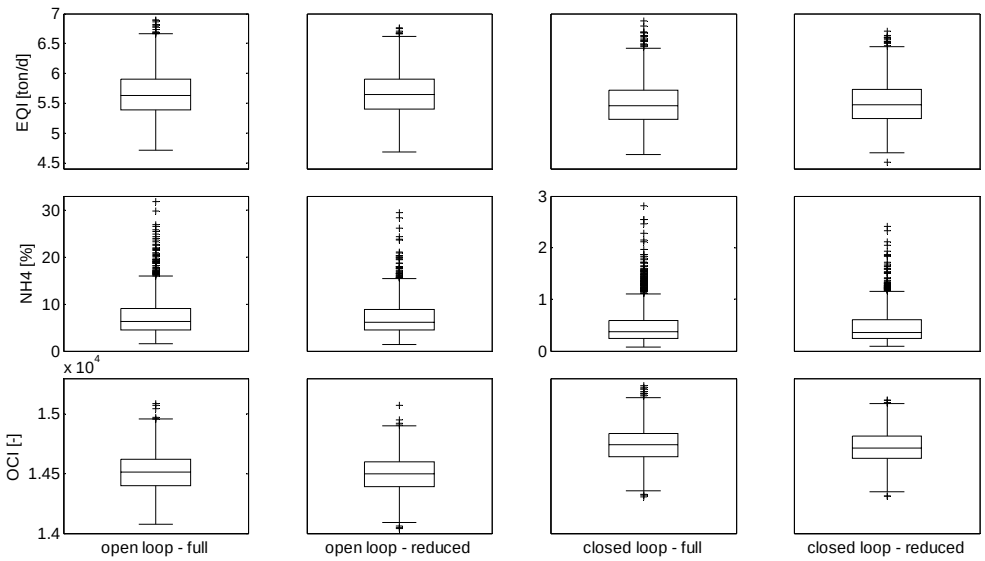


Figure 2. Variability box plots of the three BSM2 evaluation criteria for open loop and closed loop, with all wastewater and sludge treatment parameters ("full") and with the reduced set of uncertain parameters ("reduced").

4. CONCLUSIONS

Given the complexity of the BSM2 and the MC computational load, it is found useful to perform some preliminary numerical solver optimisation by means of solver setting exploration and downsampling of the output file. Proper solver selection could reduce the time required for computation by a factor of 5. This involved the use of the CVODE solver with specific settings for IterationMethod (Newton), LinearMultistepMethod (Adams), LinearSolver (SPGMR) and Accuracy (10^{-3}).

The required number of MC simulations was found to be 50 times the number of parameters to be tested.

The most sensitive BSM2 parameters belong to the design and operational group, especially for the OCI and NH_4 criteria, while for the EQI also some of the wastewater

treatment parameters are of high importance. The sludge treatment parameters have hardly any significance for the three evaluation criteria. In particular, primary settling parameters are important with respect to the economic performance of the plant.

Based on our results, the output-parameter sensitivity Jacobian can be reduced from 65 to 25 key parameters in case all parameters are considered. When a specific design and operation parameter set has to be evaluated (e.g. to assess the output variability of a control strategy), the number of wastewater and sludge treatment uncertain parameters can be reduced from 38 to 24.

These results make the execution of future sensitivity and uncertainty analysis studies more feasible.

ACKNOWLEDGEMENTS

Lorenzo Benedetti and Ingmar Nopens are post-doctoral researchers of the Special Research Fund (BOF) of Ghent University. Peter A. Vanrolleghem holds the Canadian Research Chair in Water Quality Modelling.

REFERENCES

Batstone, D.J., J. Keller, I. Angelidaki, S.V. Kalyuzhnyi, S.G. Pavlostathis, A. Rozzi, W.T.M. Sanders, H. Siegrist and V.A. Vavilin, Anaerobic digestion model no.1., IWA STR no.13, IWA Publishing, London, UK, 2002.

Batstone, D.J., P.F. Pind and I. Angelidaki, Kinetics of thermophilic, anaerobic oxidation of straight and branched chain butyrate and valerate, *Biotechnology and Bioengineering*, 84(2), 195-204, 2003.

Batstone, D.J., M. Torrijos, C. Ruiz and J.E. Schmidt, Use of an anaerobic sequencing batch reactor for parameter estimation in modelling of anaerobic digestion, *Water Science and Technology*, 50(10), 295-303, 2004.

Benedetti, L., D. Bixio, F. Claeys and P.A Vanrolleghem, Tools to support a model-based methodology for emission/immission and benefit/cost/risk analysis of wastewater treatment systems which considers uncertainties. *Environmental Modelling & Software*, 23(8), 1082-1091, 2008.

Claeys, F., D. De Pauw, L. Benedetti, I. Nopens, and P.A. Vanrolleghem, Tornado: a versatile efficient modelling & virtual experimentation kernel for water quality systems, iEMSs 2006 Conference, Burlington, VT, 2006a.

Claeys, F., P.A. Vanrolleghem and P. Fritzson, Boosting the efficiency of compound virtual experiments through a priori exploration of the solver setting space, 2006 European Modelling and Simulation Symposium (EMSS), October 4-6, Barcelona, Spain, 2006b.

Copp J.B. (ed.) The COST simulation benchmark: description and simulator manual. Office for Official Publications of the European Community, Luxembourg. 154 pages, 2002.

Hindmarsh, A.C., P.N. Brown, K.E. Grant, S.L. Lee, R. Serban, D.E. Shumaker, and C.S. Woodward. SUNDIALS: suite of nonlinear and differential/algebraic equation solvers, *ACM Transactions on Mathematical Software*, 31(3), 363-396, 2005.

Jeppsson, U., M.-N. Pons, I. Nopens, J. Alex, J.B. Copp, K.V. Gernaey, C. Rosen, J.-P. Steyer and P.A. Vanrolleghem, Benchmark simulation model no 2: general protocol and exploratory case studies, *Water Science and Technology*, 56(8), 67-78, 2007.

Forsythe, G.E., M.A. Malcolm and C.B. Moler, Computer methods for mathematical computations, Englewood Cliffs, NJ Prentice-Hall, 1977.

Henze, H., C.P.L. Jr Grady, W. Gujer, G.v.R. Marais and T. Matsuo, Activated sludge model no.1. IWA STR no.1, IWA Publishing, London, UK, 1987.

Manache, G. and C.S. Melching, Identification of reliable regression- and correlation-based sensitivity measures for importance ranking of water-quality model parameters, *Environmental Modelling & Software*, 23(5), 549-562, 2008.

Morrison, D.F., Multivariate statistical methods, Mc-Graw Hill, Singapore, 1984.

Nopens, I., D. Batstone, J.B. Copp, U. Jeppsson, E. Volcke, J. Alex and P.A. Vanrolleghem, A new interface for ASM-ADM for use in the Benchmark simulation model no.2., *Water Research* (submitted), 2008.

Otterpohl, R., M. Raak and T. Rolfs, A mathematical model for the efficiency of the primary clarification, IAWQ 17th Biennial International Conference, 24–29 July, Budapest, Hungary, 1994.

Ráduly, B., K.V. Gernaey, A.G. Capodaglio, P.S. Mikkelsen and M. Henze, Artificial neural networks for rapid WWTP performance evaluation: methodology and case study. *Environmental Modelling & Software* 22(8), 1208-1216, 2007.

Rousseau, D., F. Verdonck, O. Moerman, R. Carrette, C. Thoeye, J. Meirlaen and P.A. Vanrolleghem, Development of a risk assessment based technique for design/retrofitting of WWTPs. *Water Science and Technology* 43(7), 287-294, 2001.

Saltelli A., K. Chan and E.M. Scott (eds.), Sensitivity analysis, Wileys series in probability and Statistics, 2000.

Siegrist H., D. Vogt, J. Garcia-Heras and W. Gujer, Mathematical model for meso and thermophilic anaerobic sewage sludge digestion, *Environmental Science and Technology* 36, 1113-1123, 2002.

Takács, I., G.G. Patry and D. Nolasco, A dynamic model of the clarification thickening process. *Water Research*, 25(10), 1263-1271, 1991.

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 Sensitivity Analysis of Control Strategies using the Benchmark Simulation Model N°1 (BSM1)

Xavier Flores-Alsina[1,2], Ignasi Rodriguez-Roda[1], Gurkan Sin[2] and Krist V. Gernaey[2]

[1] *Laboratory of Environmental and Chemical Engineering, Faculty of Sciences, University of Girona, Campus Montilivi s/n 17071. Girona (Spain) {xavi, ignasi}@lequia.udg.es*

[2] *Center for BioProcess Engineering, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, DK-2800 Kgs. Lyngby (Denmark){xfa, gsi, kvg}@kt.dtu.dk*

Abstract: The objective of this paper is to perform an uncertainty analysis of the predictions of the Benchmark Simulation Model (BSM) N° 1, when comparing four activated sludge control strategies. The Monte Carlo procedure – an engineering standard, was used to evaluate the uncertainty in the predictions of the BSM1. As input uncertainty of the BSM1, the biokinetic parameters and influent fractions of ASM1 were considered, while for the model predictions the Effluent Quality (EQ) and Operational Cost (OCI) indexes were focused on. The resulting Monte Carlo simulations were presented using descriptive statistics indicating the degree of uncertainty in the predicted EQ and OCI. The Standard Regression Coefficient (SRC) method was used for sensitivity analysis to identify which input parameters influence the uncertainty in the EQ predictions the most. The results show that control strategies including an ammonium (S_{NH}) controller reduce uncertainty in both overall pollution removal and effluent total Kjeldahl nitrogen. Also, control strategies with an external carbon source reduce the effluent nitrate (S_{NO}) uncertainty, but increasing the economical costs and their variability as a trade-off. Finally, the maximum specific autotrophic growth rate (μ_A) was found responsible for causing the majority of the variance in the effluent for all the evaluated control strategies. The influence of denitrification related parameters, e.g. η_g (anoxic growth rate correction factor) and η_h (anoxic hydrolysis rate correction factor), becomes less important when a S_{NO} controller manipulating an external carbon source is implemented. These results are meaningful (and expected in a way) from a control engineering point of view: Properly tuned feedback controllers will make the process more robust towards input disturbances, attempting to maintain the process at a predefined setpoint despite input uncertainty (input disturbances), thus ensuring that the output uncertainty of the process is lower compared to for example an open-loop plant. Overall it is found useful to perform uncertainty and sensitivity analysis when comparing different control strategies based on model predictions.

Keywords: Uncertainty, sensitivity, control strategies, activated sludge plants, benchmarking, pollution removal efficiency.

1. INTRODUCTION

The benchmark simulation model (BSM) is a standardized simulation and evaluation procedure including plant layout, simulation models and model parameters, a detailed description of the disturbances to be applied during testing, and evaluation criteria for testing the relative effectiveness of simulated control strategies in activated sludge plants (Copp, 2002). Computer codes implementing the International Water Association (IWA) activated sludge model (ASM) family (Henze *et al*., 2000), secondary settling (Takacs *et al.*, 1991), and anaerobic digestion tank models (Batstone *et al*., 2002) are employed to

support the decision making on implementation of several technological alternatives such as: control strategies (Vanrolleghem and Gillot, 2002), the addition of new units (Pons and Corriou, 2002, Flores *et al*., 2007) or operational modes (Ingildsen *et al.,* 2002)

Uncertainty is a central concept when dealing with activated sludge models, whose parameters are inherently subjected to large natural variations. However, the traditional procedure for control strategy evaluation – as e.g. outlined in either the BSM1 (Copp, 2002) or the BSM2 (Jeppsson *et al*., 2006) – assumes constant rather than variable model parameters, and is thus not capable of taking into account their inherent randomness. Examples of uncertain parameters are the parameters describing the influent COD fractionation, or the parameters describing the effect of temperature or toxic compounds on the model kinetics, which will both have a significant influence on the model predictions.

The Monte Carlo procedure is an engineering standard which is commonly used for evaluating uncertainty in the predictions of simulation models (Helton and Davis, 2003). Monte Carlo simulations are based on a probabilistic sampling method of input uncertainties followed by determination and analysis of the propagation of input uncertainty to model outputs (Helton and Davis, 2003). Even though the topic of uncertainty has been dealt with before in the wastewater treatment field, see e.g. Benedetti *et al.* (2006), these studies were particularly focused on plant design rather than controller evaluation – the prime focus in this study.

The objective of this paper is to perform an uncertainty analysis of the predictions of the BSM1. Firstly, the Monte Carlo procedure is used to estimate the uncertainty in the predictions of the BSM1. As input uncertainty of the BSM1, the biokinetic parameters and the influent fractions of the ASM1 are considered, while for the model predictions the effluent quality (EQ) and operational costs (OCI) indexes are the main focus. The resulting Monte Carlo simulations were analysed using descriptive statistics indicating the degree of uncertainty in the predicted EQ. Secondly; the Standard Regression Coefficient (SRC) method is used as sensitivity analysis method to identify which input parameters are most influential on the uncertainty in the predictions of EQ.

2. METHODS

2. 1 Plant layout, implemented control strategies and evaluation criteria

The BSM1 plant layout is the activated sludge plant under study. The plant has a modified Ludzack-Ettinger configuration (see Metcalf & Eddy, 2003) with five reactors in series (tanks ANOX1 & ANOX2 are anoxic with a total volume of 2000 m^3, while tanks AER3, AER4 and AER5 are aerobic with a total volume of 4000 m^3). These are linked with an internal recycle between the 3rd aerobic (AER3) and the 1st anoxic (ANOX1) tank. The secondary settler has a surface area of 1500 m^2 and a total volume of 6000 m^3. Further details about the BSM1 design and default (open loop) operational settings can be found in Copp (2002).

Several control strategies $[A_1,..,A_j,\ldots,A_5]$ have been implemented in the activated sludge section and were compared to the default open loop base case (see **Table 1**). The oxygen (S_O) sensor is assumed to be ideal, without noise and delay. On the other hand, nitrate nitrogen (S_{NO}) and ammonium nitrogen (S_{NH}) sensor models have a delay of 10 minutes with a zero mean white noise (standard deviation of 0.1 $gN \cdot m^{-3}$).

The evaluation of the overall pollution removal of the plant is obtained by calculating the effluent quality index (EQ) (**Eq. 1)**. Compared to Copp (2002), the EQ was modified to emphasize the effect of ammonia (included in the Kjeldahl nitrogen) on the receiving water:

$$E.Q = \frac{1}{T \cdot 1000} \int_{t=7\,days}^{t=14\,days} \left(\begin{array}{l} B_{SS} \cdot TSS_e(t) + B_{COD} \cdot COD_e(t) + B_{TKN} \cdot \\ TKN_e(t) + B_{NO} \cdot S_{NO,e}(t) + B_{BOD5} \cdot BOD_e(t) \end{array} \right) Q_e(t) \cdot dt \quad \text{Eq1}$$

Table 1. Control strategies evaluated in this case study

Characteristics	3 DO	Ammonium controller	Q_{intr} controller	Q_{carb} controller
Measured variable(s)	S_o in AER1, 2 & 3	S_{NH} in AER3	S_{NO} in ANOX2	S_{NO} in ANOX2
Controlled Variable(s)	S_o in AER1, 2 & 3	S_o in AER3	S_{NO} in ANOX2	S_{NO} in ANOX2
Setpoint/critical value	2, 2 & 2 g (-COD)·m^{-3}	1 g N·m^{-3}	1 g N·m^{-3}	1 g N·m^{-3}
Manipulated variable	K_La	S_o setpoint in 3DO strategy	Q_{intr}	Q_{carb}
Control algorithm	PI	Cascaded PI	PI	PI
Applied in option	A_2, A_3,A_4 & A_5	A_4 & A_5	A_2 & A_4	A_3 & A_5

where TSS_e, COD_e, TKN_e, SNO_e and BOD_e represent, respectively, the total suspended solids, the chemical oxygen demand, the total Kjeldahl nitrogen, the nitrate nitorgen concentration and the biochemical oxygen demand in the effluent. Q_e is the effluent flowrate and T the time horizon (= 7 last days of simulation). B_{SS} = 2, B_{COD} = 1, B_{TKN} = 30, B_{NO} = 10 and B_{BOD5} = 2.

The operational costs index (OCI) is calculated adding the aeration (AE), pumping (PE), mixing (ME), chemical (CS) and sludge production (P_{sludg}) costs as states **Eq2**

$$OCI = AE + PE + ME + 3 \cdot CS + 5P_{sludg} \quad \text{Eq2}$$

2. 2. Uncertainty analysis with Monte Carlo technique

For notational convenience, the BSM1 model is represented by f, the output vector by y, the state vector by x, the input variables by u, the input parameter vector by θ and time is represented by t (see details in **Eq 3**)

$$y(t,\theta) = f(x,u,t,\theta) \quad \text{Eq3}$$

Monte-Carlo analysis of uncertainty involves 3 steps: (1) specifying input uncertainty (2) sampling input uncertainty and (3) propagating the sampled input uncertainty through f to obtain output uncertainty for y. An expert review process was used to define the input uncertainty around the biokinetic parameters and influent fractions of ASM1, while for step 2 we have used Latin Hypercube Sampling (LHS) (McKay *et al.*, 1979).

To carry out this analysis, the uncertainty associated to the ASM1 parameters [U = U_1,…,U_y,,…U_{32}] was characterized by a set of probability distributions [D = D_1,…,D_y,,…D_{32}]. These distributions were assumed to characterize a degree of belief with respect to where the appropriate values for the elements of [U] are located for use in the simulation of the BSM1. When used in this manner, these distributions are providing a quantitative representation of what is referred as subjective or epistemic uncertainty (Helton and Davis, 2003).

Three uncertainty classes are distinguished [C = C_1, C_2, C_3] to allow the representation of the parameter uncertainty in a structured way, and each uncertainty parameter U_y was assigned to a certain class C_c depending on the extent of knowledge available in the literature about this specific parameter value. The first class corresponded to low uncertainty and included mostly stoichiometric parameters. In this class (C_1), the parameters were assumed to have a 5 % upper and lower bound around their default values [U_1,…, U_{10}]. The second class (C_2), corresponded to medium uncertainty and involves kinetic parameters such as the maximum specific growth rate and the affinity constants [U_{11},…, U_{24}]. In this class, 25 % upper and lower bounds around the default values were assumed. For simplification, all the kinetic and stoichiometric parameters were supposed to be independent although the authors are aware of possible correlations amongst several parameters e.g. the maximum specific growth rate and the half saturation constants. **Table 2** summarizes these parameters.

Table 2. An expert review of input uncertainty of biokinetic parameters of ASM1 including default parameter values, uncertainty class and the corresponding variation percentage

Uncertainty parameter	Symbol	Default value	Class	Δ	Units
autotrophic yield	Y_H	0.67	1	0.067	$gCOD \cdot gN^{-1}$
heterotrophic yield	Y_A	0.24	1	0.024	$gCOD \cdot gCOD^{-1}$
fraction of biomass to particulate products	f_P	0.08	1	0.008	Dimensionless
fraction of nitrogen in biomass	i_{XB}	0.08	1	0.008	$gN(gCOD)^{-1}$ in biomass
fraction of nitrogen in particulate products	i_{XP}	0.06	1	0.006	$gN(gCOD)^{-1}$ in X_P
conversion from COD particulates	X_{I2TSS}	0.75	1	0.075	$gTSS.(gCOD)^{-1}$
conversion from COD particulates	X_{S2TSS}	0.75	1	0.075	$gTSS.(gCOD)^{-1}$
conversion from COD particulates	X_{BH2TSS}	0.75	1	0.075	$gTSS.(gCOD)^{-1}$
conversion from COD particulates	X_{BA2TSS}	0.75	1	0.075	$gTSS.(gCOD)^{-1}$
conversion from COD particulates	X_{U2TSS}	0.75	1	0.075	$gTSS.(gCOD)^{-1}$
maximum specific heterotrophic growth rate	μ_H	4.00	2	2.00	day^{-1}
half saturation (hetero. growth)	K_S	10.00	2	5.00	$gCOD.m^{-3}$
half saturation (hetero. oxygen)	K_{OH}	0.20	2	0.10	$gCOD.m^{-3}$
half saturation (nitrate)	K_{NO}	0.50	2	0.25	$gN.m^{-3}$
heterotrophic specific decay rate	b_H	0.30	2	0.15	day^{-1}
maximum specific autotrophic growth rate	μ_A	0.50	2	0.25	day^{-1}
half saturation (auto. growth)	K_{NH}	1.00	2	0.50	$gN.m^{-3}$
half saturation (auto. oxygen)	K_{OA}	0.40	2	0.20	$gCOD.m^{-3}$
autotrophic specific decay rate	b_A	0.05	2	0.025	day^{-1}
anoxic growth rate correction factor	η_g	0.80	2	0.40	Dimensionless
amonification rate	k_a	0.05	2	0.025	$m^3(gCOD.day)^{-1}$
maximum specific hydrolysis rate	k_h	3.00	2	1.50	$gX_S(gX_{BH}COD \cdot day)^{-1}$
half saturation (hydrolysis)	K_X	0.10	2	0.05	$gX_S(gX_{BH}COD)^{-1}$
anoxic hydrolysis rate correction factor	η_h	0.80	2	0.40	Dimensionless

The third class of uncertainty (C_3) corresponded to high uncertainty and included the influent fraction related parameters, assuming upper and lower bounds equal to 50 % of the default parameter values (results not shown). Several class 3 uncertainty factors (from 0.5 to 1.5) were applied to the COD and N fractions obtained from the influent file included in the BSM1 used to calculate the different ASM1 influent organic and inorganic matter state variables. The different fractions are represented as α_i where i is the component e.g. soluble organic matter S_s, inorganic soluble matter S_i. A similar method was applied to influent nitrogen *[U_{29},…,U_{32}]*, where the fraction coming from particulate products and biomass was removed first, to finally obtain the inorganic (ammonium, S_{NH}) and organic influent nitrogen compounds concentrations (either soluble or particulate, S_{ND} and X_{ND}).

The input uncertainty space was sampled using the LHS method (McKay *et al*., 1979; Iman *et al*., 1981). In this study, 1000 samples $[U_y = U_{y,1},\ldots,U_{y,f},,\ldots U_{y,1000}]$ were generated to ensure that the input uncertainty space was covered uniformly. Each Latin hypercube sample contains one randomly selected value $U_{y,f}$ from each of the previously defined probability distributions D_y. The Monte Carlo simulations were performed by evaluating the BSM model for each one of the generated Latin hypercube samples, solving the entire model and quantifying the defined EQ criterion for each tested alternative [A]. The solution of the model for each parameter combination resulted in a distribution of possible values for the defined evaluation criteria (**Eq 1** and **2**). These distributions reflected the possible variation of the performance criterion taking into account the input uncertainty.

2.3. Sensitivity Analysis

The sensitivity analysis phase involves the construction (see **Eq 4**) of a regression model that reveals the relationships between the elements associated to the input uncertainty $[U = \{U_1,\ldots,U_k,\ldots,U_{32}\}]$ and the elements associated to the output uncertainty i.e. in this case EQ (Saltelli *et al*., 2005).

$$X_j = b_0 + \sum_{k=1}^{nS} b_k U_k \qquad \text{Eq4}$$

The higher the value of the regression coefficient (b_k) the stronger the relationship is, i.e input uncertainty values are auto-scaled. The absolute values of the regression coefficients are ranked and categorized by *k-means* clustering (Hair *el al*., 1998).

3. RESULTS

All dynamic simulations were preceded by a steady state simulation to ensure an appropriate starting point for the dynamic simulations and to eliminate bias due to the selection of the initial conditions on the dynamic modelling results (Copp, 2002). Even though the length of the dynamic influent file used to carry out the simulations was 28 days, only the data generated during the last seven days were used to evaluate the plant performance.

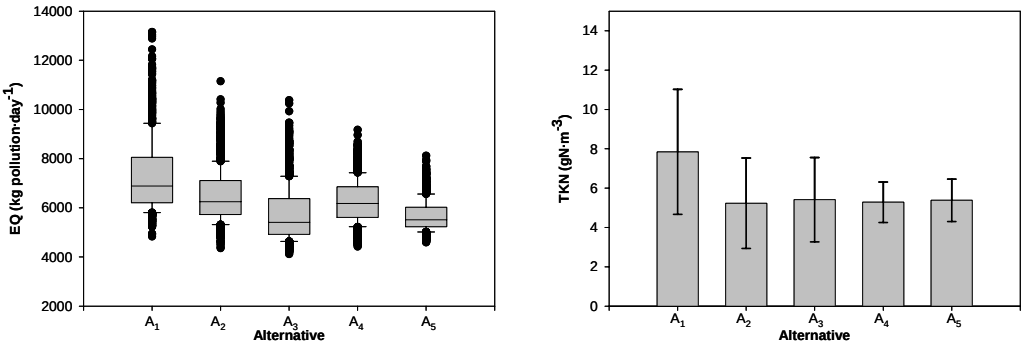


Figure 2. Representation of the output uncertainty for the evaluated control strategies using multiple box plots for EQ (left) and error bar charts for the effluent TKN (right)**.**

The results of the uncertainty analysis are shown using multiple box-plots for the EQ after running 1000 Monte Carlo simulations for each of the control strategies (**Figure 2**). From the results depicted in **Figure 2** (left), it can be noticed that there is a clear difference between the alternatives with (A_2, A_3, A_4 and A_5) and without (A_1) controllers in terms of both absolute value and variance. Also one observes the effect of the S_{NH} controller (A_4 and A_5) in reducing even more the variation of the potential impact on the receiving water. This is mainly due to the continuous adaptation of the S_O set-point via the cascade controller which ensures a constant effluent S_{NH} level as shown by the error bar chart in **Figure 2** (right)**.**

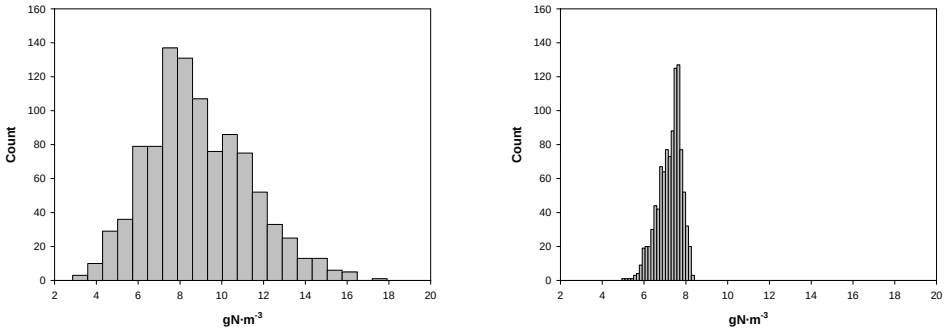


Figure 3. Histograms of the effluent S_{NO} when alternative A_1 (left) and A_3 (right) are evaluated under uncertainty

Also, it is important to point out the effect of the external carbon source controller on the overall nitrogen removal. The external carbon source addition provides the limited readily biodegradable substrate for denitrification, thereby enhancing the total nitrogen removal by improving the reduction of the nitrate to nitrogen gas. In this way, the external carbon controller ultimately decreases the variation of the effluent S_{NO} as shown by the histograms of **Figure 3**. Again, this result was expected for a properly functioning controller, which will make the system more robust and thus reduce the output uncertainty.

Table3. Mean and standard deviation of the operation costs breakdown for the different generated alternatives under uncertainty

	A_1	**A_2**	**A_3**	**A_4**	**A_5**
AE	3495,75	3489,23	3630,63	3773,32	4047,59
	0,00	218,03	220,84	631,55	681,66
PE	388,17	250,20	388,17	266,12	388,17
	0,00	48,82	0,00	57,38	0,00
3·CS	0,00	0,00	2652,46	0,00	2455,84
	0,00	0,00	898,97	0,00	1109,36
5·SP	12198,21	12228,04	13579,12	12207,36	13511,42
	912,38	905,39	1200,85	1057,24	1260,23
OCI	16322,13	16207,47	20490,37	16486,33	20643,02
	912,38	779,79	1819,72	1229,88	2504,78

The results of **Table 3,** where it is represented the mean and standard deviation of the breakdown of the operational costs; demonstrate a clear difference between the control strategies with and without external carbon source addition. The periodic purchase of an external carbon source implies a subsequent increase of both quantity and variation of the sludge production and the overall operating cost index. Hence, it can be said that the addition of external carbon source reduces the effluent nitrate and its variability as a trade-off to an increase of the operation costs and their variability. ME doe not present variation from one alternative to the others and thus not included in the table. Control strategies with a S_{No} controller manipulating the internal recycle and S_{NH} manipulating the oxygen setpoint have higher variability in pumping and aeration energy because the controllers in order to maintain the nitrate concentration in ANOX2 and AER3 to the desired setpoint $1gN \cdot m^{-3}$. From a control engineering point of view, this simply tells that the controller does fulfil its function which is to maintain a stable process performance despite input disturbances, meaning less variability in the effluent compared to the open-loop case.

Table 4. Summary of sensitivity analysis results: Standardized Regression Coefficients (SRC) of parameters clustered into three classes for all the evaluated control strategies

	A_1	**A_2**	**A_3**	**A_4**	**A_5**
Strong	μ_A	μ_A	μ_A	μ_A	μ_A
Medium	b_A, K_{OA}, b_H, K_{NH}, η_g, Y_H and η_h	b_A, α_{SNH}, α_{XND}, K_{OA}, K_{NH} and η_g,	b_A, α_{SNH}, α_{XND}, K_{OA} and K_{NH}	b_A, η_h, η_g, Y_H, K_{OA}, K_{NH} and K_{OH}	b_A, α_{SNH}, K_{NH}, α_{XND} and K_{OA}

Using the Monte Carlo simulation results for EQ, the standardized regression coefficients for each input were calculated, and are presented in **Table 4** for each control strategy. From the results generated during the sensitivity analysis one observes that the parameters related to the nitrification process, e.g. μ_A, and the denitrification process, e.g. η_g, are identified as the most influential. This is understandable since TKN (30) and NO (10) are the most important factors in the EQ calculation. Moreover, it is generally known that the BSM1 plant (Copp, 2002) is highly loaded with nitrogen, meaning that any change in ASM1 parameters influencing nitrogen removal will be noticed in the EQ. In **Figure 4**, a bi-plot shows the correlation between the EQ (**Eq. 1**) and the most influential factor found for all the evaluated control strategies (μ_A).

On the other hand, for those control strategies with an external carbon source addition (A_3 and A_5), denitrification related parameters are classified as weak, i.e. non-influential. This fact is attributed to the improvement of the overall nitrate removal rate due to the extra carbon source added by this controller, thereby reducing the importance of this denitrification process parameter. In other words, the additional input of readily biodegradable organic matter makes the system less dependent on the anoxic correction factors for both hydrolysis (η_h) and growth rate (η_g), thus decreasing the variability of the effluent S_{NO}.

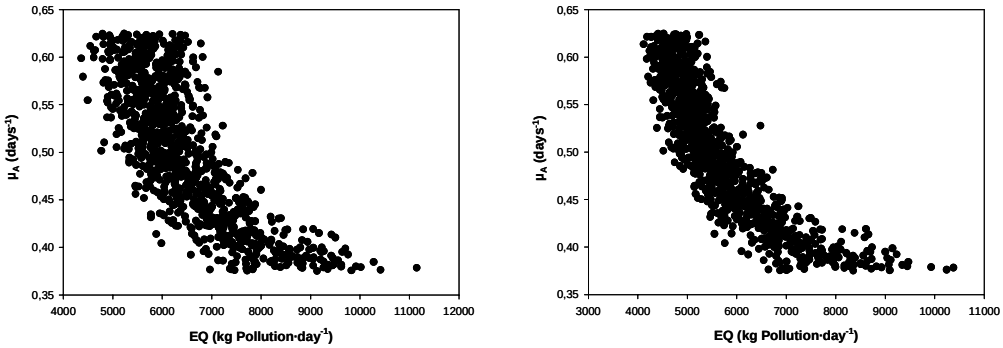


Figure 4. Correlation between the output (EQ) and input (μ_A) for alternatives A_2 (left) and A_3 (right).

It is important to point out that despite the apparent advantages of the formal assessment of the uncertainty, one should be aware that the conclusions arising from this case study considering uncertainty can only be as good as the underlying assumptions. Thus, the results of the uncertainty analysis will to a large extent depend on the characteristics of the defined uncertainty distributions, similar to a base case performance i.e. without uncertainty, where the results will depend on the model selection

4. CONCLUSIONS

This paper has presented the results of performing uncertainty and sensitivity analysis of several activated sludge control strategies using the BSM1. Different combinations of oxygen, nitrate and ammonium controllers were evaluated considering uncertainty in the ASM1 parameters. The input uncertainty, defined by probability distributions, was propagated trough the BSM1 using a Monte Carlo approach which quantified the uncertainty (probability distribution) in the model predictions such as the effluent quality index.

From the evaluated controllers, alternatives with an ammonium controller (Alternatives A_4 and A_5) reduced the uncertainty in the variation of nitrogen removal efficiency mainly due to an improvement in the simulated nitrification rates. The alternatives with a nitrate controller (A_3 and A_5) also reduced the effluent nitrate and its variation, but increasing the operation costs and their variability as a trade-off. This was mainly due to the added carbon source acts as an extra electron donor enhancing the total nitrogen removal by improving the reduction of the produced nitrate to nitrogen gas.

Finally, the results of the sensitivity analysis revealed that the autotrophic specific growth rate (μ_a) is the parameter with the strongest influence on the effluent quality variability. This was attributed to the relatively high weight of TKN in the EQ calculation and to the fact that the plant is highly loaded in N. Also, denitrification related parameters became less important when an external carbon source was added because the additional input of readily biodegradable organic matter made the system less dependent on the anoxic correction factors for hydrolysis (η_h) and growth rate (η_g)

The uncertainty analysis results showed input uncertainty may propagate differently to model outputs when applying different control strategies. While some controller type ensures the effluent ammonium concentration is less variable (low uncertainty), the other ensures the effluent nitrate is less variable (less uncertainty). Hence making a decision about which control strategy is better becomes a subject of multi-objective weighting. All in all uncertainty analysis is certainly deemed useful as it provides a more realistic interpretation of the model simulations: hence better informed decisions can be made.

REFERENCES

Batstone D.J., Keller J., Angelidaki I., Kayuznyi S.V., Pavlostathis S.G., Rozzi A., Sanders W.T.M., Siegrist H. and Vavilin V.A. Anaerobic Digestion Model No 1. IWA STR No 13, IWA Publishing, London, UK, 2002.

Benedetti L., Bixio D., Vanrolleghem P.A. Assessment of WWTP design and upgrade options: balancing costs and risks of standards' exceedance. Wat. Sci. Tech. 54(6-7):371. 2006.

Copp, J.B. The COST Simulation Benchmark: Description and Simulator Manual. Office for Official Publications of the European Community, Luxembourg. 2002.

Flores X., Poch M., Rodríguez-Roda I., Bañares-Alcántara R. and Jiménez L. Systematic procedure to handle critical decisions during the conceptual design of activated sludge systems. Ind. Eng. Chem. Res. 46(17):5600. 2007.

Helton J.C. and Davis F.J. Latin Hypercube Sampling method and the propagation of uncertainty in analyses of complex systems. Reliab. Engng. Syst. Saf. 81:23. 2003.

Hair J. Fr., Andersen R.E., Tatham R.L. and Black W.C: Multivariable Data Analysis. Fifth Edition. Prentice-Hall. London. 1998.

Henze M., Gujer W., Mino T. and van Loosdrecht M.C.M. Activated Sludge Models ASM1, ASM2, ASM2d and ASM3. IWA Scientific and Technical Report No. 9, IWA Publishing, London, UK, 2000.

Iman R.L., Helton J.C. and Campbell J.E.. An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment. J. Qual. Techno1. 13(3):174. 1981.

Ingildsen P., Olsson G. and Yuan Z. A hedging point strategy - balancing effluent quality, economy and robustness in the control of wastewater treatment plants. Wat. Sci. Tech. 45(4-5):317. 2002

Jeppsson U., Rosen C., Alex J., Copp J., Gernaey K.V., Pons M.N. and Vanrolleghem P.A. Towards a benchmark simulation model for plant-wide control strategy performance evaluation of WWTPs. Wat. Sci. Tech. 53(1):287. 2006

McKay M.D., Conover W.J. and Beckman R.J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21:239. 1979.

Metcalf and Eddy. Wastewater Engineering: Treatment, Disposal and Reuse (3rd edition). New York: Mc-Graw-Hill. Inc., 2003.

Pons M.N and Corriou J.P. Implementation of storage tanks on the COST 624 benchmark, Wat. Sci. Tech. 45(4-5):159. 2002.

Saltelli A. Sensitivity analysis in practice. A guide to assessing scientific models. Wiley. Chichester. UK.2004.

Takács I., Patry G.G. and Nolasco D. A dynamic model of the clarification-thickening process. Wat. Res. 25(10):1263. 1991.

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 data availability influence on integrated urban drainage modelling uncertainty

G. Freni[a], G. Mannina[a], G. Viviani[a]

[a] *Dipartimento di Ingegneria Idraulica ed Applicazioni Ambientali, Università di Palermo, Viale delle Scienze, 90146 Palermo, ITALY (freni/mannina/gviv@idra.unipa.it)*

Abstract: In urban water quality management, several models are connected and integrated for analysing the fate of pollutants from the sources on the urban catchment to the final recipient; classical problems connected with the selection and calibration of parameters are amplified by the complexity of the modelling approach increasing their uncertainty. The present paper aims to study the influence of dataset extension on the modelling response uncertainty with respect to the different integrated modelling outputs (both considering quantity and quality variables). At this scope, a parsimonious home-made integrated modelling approach has been used allowing for analysing the combinative effect between sewer system, treatment plant and receiving water body; the uncertainty analysis approach has been applied to an experimental catchment in Bologna (Italy). The number of available data points has been fictitiously reduced obtaining datasets with extension ranging between 25% and 100% of the measured data. For each of the datasets, uncertainty analysis has been performed and its propagation from upstream submodel to the downstream ones has been assessed. The results are interesting and show a strong influence of dataset extension on model uncertainty.

Keywords: Environmental modelling, Integrated urban drainage systems, Uncertainty analysis, Receiving water body, Waste water treatment plant.

1. INTRODUCTION

Nowadays, water quality improvement, mainly focused on the overall management of river basins, is gaining more importance during the last years [Blöch, 1999]. As a result of the publication of the European water framework directive (WFD), integrated management of river basins has become an important research area [Lindenschmidt, 2006]. The main elements of integrated urban water system are: catchment area and sewer system (SS); waste water treatment plant (WWTP); receiving water body (RWB). Dry weather flow along with surface runoff is transported from the catchment area to the WWTP for the removal of pollutants and subsequent release into the receiving water. Only if the amount of runoff exceeds the given hydraulic capacity of the plant, a mixture of wastewater and rainwater is discharged to the receiving water directly.

Indeed, due to the fact that integrated approaches are basically a cascade of sub-models (simulating SS, WWTP and RWB), uncertainty produced in one sub-model propagates to the following ones depending on the model structure, the estimation of parameters and the availability and uncertainty of measurements in the different parts of the system. For this reason, an important issue in integrated modelling is connected with the balance between sub-models so that each of them will not introduce too much uncertainty depending on the available field data. Furthermore, such problem, coupled with the consequent uncertainty

accumulation, may lead to get very uncertain and sometimes useless results [Willems, 2000].

Regarding field data for integrated modelling calibration/validation, a limited data base is generally available especially looking at the water quality one [Vanrolleghem et al., 1999]. Indeed, this consideration depends on the fact that contemporary monitoring campaigns on the different compartments of the integrated urban water system are complex and they require large technical and economic efforts. The situation is even worse when looking at wet weather impact on RWB because of the small scale at which polluting impacts take part adding more difficulties in providing reliable monitoring campaigns. As a matter of fact, several Authors highlight data availability as one of the main limitation to integrated modelling application and they adopt hypothetic or semi-hypothetic case studies where the different parts of the system are simulated using data coming from real and well documented case studies not really linked together [Schütze et al. 1999; Rauch and Harremoës, 1996].

The present paper aims to study the dataset extension influence on the modelling response uncertainty and, at this scope, a parsimonious home-made integrated model has been used allowing for analysing the combinative effect between SS, WWTP and RWB [Mannina et al., 2004, 2005]. In details, the uncertainty analysis has been applied to an experimental catchment in Bologna (Italy).

The dataset has been fictitiously reduced in the extension by uniformly decreasing the number of data points available for each of the monitored events as it will be better discussed in the following paragraphs.

2. MODELS AND METHODS

2. 1 The adopted model

In order to carry out the survey, a home-made quality-quantity integrated urban drainage model has been employed. For sake of conciseness, in the following will be discussed only the model structure remanding to the literature for further details Mannina et al. [2004] and Mannina [2005]. The model is able to estimate both the interactions between the different systems (SS, WWTP and RWB) and the modifications, in terms of quality, that urban stormwater causes inside the RWB. Such a system is made up mainly of three sub-models:

- the rainfall-runoff and flow propagation sub-model, which is able to evaluate the quality - quantity features of SS outflows;
- the WWTP sub-model, which is representative of the treatment processes;
- the RWB sub-model that simulates the pollution transformations inside the river.

The first sub-model, reproducing the physical phenomena which take place both in the catchments and in the sewers, allows to determine the hydrograph and pollutograph in the sewer. This sub-model is divided into two connected parts: a hydrological - hydraulic module, which calculates the hydrographs at the inlet and at the outlet of the sewer system, and a water quality module, which calculates the pollutographs at the outlet for different pollutant species (TSS, BOD and COD). The hydrological - hydraulic module starts to evaluate the net rainfall, from the measured hyetograph, by a loss function (taking into account surface storage and soil infiltration). From the net rainfall, the model simulates the net rainfall-runoff transformation process and the flow propagation with a cascade of one linear reservoir and a linear channel (representing the catchment) and a linear reservoir (representing the sewer network). The solid transfer module reproduces the build-up and wash-off of pollutants from the catchment and the propagation of solids in the sewer network considering also their sedimentation and re-suspension.

The second sub-model is aimed to the analysis of WWTP during both dry and wet weather periods. The WWTP inflow has been computed taking into account the presence of a CSO device and its efficiency. The WWTP sub-model simulates the behaviour of the part of the

plant composed by an activated sludge tank and a secondary sedimentation tank. For the activated sludge tank model, mass balance equations derived from Monod's theory have been used in order to reproduce BOD removal. On the other hand, the sedimentation tank has been simulated using the solid flux theory according to the methodology proposed by Vitasovic et al. [1997]. In particular, the solids concentration profile has been obtained by dividing the settler into 50 horizontal layers of constant thickness. Within each layer the concentration is assumed to be constant and the dynamic update is performed by imposing a mass balance for each layer. Further, the sedimentation model proposed by Takács et al. [1991] has been employed in order to evaluate the different types of sedimentation.

The third sub-model examines the assessment of RWB. The simplified form of the De Saint Venant equation (cinematic wave) is used for the quantity module and water de-oxygenation and re-oxygenation phenomena are simulated by using the advection - dispersion equation; this approach allows for evaluating the effects of stormwater on the RWB, both at a single event scale and during long term simulation. The pollutant sources, which cause a worsening in the characteristics of RWB, are mainly two: the WWTP continuous discharges and the intermittent discharges coming from the combined storm overflows (CSOs). The sub-model also allows for simulating the effects of the insertion of CSOs control measures for temporary accumulation of stormwater during the rainfall event.

2.2 Uncertainty analysis based on Generalised Likelihood Uncertainty Estimation (GLUE) approach

The GLUE methodology [Beven and Binley, 1992] is aimed to the identification of modelling uncertainty related to parameter estimation. In the present study, it will be used for assessing the effect of imperfect knowledge of the system on this specific kind of uncertainty. System knowledge will be fictitiously "altered" by reducing the number of available data points in the measurements dataset: in each system compartment (SS, CSO and RWB), data availability will be reduced to 25%, 50% and 75% respectively. Imperfect knowledge in each system compartment will be analysed separately in order to understand the weight to be given to measuring campaigns in the parts of the system. For applying the GLUE methodology, the model has been solicited by randomly sampled parameter sets throughout Monte Carlo simulations. By means of a likelihood measure, E, parameter sets are classified and sets with poor likelihood weights, with respect to a user-defined acceptability threshold (Tr), are discarded as "non-behavioural". All parameters sets coming from the behavioural simulation runs are retained and their likelihood weights are re-scaled so that their cumulative total sum is equal to 1. The likelihood measure, E, represents the ability of the model to fit real data. On the other hand, the acceptability threshold, Tr, represents a user-defined critical value indicating the minimum value of E that each modelling simulation should have to be representative of the model behaviour with respect to the analysis aim. Tr is usually set equal to zero. In the present study, the Nash and Sutcliffe efficiency index has been used as likelihood measure [Nash and Sutcliffe, 1970]. Nash-Sutcliffe likelihood measure is analytically defined in the range $]-\infty; 1]$ although behavioural simulations are defined in the range $]Tr; 1]$.

Treating the distribution of likelihood values as a probabilistic weighting function for the predicted variables, it is possible to assess the uncertainty associated with the predictions, conditioned on the definition of the likelihood function, of the input data and model structure. A method of deriving predictive uncertainty bands using the likelihood weights from the behavioural simulations has been shown by Beven and Binley [1992]. The uncertainty bands are calculated using the 5% and 95% percentiles of the predicted output likelihood weighted distribution. Wider bands mean higher uncertainty in the estimation of the modelling output and thus lower confidence in the model results.

Leaving details to literature [among others: Beven and Binley, 1992; Gupta et al., 2005; Freni et al., 2008a], GLUE procedure can be applied according to the following steps:

1. "Uncertain" parameters are selected: the adopted integrated model is controlled by 27 parameters. As pointed out in other studies, the model is not sensitive to some of them in the presented case study and, thereafter, nineteen parameters have been considered in

the present study [Freni et al., 2008b]. Their brief explanations are provided in previous studies [Mannina, 2005; Freni et al., 2008a, b].

2. variation ranges are selected for each model parameter which can be considered affected by uncertainty: the variation range for each of the selected parameters has been obtained by the calibration of the five fully monitored events [Freni et al., 2008b]; more specifically, initially wide ranges have been assigned for uncertain parameters based on both their physically feasible ranges and previous literature researches. Thereafter, five calibration processes has been carried out based on data coming from single fully monitored events. Parameters variation ranges were thus refined limiting them between the maximum and the minimum value obtained on the different calibration processes.
3. random sets are drawn from the parameter variation ranges by Monte Carlo sampling according to an *a priori* distributions: in this case, the uniform distribution has been assumed not having adequate prior knowledge for selecting another distribution; correlation between parameters, that may affect their mutual distribution, has been neglected because no evidence has been reported in the model calibration process showing the relevance of such aspect [Freni et al., 2008b].
4. the model is run for each random parameter set and likelihood measure (in this study, Nash-Sutcliffe efficiency index) is calculated for all analysed modelling variables (discharges and water quality variables at the SS ending pipe, at CSO outlet and at the RWB closing cross-section); if likelihood measure is positive, the model simulation is considered "behavioural" and saved for the following analysis steps;
5. step 2 is repeated until a user-specified number of behavioural simulations is reached (the present study is based on 1,000 behavioural simulations) linking the parameter set and the correspondent modelling outputs to a likelihood measure value;
6. analogously to the procedure used in statistics for obtaining cumulated probability from probability density, likelihood measures can be cumulated for each parameter and for each modelling output obtaining cumulated likelihood distributions;
7. for each modelling output, 5% and 95% percentiles cumulated likelihood distribution represent the uncertainty bands; for each model parameter, 5% and 95% percentiles represent the region in the parameter space where most likely the "real" parameter values should be (with a confidence level equal to 0.1);

3. THE CASE STUDY

The integrated model and the uncertainty analysis have been applied to the catchment of the Savena river (Italy). The sewer system and the river studied in this work concern a part of the sewer network of Bologna, studied within the European Union research project INNOVATION 10340I [Artina et al., 1999]. The studied river reach is long about 6 km and it receives discharge from 6 CSOs deriving from the Bologna sewer network. The sewer network is a part of the combined system serving the whole city of Bologna, which can be considered as hydraulically divided into many independent catchments, all connected to a WWTP. The part of Bologna connected to the studied river has an area of more than 450 ha, with an impervious percentage of about 66% and about 60,000 inhabitants. During experimental survey, carried out within the INNOVATION European Research Project, from December 1997 to July 1999, about 50 events have been recorded, but, for only 5 of these, water quality aspects have been analysed regarding both RWB and SS. The monitoring infrastructure consisted of 3 raingaues, 8 sonic level gauges (6 in the drainage system and 2 in the receiving river) and 6 automatic 24-bottles sampler (3 in the sewer system and 3 in the river). The study has been focused on BOD_5, TSS, COD and DO, even if analogous considerations may be extended to other parameters. In this study, only a part of the Savena river has been simulated (400 meters downstream the CSO No. 6) because the contribution of this CSO to river pollution has been determined much more relevant in respect to all the others. The contribution of other polluting sources has been considered by monitoring river pollution load in the first cross-section upstream of CSO

No. 6 and introducing this information as input in the models. Savena is an ephemeral river since there are wide flow variations during the different seasons and the river base flow is comparable with the CSO discharge. Further details on the monitoring campaign can be found in Artina et al. [1999].

4. RESULTS ANALYSIS

The comparison among system knowledge scenarios will be provided according to three criteria: the effect on parametric uncertainty will be assessed by comparing uncertainty bands for different modelling outputs; the effect on single parameters weight in uncertainty propagation will be identified by analysing likelihood distributions; the effect on model calibration effectiveness will be derived from modelling efficiency distributions. Figure 1 shows uncertainty bands obtained the RWB downstream cross-section reducing data availability reducing data availability in different upstream sub-models: data availability is reduced in one sub-system per time so trying to separate the influence of each sub-system in the overall modelling uncertainty. Uncertainty bands width demonstrates that SS sub-model and RWB sub- model have higher importance in uncertainty propagation.

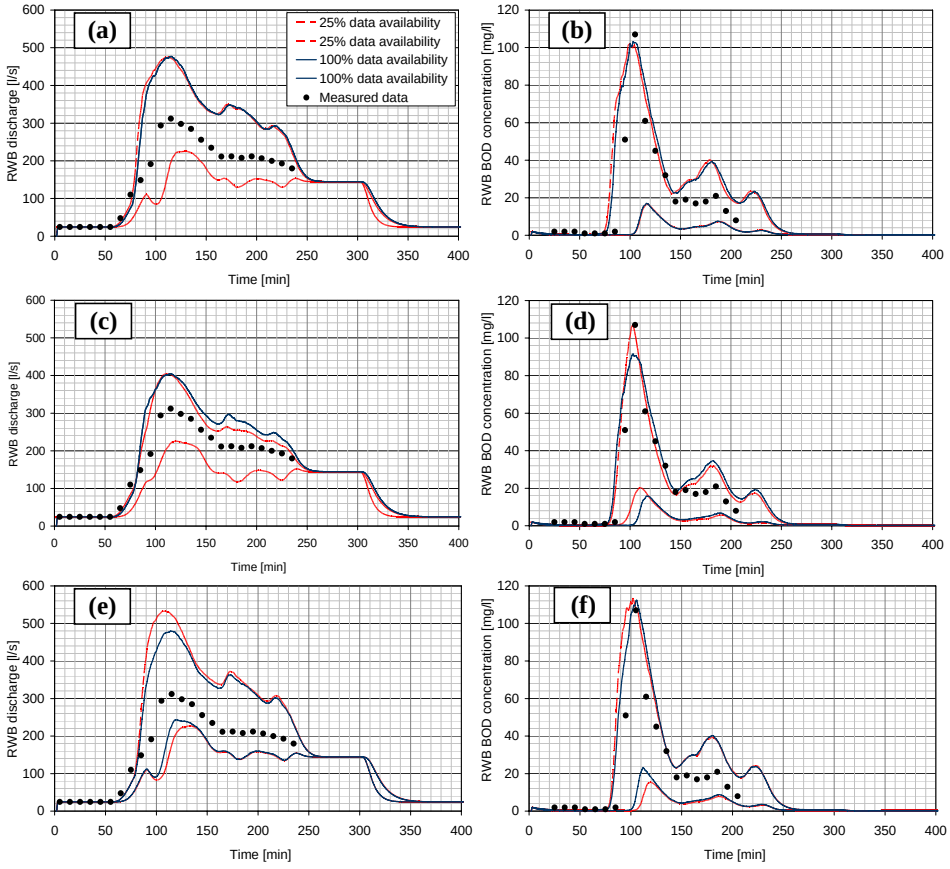


Figure 1. Uncertainty bands in the RWB modelling outputs obtained reducing data availability in SS sub-model (a-b), in CSO sub-model (c-d) and in RWB sub-model (e-f).

Uncertainty bands generated during CSO analysis (Figures 1c-1d) are generally smaller independently from data availability. Data availability in SS sub-model practically does not affect modelling output uncertainty (Figures 1a-1b); this behaviour can be explained by the uncertainty mitigation effect provided by the two downstream sub-models that can supply to the imperfect SS knowledge with their data. The effect of data availability is progressively higher moving to downstream sub-models (Figures 1c-1d and 1e-1f):

reducing data availability to 25% percent of measured data, uncertainty bands maximum width is increased by 30% for RWB discharge and 20% for BOD concentrations.

Data availability may affect the importance of single parameters in modelling uncertainty propagation: one parameter, that can greatly affect modelling efficiency when all measured data are available, can lose its importance when data availability is lower thus contributing to produce scattered modelling output and consequently increase uncertainty. The change in parameters behaviour can be analysed by the mean of likelihood distributions (Figure 2). The analysis of all parameters demonstrated that data availability does not affect parameters behaviour in terms if modelling uncertainty contribution. Figure 2 shows likelihood distributions for some of the modelling parameters adopted as examples. In particular, in Figure 2, the following model parameters are reported: the initial hydrological losses (W_0), the catchment runoff coefficient (Φ), the catchment reservoir constant (K_1), the decay rate in the Alley-Smith model (Disp), the CSO dilution factors (R_{d1}-R_{d2}), the river bed roughness (Gauckler–Strickler) (K_s), the deoxigenation coefficient (Kd) and the sediment oxygen demand (ksed).

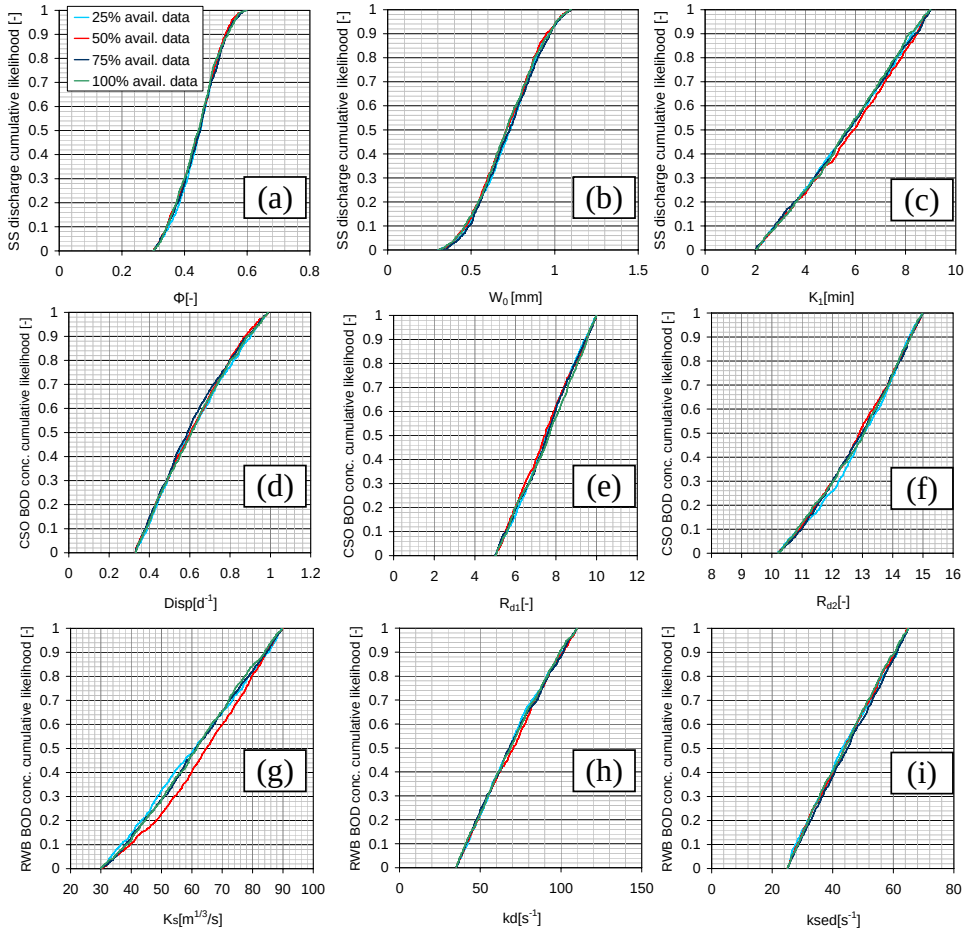
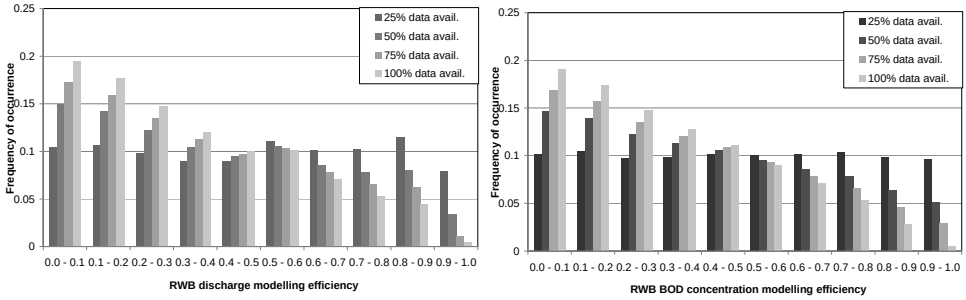


Figure 2. Cumulated likelihood distributions for some relevant parameters in SS sub-model (a-c), in CSO sub-model (d-f) and in RWB sub-model (g-i).

The distributions obtained for all data availability scenarios are substantially superimposed with some differences only in RWB sub-model that justify the global increase in modelling uncertainty presented in Figure 1.

Another important effect of data availability has been detected by Beven and Freer [2001]: the basic idea of GLUE is that given our inability to represent exactly in a mathematical model how nature works, there will always be several different models that mimic equally well an observed physical process (such as sewer discharge or pollutant concentration).

Such concept takes to model equifinality (more than one model structure can fit the monitored reality) and to model parameters identifiability problems (once fixed the modelling structure, several sets of parameters are able to give similar model outputs). Such behaviour can be amplified by lack of data for efficient characterization of over-parameterized models. The increase in model identifiability problems can be demonstrated by examining modelling efficiency frequency density plots (Figure 3).



(a) (b)

Figure 3. Efficiency frequency density plots for the estimation of RWB discharge (a) and of RWB BOD concentration (b)

Such diagrams show the frequency that a model simulation efficiency can be included in a specific interval and it gives an idea of the shape of the modelling efficiency surface (i.e. the surface representing modelling efficiency as a function of parameters values in the parameters confidence region). If frequency is uniformly distributed in all intervals, the shape of efficiency surface is not steep and several parameters sets can be find obtaining a good modelling performance; if frequency is progressively reducing from lower efficiency intervals to higher ones, the efficiency surface is steep and the identification of a single calibration parameters set is easier. Figure 3 shows that data availability affects parameters identifiability problems of the adopted model especially for water quality aspects that are probably characterised by higher uncertainty.

5. CONCLUSIONS

The study analysed the effect of data availability on urban drainage integrated modelling reliability. Such effect has been analysed by the mean of GLUE uncertainty analysis approach. The study allowed for the following considerations:

- Modelling uncertainty is affected by data availability even if some differences can be highlighted among different sub-models; when RWB analysis is the aim of the study, imperfect knowledge in upstream sub – systems (SS and CSO) may be partially corrected by the model thanks to the higher amount of information available in the downstream one.
- The importance of single parameters in uncertainty propagation is not greatly affected by data availability so increasing modeller confidence in the fact that a first modelling analysis when low amount of data is available can help in focusing further monitoring campaigns towards the most sensitive parts of the system.
- Data availability greatly can greatly modelling parameter identifiability thus reducing modeller confidence in the calibration process; this aspect also affects the possibility that initial calibration (based on low amount of data) may be confuted by following data acquisition.

The study is based on a specific case study and model, thus requiring extended applications for generalising the study results. Nevertheless, the study suggests a method for evaluating the importance of data in integrated modelling identifying where larger efforts are needed with the aim of collecting more information.

ACKNOWLEDGEMENTS

Authors gratefully acknowledge Prof. Sandro Artina and Dr. Marco Maglionico (D.I.S.T.A.R.T. – University of Bologna, IT) for providing data of Savena catchment.

REFERENCES

Artina S., Bardasi G., Borea F., Franco C., Maglionico M., Paoletti A., Sanfilippo U., Water quality modelling in ephemeral streams receiving urban overflows. The pilot study in Bologna, 8th International Conference Urban Storm Drainage, ICUSD, Sydney, 30 Agosto-3 Settembre 1999, Joliffe Ian B. & James E. Ball (eds.), Sydney, Australia, Vol. 3, 1589-1596, 1999.

Beven K.J., Binley A.M., The future of distributed models - model calibration and uncertainty prediction, Hydrological Processes, , Vol. 6(3), pp. 279-298, 1992

Beven K.J., Freer J. Equifinality, data assimilation and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of Hydrology, 249: 11-29, 2001

Blöch, H., The European water framework directive: taking European water policy into the next millennium. Water Science and Technology 40 (10), 67-71, 1999.

Freni, G., Mannina, G., Viviani, G., Uncertainty in urban stormwater quality modelling: The effect of acceptability threshold in the GLUE methodology, Water Res., Volume 42, 8-9, 2061-2072, 2008a

Freni, G., Maglionico, M., Mannina, G., Viviani, G., Comparison between a detailed and a simplified integrated model for the assessment of urban drainage environmental impact on an ephemeral river, Urban Water J. 5 (2), 21–31, 2008b.

Gupta, H. V., Beven, K. J., Wagener, T., Model Calibration and Uncertainty Estimation." Encyclopedia of Hydrological Sciences. Edited by M. Anderson. John Wiley & Sons, Ltd, 2005.

Lindenschmidt K. E., The effect of complexity on parameter sensitivity and model uncertainty in river water quality modelling, Ecological Modelling 190, 72–8672-86, 2006

Mannina, G., Integrated urban drainage modelling with uncertainty for stormwater pollution management, Ph.D. Thesis, Universita` di Catania, Italy, 2005.

Mannina, G., Viviani, G., Freni, G., Modelling the integrated urban drainage systems, In: Bertrand-Krajewski, J.-L., Almeida, M., Matos, J., Abdul-Talib, S. (Eds.), Sewer Networks and Processes within Urban Water Systems. IWA Publishing, London, 3–12, 2004.

Nash, J.E. and Sutcliffe, J.V., River flow forecasting through conceptual models. J. of Hydr., 10(3), 282–290, 1970.

Rauch W. and Harremoës P., The importance of the treatment plant performance during rain to acute water pollution. Wat. Sci. Tech., 34 (3-4), 1-8, 1996.

Schütze, M., Butler, D., and Beck, M. B., Optimisation of control strategies for the urban wastewater system. An integrated approach, Water Science and Technology ,39(9), 209-216, 1999.

Takács, I., Patry, G.G., Nolasco, D. (1991). A Dynamic Model of the Clarification-Thickening Process. Wat. Res., 25(10), 1263-1271.

Vanrolleghem P., Schilling W., Rauch W., Krebs P., Aalderink H., Setting up measuring campaigns for integrated wastewater modelling" - Wat. Sci. and Tech., vol. 39, n. 4, pp. 257-268, 1999.

Vitasovic, Z. C., Zhou, S., McCorquodale, J. A., and Lingren, K., Secondary clarifier analysis using data from the Clarifier Research Technical Committee protocol. Water Environment Research 69, 999 – 1007, 1997.

Willems, P.. Probabilistic immission modelling of receiving surface waters. Ph.D. Thesis, K.U. Leuven, 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

Comparison of control strategies for multi-objective control of urban wastewater systems

G. Fu, D. Butler and S.-T. Khu

Centre for Water Systems, School of Engineering, Computing and Mathematics, University of Exeter, North Park Road, Harrison Building, Exeter EX4 4QF, UK (g.fu/d.butler/s.t.khu@exeter.ac.uk)

Abstract: In recent years much attention has been paid to integrated management and control of urban wastewater systems. With the application of integrated system modelling tools, overall system performance can be improved to a great extent in terms of receiving water quality, through development of optimal control strategies. Most studies to date, however, have used a single objective to demonstrate the potential benefits. Control of urban wastewater systems is actually a multiple objective optimisation problem, involving balancing different, possibly conflicting objectives required by stakeholders with different interests. This paper compares three different control strategies for multi-objective optimal control of the urban wastewater system, including one global control strategy and two integrated control strategies. A popular multiple objective evolutionary algorithm, NSGA II, is applied to derive the Pareto optimal solutions for the three strategies. The comparative results show the benefits of application of integrated control in achieving an improved system performance in terms of dissolved oxygen and ammonium concentrations in the receiving river. The simulation results also illustrate the effectiveness of NSGA II in deriving the optimal control strategies with different complexities.

Keywords: Evolutionary algorithms, Integrated control, Multi-objective optimisation, NSGA II, Urban wastewater system.

1. INTRODUCTION

There is growing recognition of the need for and benefits of integrated simulation of the sewer system, wastewater treatment plant, and receiving water body in order to achieve a better receiving water environment [Rauch et al., 2002; Schütze et al., 2002; Butler & Schütze, 2005; Vanrolleghem et al., 2005]. In recent years, several simulation tools and methods have been developed, for example, SYNOPSIS [Schütze et al., 2002], SIMBA [IFAK, 2005], WEST [Vanhooren et al., 2003], and CITY DRAIN [Achleitner et al., 2007], and this provides the opportunity to optimise the urban wastewater system as a whole. The benefits of integrated simulation and control include 1) simultaneous optimisation of various components in the three subsystems; 2) evaluation of the performance of the urban wastewater system directly using receiving water quality indicators, rather than by reference to surrogate criteria such as CSO discharge frequency/volume or treatment plant effluent quality, and 3) control of one subsystem based on the information from other subsystems, termed as 'integrated control', which makes the best use of potential interactions between subsystems to further improve system performance [Schütze et al., 1999].

In the context of integrated simulation of urban wastewater systems, the improvement in system performance has been demonstrated through development of optimal control strategies. Although most of the effort has been focused on a single objective in the receiving water (e.g. DO and Ammonium concentrations), there are a few studies considering multi-objectives to develop optimal control strategies for urban wastewater systems [e.g., Schütze et al., 2002; Fu et al., 2008]. This paper will describe the development of three different control strategies for multi-objective optimal control,

including two integrated control strategies, and the parameter optimisation using a multi-objective evolutionary optimisation method.

2. MULTI-OBJECTIVE CONTROL

Control of urban wastewater systems is actually a multi-objective optimisation problem in practice, which has to balance different objectives in order to meet the requirements by stakeholders with different interests. Mathematically, the optimal control problem can be described as follows:

$$\text{Min } F(\mathbf{x}) = \{f_1(\mathbf{x}), \cdots, f_m(\mathbf{x})\} \quad (1)$$

where **x** is the variable vector which defines a specific control strategy in the feasible solution space, and $f_1, \cdots, f_m$ are the m objective functions to be simultaneously minimized. These objectives could arise from different parts of the system and possibly are conflicting with each other in nature. For example, avoiding and reducing sewer flooding and CSO discharges, ensuring treatment plant effluent quality complies with legislative requirements, maintaining or improving the water quality in receiving water bodies. In this paper, two water quality indicators for the receiving river are considered: minimum DO concentration (DO-M) and maximum ammonium concentration (AMM-M) for all the river reaches.

In general, different control strategies may achieve very different results for each of the objectives. Due to the non-linear, complex behaviours of urban wastewater systems, the objectives can only be evaluated through a simulation model. An existing model was used in this research to simulate various hydraulic and biochemical processes in the integrated system and thus to evaluate the chosen objectives [Fu et al., 2008]. This model was developed using the SIMBA tool, which is based on the SIMULINK® environment. This model allows a system dynamic modelling of the different processes in various subsystems and the interactions between them, which makes it possible to apply an integrated control with information interaction between different subsystems.

It is important to choose an appropriate optimisation method in order to reveal the whole trade-off relationship between objectives, which could help decision makers to make an informed decision. Evolutionary algorithms (EAs) have been regarded as promising to derive the optimal control strategies, in comparison with the conventional optimisation techniques [Rauch and Harremoës, 1999; Muschalla et al., 2006]. In this research, a state of the art multiobjective genetic algorithm, NSGA II [Deb et al., 2002], was chosen to derive the Pareto optimal control strategies.

3. THE INTEGRATED CASE STUDY

The approach will be demonstrated by a semi-hypothetical case study, consisting of a combined sewer system, a treatment plant and receiving river. This integrated case study was originally defined by Schütze [1998] and has been studied in detail for real time control optimisation [Schütze et al., 2002; Butler and Schütze, 2005; Fu et al., 2008].

Figure 1 shows the schematic representation of the integrated system. The sewer system has seven sub-catchments with a total area of 725.8 ha, and four on-line pass-through storage tanks linked to sub-catchments SC2, 4, 6 and 7 respectively, which are controlled by a pump. The storage tanks have a total volume of 13,200 m^3. The wastewater treatment plant includes an off-line pass-through storm tank with a volume of 6750 m^3, a primary clarifier, aerator, and secondary clarifier. The storm tank is controlled as follows: filling starts when the inflow to the primary clarifier reaches its maximum value and emptying is triggered when the inflow drops below a threshold value. The treatment plant effluent and storm tank overflow are discharged to the river at Reach 10. The river is divided into 45 reaches, each 1km in length. For the detailed set-up of the case study, the reader is referred to Schütze [1998] and Schütze et al. [2002].

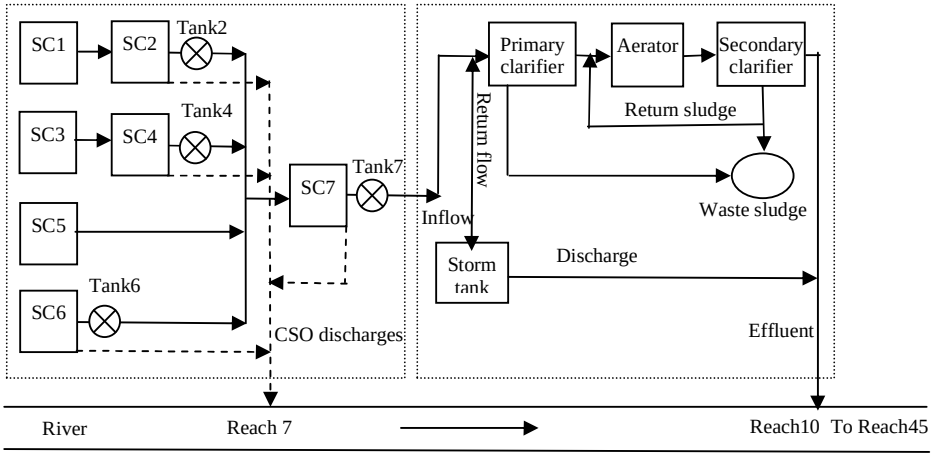


Figure 1. Schematic representation of the integrated urban wastewater system (adapted from Fu et al. [2008]. SCx represents the xth sub-catchment, and the dash lines show CSO discharges from the four storage tanks.

4. DEVELOPMENT OF CONTROL STRATEGIES

Development of control strategies for urban wastewater systems can be divided into two key stages: control strategy setup that determines what processes to be influenced by what measurements; and control strategy optimisation that aims to derive the optimal control parameters using an appropriate optimisation approach. In this research, three control strategies, originally developed by Schütze et al. [2002], were adapted here and compared for multi-objective control of the integrated system.

4.1 Global Control

A global control strategy was developed to control both of the sewer system and treatment plant, however, only local information was used in this control strategy setup, i.e., no information interaction between subsystems. According to the sensitivity analysis by Schütze et al. [2002], the most sensitive variables include the maximum outflow rate of Tank7 (x_1), maximum inflow rate to treatment plant (x_2), the threshold for emptying the storm tank (x_3) and its emptying flow rate (x_4). These variables were chosen in this global control strategy set up and are operated in the constant settings in which their optimal values are to be optimised. The feasible ranges used in this research is shown in Table 1, according to Schütze et al. [2002].

Table 1. Control variables and their ranges. The variables x_1 to x_4 are used for global control, x_1 to x_7 for integrated control (a), and x_1 to x_8 for integrated control (b).

Variables	Description	Ranges
x_1 (×DWF)	The maximum outflow rate of Tank7	[3, 8]
x_2 (×DWF)	The maximum inflow rate to treatment plant	[2, 5]
x_3 (m3/s)	Threshold triggering emptying the storm tank	[0.19, 0.36]
x_4 (m3/s)	The emptying flow rate of storm tank (m3/s)	[0.08, 0.28]
x_5 (%)	The threshold of filling degree for 'heavily loaded Tank 7'	[50, 100]
x_6 (×DWF)	The threshold of influent for 'heavily loaded treatment plant'	[3, 5]
x_7 (%)	The threshold of filling degree for 'heavily loaded storm tank'	[50, 100]
x_8 (×DWF)	The threshold of flow rate for 'heavily loaded river'	[4, 5]

4.2 Integrated Control (a)

This strategy is defined as integrated control because it uses measured information from one subsystem to control another subsystem [Schütze et al., 2002]. This can make the best of the available capacity to achieve a better overall system performance. This control strategy is an extension of the global control strategy defined above, through further exploration of the spare capacity of tanks in sewer system or treatment plant when one subsystem is heavily loaded.

Only the sewer system and treatment plant's states are considered in this control strategy set up. As different state variables can be used to indicate each subsystem's states, and the choice could have a major impact on system performance. In this research, the state of sewer system is measured by the filling degree of Tank7 (s_1), and the state of treatment plant measured by the flow rate of treatment plant influent (s_2). The storm tank is regarded as a separate component here in order to explore its potential in retaining wastewater for treatment, and its state is indicated by its filling degree (s_3). For simplicity's sake, only two states are considered for each of the components, i.e., 'heavily loaded' and 'not heavily loaded'. A threshold value is used to determine which state each component lies in, i.e., the component is regarded as 'heavily loaded' if the value of its state variable is bigger than its threshold. In this research, the optimal thresholds x_5 to x_7 are derived by NSGA II, and its feasible ranges are listed in Table 1.

Considering all the possible combinations of the states, there are 8 if-then rules in total of which the control strategy set up is comprised, as shown in Table 2. For example, Rule 1 for the case that all the three components are not heavily loaded, is interpreted as

IF $s_1 < x_5$ and $s_2 < x_6$ and $s_3 < x_7$ THEN $x'_1 = x_1$ and $x'_2 = x_2$

x'_1 and x'_2 will override x_1 and x_2, respectively. The other control variable x_3 and x_4 keep the same for the rules but will also be optimised.

4.3 Integrated Control (b)

On the basis of integrated control (a), this control strategy setup further considers the state of river to improve the river water quality. The river state is also regarded as 'heavily overloaded' if its flow rate is bigger than its defined threshold x_8, otherwise, not heavily overloaded. There are totally 16 rules defined for this control strategy setup. The 8 rules in which the river is 'not heavily overload', are defined as the same as integrated control (a), (b) in Table 2 shows the other 8 rules with the 'heavily overloaded' river.

Table 2. The rules for the integrated control strategies (a) and (b).

Rule number	Sewer system over-loaded?	Storm tank over-loaded?	Treatment plant over-loaded?	$[x'_1, x'_2]$ for (a)	$[x'_1, x'_2]$ for (b)*
1	No	No	No	$[x_1, x_2]$	$[x_1, x_2]$
2	No	No	Yes	$[x_1, x_2]$	$[x_1, x_2]$
3	No	Yes	No	$[x_1, x_2]$	$[x_1, x_2]$
4	No	Yes	Yes	$[x_1\text{-}1, x_2]$	$[x_1\text{-}1, x_2\text{+}1]$
5	Yes	No	No	$[x_1\text{+}1, x_2]$	$[x_1\text{+}1, x_2]$
6	Yes	No	Yes	$[x_1\text{-}1, x_2]$	$[x_1\text{-}1, 2]$
7	Yes	Yes	No	$[x_1\text{+}1, x_2\text{+}1]$	$[x_1\text{+}1, x_2\text{+}1]$
8	Yes	Yes	Yes	$[8, x_2\text{+}1]$	$[8, x_2\text{+}1]$

*These 8 rules are for the river state of 'heavily overloaded', and the other 8 rules for 'not heavily overloaded' are defined as the same as control (a).

5. RESULTS AND DISCUSSION

5.1 Runtime convergence

A runtime convergence indicator was used to measure the algorithm's performance [Deb and Jain, 2002]. This indicator measure the average Euclidean distance between a solution set *S* and a reference set $P = \{p_1, \quad p_2, \quad \cdots \quad p_n\}$ and is calculated as follows

$$C = \frac{1}{|S|}\sum_{i=1}^{|S|} d_i \tag{2}$$

where the smallest Euclidean distance d_i of each solution s_i in S to P is calculated as

$$d_i = \min_{j=1}^{n} \sqrt{\sum_{k=1}^{m}\left(\frac{s_i(k) - p_j(k)}{p_{\max}(k) - p_{\min}(k)}\right)^2} \tag{3}$$

$p_{\max}(k)$ and $p_{\min}(k)$ are the maximum and minimum objective values of *k*-th objective in the reference set. The reference set is made up of the non-dominated solutions from all the solutions visited in a simulation run. The obtained indicator is normalized to the range between 0 and 1, which represent a poor and perfect performance, respectively.

Figure 2 shows the runtime convergence for each of the control strategy setups. The parameters of NSGA II are set to the same values for each run, amongst them the population size and number of generations are set to 100. For the global control setup, the algorithm can converge rapidly within 30 generations and it converges slower for the other control strategy setups. For all the three setups, this algorithm reaches its ultimate convergence after 60 generations.

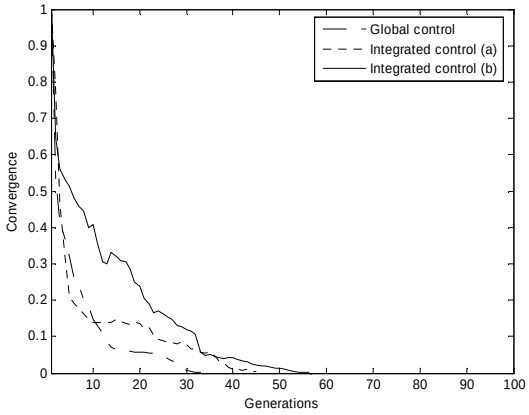


Figure 2. The run time convergence for the three control strategy set-ups.

Comparing the three setups, NSGA II attains the best for the global control in approaching the true Pareto front, and performs the poorest for the integrated control (b). The performance here is probably related to the complexity of control strategy setups, i.e., only 4 parameters were used for the global control in total, however, 7 and 8 parameters for the two integrated controls, respectively.

5.2 The Pareto solutions

Figure 3 shows the sets of Pareto optimal solutions from the three control strategy set-ups derived by NSGA II. For each set-up, a clear trade-off curve can be observed between DO-

M and AMM-M, in which a solution cannot be improved upon without deteriorating either of the objectives. Understanding the trade-off relationship between these two objectives will give decision makers an insight into the implications of different control strategies and thus help them make an informed decision.

Comparing the global control and two integrated control strategies, Figure 3 shows that the set of Pareto solutions from the global control is completely dominated by those of the integrated control strategies, which means that the integrated controls can achieve a better system performance in terms of both river DO and ammonium concentrations. This confirms the benefit of considering information interaction in system control. It can also be seen that the set from integrated control (b) dominates that of integrated control (a). This shows the importance of choosing the right sensors and controllers in the control strategy setup process. If the critical control variables or measures are not included in control strategies, it might be not able to achieve the best system performance. As shown in this case study, in addition to the states of components in sewer system and treatment plant, the real time river information plays a significant role in improving system performance.

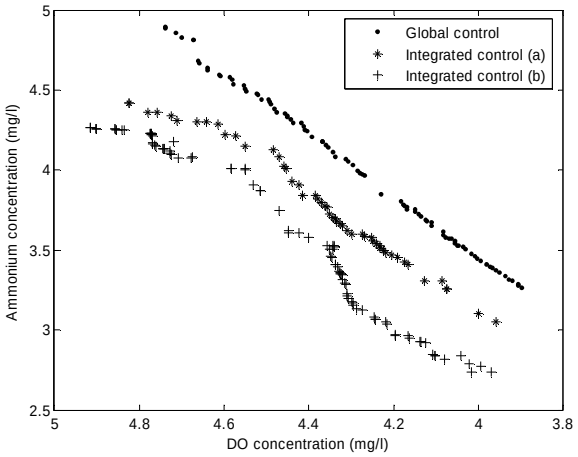


Figure 3. The Pareto optimal solutions from the three control strategy set-ups.

6. CONCLUSIONS

Recent advances in integrated modelling of the urban wastewater system make it possible to apply integrated control to achieve better overall system performance in terms of receiving water quality. While the receiving water quality can be measured by different determinants, there are also other objectives that might need to be considered to address interests from different stakeholders. So there is a real need to apply multi-objective optimisation methods to reveal any possible trade-offs between the objectives. This paper discussed three control strategy setups based on different levels of information interactions between the subsystems. A popular multiple objective evolutionary algorithm, NSGA II, was applied to derive the Pareto optimal solutions for the three strategies. Optimisation results from this research show the necessity for the use of multi-objective optimisation to derive Pareto optimal solutions. This provides decision makers a detailed understanding of the trade-off relationships when balancing the control objectives to meet various needs in practice. Further, the potential benefits of integrated control can be illustrated in comparison with global control in terms of both DO and ammonium concentrations. The simulation results also illustrate the effectiveness of NSGA II in deriving the optimal control strategies with different complexities.

ACKNOWLEDGEMENTS

The study is partly funded by UK EPSRC grant no GR/S86846/01 and is also part of the Integrative Systems and the Boundary Problem (ISBP) project (www.tigress.ac/isbp), supported by the European Union's Framework 6 Programme New and Emerging Science and Technology Pathfinder initiative.

REFERENCES

Achleitner, S., Möderl, M., Rauch, W., CITY DRAIN © – An open source approach for simulation of integrated urban drainage systems. *Environmental Modelling & Software*, 22 (8), 1184-1195, 2007.

Butler, D. & Schütze, M., Integrating simulation models with a view to optimal control of urban wastewater systems. *Environmental Modelling & Software*, 20, 415-426, 2005.

Deb, K**.,** Jain, S., Running performance metrics for evolutionary multi-objective optimization. Kangal Report No. 2002004, 2002.

Deb, K., Pratap, A, Agarwal, S., Meyarivan, T., A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transaction on Evolutionary Computation*, 6 (2) 181-197, 2002.

Fu, G., Butler, D., & Khu, S. T., Multiple objective optimal control of integrated urban wastewater systems. *Environmental Modelling & Software*, 23(2), 225-234, 2008.

IFAK, SIMBA 5.0: Simulation of Wastewater Systems, User's Guide. Institut für Automation und Kommunikation e.V. Magdeburg, Germany, 2005.

Muschalla, D., Schrφter, K., Schütze, M., Multi-objective evolutionary algorithms in the field of urban drainage. Seventh International Conference on Hydroinformatics, Nice, 2006.

Rauch, W., Harremoës, P., On the potential of genetic algorithms in urban drainage modelling. *Urban Water*, 1, 79-89, 1999.

Rauch, W., Bertrand-Krajewski, J.-L., Krebs, P., Mark, O., Schilling, W., Schütze, M. and Vanrolleghem, P.A., Mathematical modelling of integrated urban drainage systems. *Water Science and Technology*, 45(3), 81-94, 2002.

Schütze, M., Integrated Simulation and Optimum Control of the Urban Wastewater System. PhD thesis. Department of Civil Engineering; Imperial College of Science, Technology and Medicine, University of London, London, 1998.

Schütze, M., Butler, D., & Beck, M.B., Optimisation of control strategies for the urban wastewater system - an integrated approach. *Water Science and Technology*, 39 (9), 209-216, 1999.

Schütze M., D. Butler and M.B. Beck. Modelling Simulation and Control of Urban Wastewater Systems, Springer, Berlin, 2002.

Vanhooren H., Meirlaen J., Amerlink Y., Claeys F., Vangheluwe H. and Vanrolleghem P., WEST: modelling biological wastewater treatment. Journal of Hydroinformatics 5(1), 27-50, 2003.

Vanrolleghem, P.A., Benedetti, L., & Meirlaen, J., Modelling and real-time control of the integrated urban wastewater system. *Environmental Modelling & Software*, 20 (4), 427-442, 2005.

AqquaStudy – enabling remote access to model-based software tools for supporting the design and operation of WWTPs

I. Irizar[a] **and M. De Gracia**[b]

[a] *CEIT and TECNUN. Manuel Lardizabal 15, 20018 San Sebastián, Spain (iirizar@ceit.es)*

[b] *ATM S.A. Epele Bailara 29, 20120 Hernani, Spain (mdegracia@atmsa.com)*

Abstract: Making use of mathematical programming algorithms, a software library (AqLib) has been built for automatically optimising the design and operation of wastewater treatment plants. Currently, AqLib implements seven configurations for nitrogen removal (DN, DN2, RDN, SDN, DRDN, RSDN and DRSDN), five configurations for phosphorous removal (A^2O, Bardenpho, Johannesburg, UCT and MUCT), and the anaerobic digestion process. Moreover, an Internet-based service (AqquaStudy) has been deployed so that AqLib can be remotely accessed from any standard web-browser.

Keywords: Web-based, Wastewater, Mathematical modelling, Optimisation

1. INTRODUCTION

Advances in computer technology have motivated significant changes in the methods that are currently being applied to solve engineering problems in general. Feasibility, robustness and calculation speed are some computing features which have encouraged knowledge being progressively transferred to decision-support software tools. As a result of the explosion of computer technologies, model-based simulation has become an effective tool for decision-makers dealing with complex situations. In this respect, model-based software tools are, more than ever, an essential asset within many water engineering companies in order to support the design and operation of wastewater treatment plants (WWTP). As a result of more stringent regulations for both treated water discharges and sludge disposal, many WWTPs are being retrofitted with modern and complex treatment technologies whose design and operating criteria are not straightforward even for experienced professionals [Larrea et al., 2007]. Likewise, the growing competitiveness in the water industry is pushing companies to refine their current WWTP design and operating manuals with the adoption of new procedures that lead to optimum solutions.

Over the last decades, an intensive research has been focused on the mathematical formulation of all those mechanisms involved in the biological treatment of wastewaters. Nevertheless, owing to the great diversity of model proposals and the inherent complexity of these systems, modellers very soon realised the importance of promoting standardisation so as to establish a common framework around mathematical modelling issues. Thus, the confirmation of the standard IWA ASM models for biological COD and nutrient removal [Henze et al., 2000]) as a valuable tool for predicting the behaviour of activated sludge (AS) systems, have meant a decisive driving force towards the progressive acceptance of model-based techniques within water companies. Later on, drawn by the success of the ASM models, the IWA ADM1 model is published and adopted as standard for dynamic modelling of the anaerobic digestion processes [Batstone et. al., 2002]). Moreover, at present, great efforts are being addressed at providing accurate descriptions for promising technologies such as biofilm systems or membrane bioreactors. In these systems, complex

approaches based on partial derivative differential equations are proposed to model typical phenomena of diffusion or attachment/detachment [Eberl et al., 2006]).

Additionally, the idea of integration stressed by the European Water Framework Directive in relation to the management of water resources has also encouraged further advances in the mathematical modelling of wastewater treatment plants. Nowadays, the criteria for design, operation and control of WWTPs are still unit-process oriented; however, it has been demonstrated that optimum solutions can only be achieved when the interactions between the water and waste lines are considered. In this context, the term "Plant-Wide Modelling" is adopted to refer to all that work that, in the last years, has dealt with the integration of the existing unit-process models [Grau et al., 2007a]. Obviously, the WWTP design, operation and control problems become much more difficult and more time-consuming when they must be solved on the basis of plant-wide model approaches. Moreover, computational loads increase drastically with the application of modern methodologies where uncertainty analyses are included and, therefore, Monte Carlo simulations required [Benedetti et al., 2008].

Probably, the transfer of model-based methodologies from the academic domain to the water industry starts with the development and commercialisation of WWTP-specific simulation software packages such as WEST, SIMBA, GPS-X or Biowin. Though simulation has been traditionally addressed at supporting the design of WWTPs, interesting applications for operation and control have been also reported. Thus, based on combining historic data with dynamic simulations, Ayesa et al. [2001] implement a specific simulator for the Badiolegi WWTP that proves its effectiveness to facilitate both the diagnose of the plant and the exploration of operational strategies. Another example is the IWA/COST simulation benchmark, a simulation protocol which has been widely used to assess and compare control strategies [Coop, 2002]. Even so, current challenges are still focused on the establishment of systematic and straightforward procedures that assist in the proper use of modelling especially among practitioners and consultants. Thus, the goal for the coming years is to spread the use of simulation techniques in multiple areas: from industry to educational sectors. In this respect, two important barriers will have to be overcome: (1) the high investment built-in costs of simulation; and (2) model-based problem-solving is not a trivial issue, especially when optimum solutions are pursued.

Fortunately, the rapid consolidation of the World Wide Web as a common framework for applications has recently led to new web-based business models for accessing simulation services (web-based simulation). The distributed and client-server nature of the WWW enables: (1) the sharing of software resources between different customers so as to significantly reduce simulation costs; (2) hardware resources to be dedicated to specific simulation sub-tasks like processing, visualisation, storage, etc. Page and Opper [2000] classify web-based simulation practice into five primary areas: (1) simulation as hypermedia; (2) simulation research methodology; (3) web-based access to simulation programs; (4) Distributed modelling and simulation; and (5) Simulation of the WWW. Although several works dealing with each area have been reported in recent years, most of them cover area (3), i.e., web-based access to simulation programs. Wiedemann [2001] describes a web-based Application Service Provider for simulation of customised models and systems. By combining Java applet interfaces with distributed simulation, Yang and Alty [2002] demonstrate the potential of web-based online simulation for creating appropriate virtual experimental frameworks intended for distance learning. Cheng and Fen [2006] present a web-based distributed problem-solving platform for engineering applications. Additionally, the literature also reports some works on web-based simulation for wastewater treatment systems. For example, Samuelsson et al. [2001] employ Java applet technology to develop a Java-based dynamic simulator (JASS) for AS systems which can be accessed from any standard web browser.

There is a consensus around the necessity of model-based tools to cope with complex problems. However, these tools by themselves are not enough, being thus crucial to complement them with effective problem-solving methods that increase the confidence of

end-users. Mathematical programming techniques represent valuable instruments in this concern, since their integration within model-based engineering problems allows optimum solutions to be found in an automatic way. These techniques have already been applied to specific AS configurations with satisfactory results [Rivas et al., 2008]. This success encourages their expansion to other wastewater treatments technologies and other unit-processes. By combining web-features and mathematical programming, this paper presents AqquaStudy, an Internet-based service that offers remote and free access to a library of software tools aimed at facilitating the optimum design and operation of WWTP unit-processes.

2. AQLIB: SOFTWARE LIBRARY OF MODEL-BASED OPTIMISATION TOOLS

So far, the utilisation of model-based software tools in the water industry has mainly concentrated on the dimensioning of conventional AS systems. Often, these tools are simply software implementations of systematic procedures such as the German ATV guidelines or similar [ATV, 2002]. Usually, model-based WWTP design combines steady-state and dynamic-state simulations: a preliminary steady-state study is carried out to estimate a first solution which, then, must be verified under non-steady conditions in order to refine the final solution. Within this two-step procedure, mathematical programming can help to automatically obtain optimum solutions for the steady-state problem. It is worth noting that the steady state of continuous flow systems corresponds to zero values for all time-derivatives; therefore, the steady-state condition becomes a non-linear algebraic equation system (instead of a differential one) that can be easily formulated and solved using mathematical optimisation tools. In contrast, concerning non-continuous technologies (e.g., SBR, Biodenitro, Biodenipho, intermittent systems …), their steady-state regime involves a periodic behaviour where time-derivatives get non-zero values. Consequently, a more different formulation is required for these technologies. Nevertheless, the application of optimisation techniques to these systems is beyond the scope of the present work.

AqLib, the software library of engineering tools described below, allows users to automatically solve studies under steady-state conditions. Rather than replacing current WWTP dynamic simulation software packages, the objective of AqLib is to complement them by providing initial estimations so as to facilitate subsequent studies under dynamic conditions. At present, only continuous flow technologies have been considered within AqLib, in particular: (1) seven AS configurations for nitrogen removal (DN, DN2, RDN, SDN, DRDN, RSDN and DRSDN); (2) five AS configurations for phosphorous removal (A^2O, Bardenpho, Johannesburg, UCT and MUCT); and (3) the anaerobic digestion process. Nevertheless, this library will be hopefully extended in the future to cover more configurations and more treatment technologies.

2. 1. Mathematical modelling and optimisation

Concerning the AS configurations, the standard ASM2d model was selected to describe the bio-chemical transformations in reactors. Moreover, a simple instantaneous settling model was set for secondary clarifiers: this model approach is good enough for steady-state studies focused on optimising the design/operation of biological tanks. In relation to the anaerobic digester software tool, it was based on the standard ADM1 model.

In general, the definition of an optimisation problem encompasses the three following items: (i) a cost function, whose value must approximate to a given objective; (ii) the independent variables that the mathematical programming algorithm can modify to reach the objective; and (iii) the constraints that the problem must satisfy. Thus, the mathematical algorithm searches the optimum values for the design/operation variables (volumes of anaerobic, anoxic, and aerated zones, internal recycling flow, waste flow, etc.) that comply with the problem constraints (maximum concentration of mixed liquor total suspended

solids, minimum practical volumes, effluent quality, etc…). Values in Table 1 give, for the optimum design/operation problems of every configuration included in AqLib, an order of magnitude about their respective problem-solving complexity. The mathematical background around model-based optimisation algorithms for design and operation of WWTPs is detailed in Rivas et al. [2008]. Similarly, Irizar et al. [2005] and De Gracia et al. [2007] set out the equations involved in the development of a model-based optimisation software tool specific for the DN process and the anaerobic digester, respectively.

Table 1. Mathematical optimisation: built-in complexity of the optimum design/operation problems for different WW treatment technologies

	Nº of Independent Variables	Nº of Design/Operation Variables	Nº of Constraints
AS configurations for Nitrogen Removal			
DN	46	5	7
DN2	70	7	9
RDN	70	7	10
SDN	69	6	9
RSDN	92	8	10
DRDN	93	9	11
DRSDN	115	10	12
AS configurations for Phosphorous Removal			
A2O	69	6	9
Bardenpho	114	9	12
Johannesburg	92	8	10
UCT	70	7	9
MUCT	92	8	10
Sludge Treatment			
Anaerobic Digester	48	2	21

2. 2. Software implementation

Many mathematical problems can be easily formulated and solved through commercial spreadsheet packages. *Microsoft Excel* is probably the most widely-used spreadsheet by engineers. Although Excel is particularly appropriate when solving steady-state models, it also includes the VBA macro programming language so that dynamic models can also be solved. Moreover, *Solver* (by *Frontlyne Systems, Inc.*) is a commercial optimisation tool which is fully compatible with Excel in such a way that optimisation problems can be formulated within Excel worksheets before being solved through this tool. Therefore, since Excel meets the requirements for solving optimisation problems, AqLib has been implemented as a set of 13 Excel workbooks (one per treatment technology). All the workbooks have been designed in accordance to a unified pattern aimed at facilitating both their implementation and the development of additional model-based tools in the future. Thus, the 13 workbooks organise user information into four separated worksheets: (1) Influent characterisation; (2) Model coefficients; (3) Problem selection; and (4) Optimisation results. A screenshot of the "Influent Characterisation" worksheet for the DN tool is shown in Figure 1.

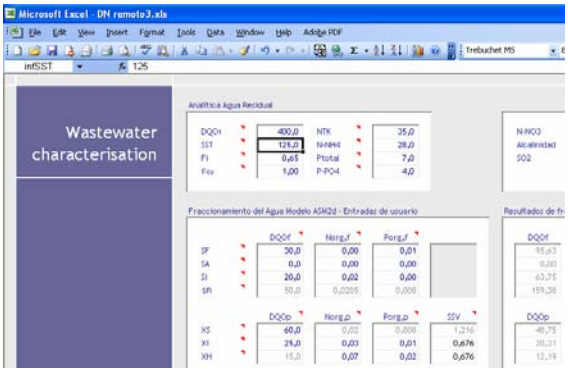


Figure 1. User input interface of the "Influent Characterisation" worksheet

In addition, every workbook includes a hidden worksheet where the design/operation problems are formulated in accordance to the specifications of the Solver optimisation tool. User-inputs in this worksheet are refreshed automatically every time a particular problem is launched to be solved. For that, within the workbook, specific VBA macros have been programmed that process user input information and send it to this worksheet before calling the Solver tool. Once the optimisation engine has solved the problem, these macros gather the solution and copy it to the "Optimisation Results" worksheet. In Figure 2, an overview of the software architecture of the implemented workbooks is illustrated.

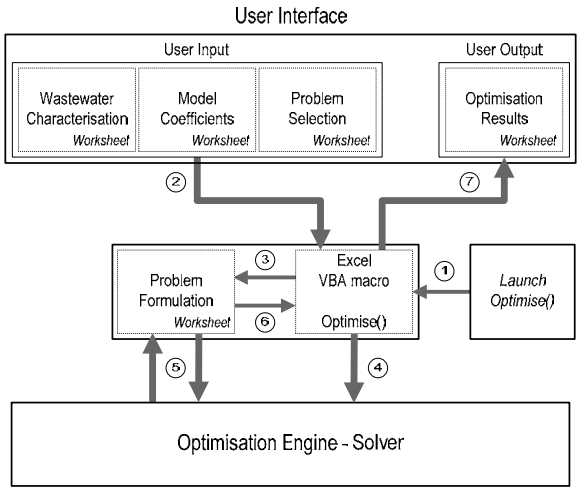


Figure 2. Software architecture of the Excel-based optimisation tools included in AqLib

3. AQQUASTUDY: INTERNET-BASED ACCESS TO THE SOFTWARE LIBRARY

Microsoft Excel integrates an OLE/COM interface that allows communication with external Windows-based applications supporting COM technology. This feature is especially attractive when the objective is to offer a remote service for accessing Excel-based engineering tools such as those contained in AqLib. The standard Web Services technology (http://www.w3.org/2002/ws/) has been used to implement a specific web service that provides client applications with a network interface for access to AqLib. The AqLib Web Service has been programmed within the Microsoft .NET environment because it supports COM technology. Through this web service, AqLib can be exploited by remote users from different client applications and hardware platforms. To achieve this, a XML schema has been defined so that clients and the web service can send each other I/O information bound to the problem-solving engineering tool. The web service architecture consists of three main layers: the interface layer; the logic layer; and the data access layer (Figure 3). Requests from clients are managed by the interface layer which dispatches them to the appropriate layer (either logic or data access). At the moment, four public functions form the interface layer:

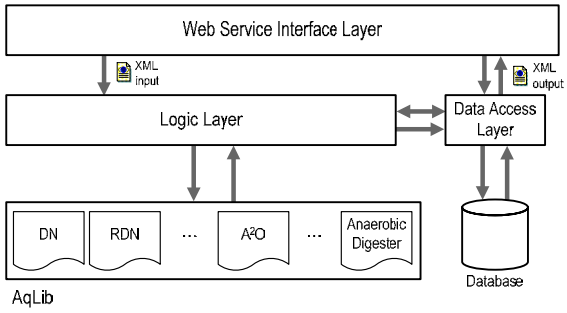


Figure 3. Architecture of the Web Service that enables remote access to AqLib

LaunchNewTask. Through this function clients invoke the web service in order for AqLib to solve a particular optimisation problem (task). *LaunchNewTask* requires two parameters from clients: (1) XML information defining the task (*XML task definition*); and (2) a *task Password* to protect access to task information. When this function is called, the interface layer sends both parameters to the logic layer and returns a *task Identifier* to the client.

IsTaskFinished. Clients make use of this function to know the current status (finished or in queue) of tasks launched previously. The function call has to include values for

two parameters: *task Identifier* and *task Password*. An error code is returned if either parameter does not match correctly.

RemoveTask. Clients are able to remove tasks in the queue at any time by calling this function with the proper values for *task Identifier* and *task Password* arguments. Then, the interface layer sends this request to the data access layer for the removal of the task from the queue.

GetTaskSolution. Every time a task is solved by AqLib, the logic layer processes optimisation results (task results) and brings them in XML format to the data access layer. Finally, the data access layer stores the XML task results in the database. *GetTaskSolution* provides clients with a mechanism to gather XML task results. An e-mail address has been considered in the XML definition task structure so that the web service can alert users about the completion of their respective launched tasks.

3. 1. Web-based Client: AqquaStudy

As mentioned above, multiple client applications might be implemented in different hardware platforms for providing users with remote access to AqLib. Nevertheless, a default Web-based client application has been deployed that enables the AqLib tool to be used from any local computer connected to the Internet (Figure 4). This Web-based service has been named AqquaStudy (AqST) and, at present, it can be freely accessed from the web site http://www.aqquas.com (under construction). AqST represents a low-cost solution for small and medium water companies to incorporate modern WWTP design procedures. AqST enables new business models to be agreed such as pay per use instead of acquiring the software. Likewise, this service is aligned with current needs from workers who, because of a growing mobility, increasingly are demanding an access to software services from anywhere.

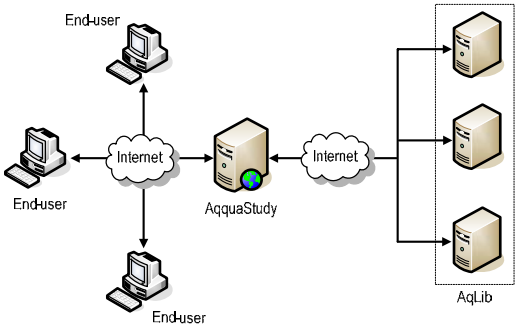


Figure 4. Hardware architecture of AqquaStudy

For now, AqST enables users to carry out model-based optimum design studies for three AS configurations: DN, DN2 and RDN. Figure 5 shows a screenshot of the user-input interface for the DN process. Initially, it has been designed a basic input interface where only the most relevant input information required to define an optimisation problem has been included. Thus, concerning the wastewater characteristics, only laboratory measurements have been considered (Total COD, Filtered COD, TSS, VSS, TKN and NH_4-N). On the other hand, default ratios have been used to obtain the different influent COD fractions, although this could be also optional for users depending on their knowledge. In addition, the following information is requested to define the optimisation problem: (1) temperature; (2) Mixed Liquor Suspended Solids in biological tanks; and (3) effluent Nitrogen requirements.



Figure 5. Screenshot of the Web-based user input interface for AqquaStudy: DN configuration

Similarly, a basic Web user output interface has been designed to give users the most significant optimisation results: (1) HRT and SRT; (2) Total Volume; (3) Anoxic and

aerobic Volume fractions; (4) Internal recycling flow-rate; (5) Wasted sludge flow-rate; (6) Sludge production; (7) Oxygen requirements; and (8) Dissolved oxygen transfer coefficient in aerated tanks.

3. 2. System performance

An optimum design problem has been defined for the DN process in order to evaluate the performance of AqST (Table 2). This problem has been solved using two different tools: (1) AqST (AS); and (2) a commercial WWTP dynamic simulation software package (DS).

Table 2. Input data for the optimum design of the DN process

Wastewater Characteristics		
TCOD	mg O_2/L	350
FCOD	mg O_2/L	150
TSS	mg/L	125
VSS	mg/L	100
TKN	mg N/L	35
NH4-N	mg N/L	28
Design requirements		
Temperature	º C	15
MLSS	mg /L	3000
NH4-N	mg N/L	1
NO3-N	mg N/L	10

Table 3. Optimisation Results(*)

		AqS	**DS**
HRT	h	13.9	13.8
SRT	d	14.2	14.1
Aerobic fraction	%	82	82
Internal Recycle	%Qinf	122	114
Waste Sludge	%Qinf	3.6	3.6
Effluent NH4-N	mg N/L	1	1
Effluent NO3-N	mg N/L	10	10
Solving Time	min	0.5	15-30

(*) Further details on the optimisation problem and the algorithm used can be found in Irizar et al. [2005]

Table 3 shows the results obtained with each tool. It can be seen that, in both cases, the optimum design was very similar. However, major differences appear when comparing the time consumed for reaching these results. While AqST took less than 1 minute to solve the problem, with the dynamic simulation tool, this value increased up to more than 15 minutes. In fact, these differences would have been more significant if the comparison, instead of using the DN process, had been done with a more complex configuration.

3. 3. Further extensions in AqquaStudy

The AqLib library should be extended in the next years with more engineering tools. Thus, further research must be addressed at incorporating additional problem-solving tools. Some examples are: (1) the design and operation of sequencing systems such as SBR, ATAD, etc.; (2) tools for wastewater characterisation [Grau et al., 2007b]; (3) the design of settling tanks; (4) the design of aeration systems, etc.

The hardware architecture of AqST can be easily expandable in the future depending on the end-user requirements. AqST supports the deployment of computational grids in the back-end (dedicated machines where the software applications of AqLib are installed) so as to increase both computational power and storage capacity and, as a consequence, to minimise the time for solving user-defined problems. At the moment, the XML task definition schema allows a single problem, only, to be specified. However, a new design for the XML task definition schema (and probably for the XML task results too) will have to be analysed for dealing with studies that demand batch simulations.

4. CONCLUSIONS

A software library has been designed for solving complex model-based engineering problems related to optimum design and operation of WWTP unit-processes. In addition, a Web-based approach (AqquaStudy) has been deployed to provide users with remote access to the software library. The implemented web service offers full utilisation of this tool from many different software clients as well as hardware platforms. AqquaStudy represents a

valuable service to complement existing WWTP-specific dynamic simulation packages by providing them with the optimum values for the design/operation variables. Consequently, interfaces that automatically export the results from AqquaStudy to these commercial tools should be developed in order to facilitate users their utilisation.

REFERENCES

ATV-DVWK-A 131E, German ATV-DVW-Standards. DVWK-A 131E, Dimensioning of Single-Stage Activated Sludge Plants, ISBN: 3-935669-96-8, 2000

Ayesa, E., G. Garralon, A. Rivas, J. Suescun, L. Larrea and F. Plaza, New simulators for the optimum management and operation of complex WWTPs, *Water Science and Technology*, 44(2-3), 1-8, 2001

Benedetti, L., D. Bixio, Claeys, F., and P.A. Vanrolleghem, Tools to support a model-based methodology for emission/inmission and benefit/cost/rosk analysis of wastewater systems that considers uncertainty, Environmental Modelling and Software (2008), doi:10.1016/j.envsoft.2008.01.001

Batstone, D.J., J. Keller, I. Angelidaki, S. Kalyuzhnyi, S.G. Pavlostathis, A. Rozzi, W. Sanders, H. Siegrist and V. Vavilin, Anaerobic Digestion Model No. 1 (ADM1), *Scientific and Technical Report No. 13*, IWA Publishing, London, 2002

Cheng H-C.n and C.S. Fen, A web-based distributed problem-solving environment for engineering applications, *Advances in Engineering Software, 37*(2), 11-128, 2006

Copp, J., The COST Simulation Benchmark. Description and Simulator Manual, *Office for Official Publications of the European Community*, Luxembourg.

De Gracia, M., E. Huete, J.L. Garcia-Heras, and E. Ayesa, Algebraic solution of the mass balanced ADM1 to predict the steady state and to optimise the design of the anaerobic digestion of sewage sludge, *Water Science & Technology, 56*(9), 127–136, 2007

Eberl, H., E. Morgenroth, D. Noguera, C. Picioreanu, B. Rittman, M. van Loosdrecht, and O. Wanner, Mathematical modelling of Biofilms. *Scientific and Technical Report No. 18*, IWA Publishing, London, 2006

Grau, P., M. de Gracia, P.A. Vanrolleghem, and E. Ayesa, A new plant-wide modelling methodology for WWTPs, *Water Research*, *41*(19), 4357-4372, 2007a

Grau, P., S. Beltrán, M. De Gracia and E. Ayesa, New mathematical procedure for the automatic estimation of influent characteristics in WWTPs, *Water Science and Technology, 56*(8), 95-106 2007b

Henze, M., W. Gujer, M. Takashi and M. van Loosdrecht, Activated Sludge Models ASM1, ASM2, ASM2d and ASM3, *Scientific and Technical Report No. 9*, IWA Publishing, London, 2000

Irizar, I., A. Castro, C. Perez, and E. Ayesa, Solving model-based WWTP design and operation related engineering problems. Stand-alone and Web-based approaches, *Proceedings of the 2005 European Simulation and Modelling Conference EMS2005*, Porto, Portugal, October 24-26, 2005

Larrea, L., J. Albizuri, I. Irizar, and J.M. Hernandez, Design and operation of SBR processes for small plants based on simulations, *Water Science & Technology, 55*(7), 163–171, 2007

Page, E.H., and J.M. Opper, Investigating the application of web-based simulation principles within the architecture for a next-generation computer generated forces model, *Future Generation Computer Systems, 17*, 159-169, 2000

Rivas, A., I. Irizar, and E. Ayesa, Model-based optimisation of Wastewater Treatment Plants design, *Environmental Modelling & Software, 23*(4), 435-450, 2008

Samuelsson, P., M. Ekman, and B. Carlsson, A java based simulator of activated sludge processes, *Journal of Mathematics and Computers in Simulation, 56*(4-5), 333-346, 2001

Wiedemann, T., Simulation Application Service Providing (SIM-ASP), *In Proceedings of the 2001 Winter Simulation Conference*, 623-628, 2001

Yang, S.H., and J.L. Alty, Development of a distributed simulator for control experiments through the Internet, *Future Generation Computer Systems, 18*, 595-611, 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

Towards an Agent-Based design for the management of Urban Wastewater Systems

M. Aulinas[a, b], M. Poch[a], U. Cortés[b]
[a]*Laboratory of Chemical and Environmental Engineering, University of Girona. Campus Montilivi s/n, E-17071, Girona, Catalonia ({aulinas,manel}@lequia.udg.cat)*
[b]*Knowledge Engineering & Machine Learning Group, Technical University of Catalonia, c/Jordi Girona 1 & 3, E- 08034, Barcelona, Spain (ia@lsi.upc.edu)*

Abstract: Urban Wastewater Systems (UWS) have an important role in nowadays growing and urbanized areas. The management of the water and wastewater cycle in these areas is complex since wastewater sources are both heterogeneous and rapidly changing. Moreover, it implies to deal with different spatial and temporal references and to adapt to the changing conditions that might appear in both natural and urban environments. New paradigms are required to be undertaken when designing and modelling this complex dynamic system, since the need to communicate the embedded elements is essential for an efficient and effective management and a better understanding of the UWS. Multi-Agent Systems (MAS) are systems composed of several *software agents* that collectively are capable of reaching goals that are difficult to achieve by individuals. The concept of MAS shapes an important step towards the possibility to design the system to reach decisions in a more collaboratively and integrated way. The paper proposes a primary but consistent design of UWS using an agent-oriented methodology (GAIA). The agents in the UWS are designed in terms of the roles they can perform, so the activities and protocols they act upon, which finally depict a detailed description of the relations in between them. A general hierarchy between the agents is also given (*e.g.* the acquaintance model). Therefore, the agent-oriented design enables to depict both micro and macro level design of the UWS management.

Keywords: Urban Wastewater System (UWS); Integrated Management; Multi-Agent Systems (MAS); Agent Design

1 INTRODUCTION

The efficient and effective management of water and wastewater resources is a complex task. In the Mediterranean Catalan region most of the urban catchments are designed as unitary systems that collect both domestic and industrial wastewaters, each one presenting different and variable flows and polluting loads. In this context, the industrial component in the Urban Wastewater System (UWS) is of special concern mainly for urban Wastewater Treatment Plant (WWTP) managers. One of the most important reasons is the variability on its composition, changing widely depending on the function and activity of the particular industry. Whereas they generate relatively constant flow rates during production, these flow rates change markedly during cleanup and shutdown. This implies that urban Wastewater Treatment Plants (WWTP) often have to assume flows and loads that are not always predictable: industrial wastewater discharges are difficult to forecast and they cause both, long-term (seasonal) and short-term (hourly, daily and weekly) variations. Consequently they may enter in conflict with the real possibilities of wastewater treatment plants' capacity to treat with a minimum quality before pouring to a receiving media (*e.g.* river).

The UWS have several features that make them difficult to represent and model: they evolve over time; they have a multi-spatial coverage (*e.g.* a discharge in the upper part of the river may affect lower parts of the river); they involve interactions between physical-chemical and biological processes; on top, many of the processes involved are stochastic [Poch *et al.,* 2004; Rizzoli *et al.*, 1997]. All of these features require a multi-disciplinary expertise as well as ideal tools to successfully cope with these characteristics. Until now there has been a long modelling tradition of sewers, treatment plants and receiving waters

describing performance in terms of individual needs and objectives and more recently integrating the different parts to allow the wastewater system to be considered as a single system [Fu *et al*., 2008; Vanrolleghem *et al.,* 2005; Butler *et al.,* 2005; Pleau *et al.,* 2005 and Schütze *et al*., 2004]. These elements have typically been built using deterministic descriptions of the fundamental mechanisms and processes (*e.g.* hydraulics in sewer systems, wastewater treatment and quality change modelling in rivers, etc.). But the complexity of environmental domains makes it difficult to wholly describe the system in terms of numerical models that have an important limitation in describing the interactions among the elements.

Some of the conceptual challenges that the present mathematical models can not tackle when performing an integrated urban wastewater management can be overcome using the *agent paradigm*. The need to cope with the dynamics and emergent situations (*e.g.* the incorporation of new industries to the system) requires application components to interact in more flexible ways. The characterization in terms of agents has proven to be a natural abstraction to many real world problems, having convinced researchers and developers in a wide variety of domains [Jennings *et al.*, 1998; Nwana, 1996; Athanasiadis, 2005] of the great potential of multi-agent solutions. The Multi-Agent System can be built describing numerous *agents* with a different degree of complexity, according the information and knowledge available. The complexity and degree of knowledge can be improved without the necessity to modify all the system as they are intended to be systems with high modularity and scalability [Luck *et al.*, 2005].

We consider UWS as a systematic process where effective decisions, policies and strategies are designed to lessen the impacts of humans and industrial activities on a river basin. All activities, including structural and non-structural measures to prevent (avoid) or to limit (mitigate) adverse effects are considered in this definition. As citizens, governments and industrialists seek for more integrated information systems the specification of UWS gets more complex. Among other: Accessibility to information, Interoperability and Coordination stand as issues to be solved.

In this paper we propose an agent-oriented approach to design the system whose final aims are *1)* in normal conditions to make the UWS work to achieve the *best* possible quality of the river and to avoid leading the system to disturbed/emergency situations, *2)* in emergency or disturbed conditions, get the minimum possible impact (stressing on disturbances of industrial activity in the system) and *3)* to improve the context of decision-making process in the prevention of risks in a basin by focusing on the efficiency of information management and assessment services, thus improving the interoperability among the involved actors and their information systems. The agent-based model for the wastewater system is further described, making special emphasis on the description of *1)* the roles and their instantiations for each agent type presented in section 3, *2)* the obligations in terms of responsibilities, permissions and constraints of each role shown in section 4, and *3)* the communication and interaction patterns among the agents (section 5). Hence, the results of applying agent-based oriented methodology for the design of a UWS are presented, highlighting the conceptual models obtained. In section 6 a summary of the main advantages offered by this new conceptualization as well as some of the most important future challenges are provided.

2 METHODOLOGY: AGENT-ORIENTED DESIGN

Agent-oriented approaches can improve the ability to model, design and build complex systems. The development of MAS requires at least three important features to be resolved [Zambonelli *et al.* 2003, Tran *et al.* 2007, Biswas 2008]: *1)* the agent internal design, that is to identify the agent types and their roles; *2)* the agent interaction design, that is to design the coherent blocks of activity in which agents will engage to realize their roles and their properties and the possible exchanged messages; and finally *3)* the MAS organization model, describing the acquaintances and hierarchical organization structure, that is the communication pathways between the different agent types.

For the proposed agent-oriented approach the extended version of GAIA Methodology [Zambonelli *et al.*, 2003] to model the agents is used. GAIA is a general methodology that supports both the micro-level (agent structure) and macro-level (agent society and organization structure) conceptualization for agent-based system development. According to this methodology (for a comparison with other existing AOSE methodologies see Tran *et al.* [2008] and Biswas [2008]), to solve a given problem requires designing a system as three components:

- *Environment*: involves analysing all the entities and resources that the agents in Multi-Agent Systems (MAS) can use when working for a common goal. Actions are performed by the agent on the environment, which in turn provides percepts to the agent [Russell and Norvig, 2003]. Consequently, for our application-specific purposes it will be very important to identify, model and shape the characteristics of the involved environment. However the environment is not always physical (including, among many other percepts, the chemical and physical measures of river quality); sometimes it will be virtual (*e.g.* legislative framework and rules).
- *Roles*: since agents interact in some organisational setting that influences the agent behaviour, roles define what an agent is expected to do in the organization itself and with respect to other agents; thus, the roles are the specific tasks that an agent has to accomplish in the context of the overall system.
- *Interactions*: define interdependencies among agents that are illustrated using protocols. The interaction model describes the characteristics and dynamics of each protocol (*e.g.* when, how and by whom a protocol has to be executed).

In the next section the agent-based model for the wastewater system is further described making special emphasis on the explanation of *1)* the roles and their instantiations for each agent type, *2)* the responsibilities, permissions and constraints of each role, and *3)* the communication and interaction patterns between the agents.

3 THE AGENT MODEL FOR THE WASTEWATER SCENARIO

In this section the results of the analysis process that enable to identify the agent types, the roles of the system and the assignment of these roles to the agents are shown. We propose the modelling of a simplified UWS, whose main elements and processes are depicted in Figure 1. Accordingly, the understanding of the proposed MAS is an organization of entities (squares in Figure 1), represented as agents interacting

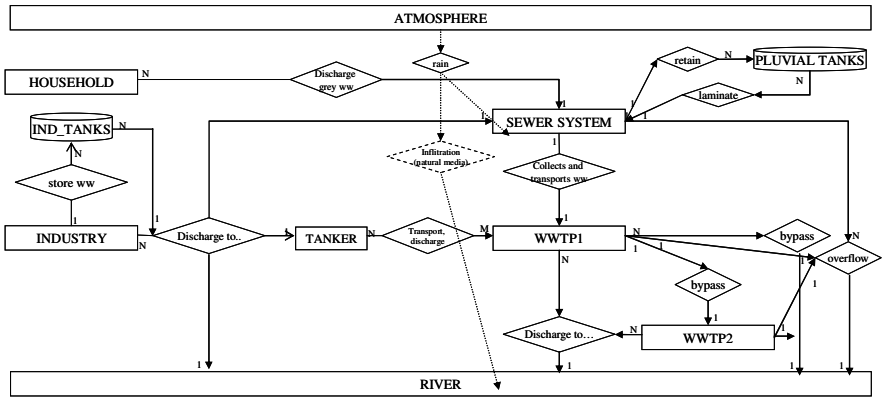


Figure 1. Urban Wastewater System: wastewater flow diagram with components and processes included in the study; each of the squares represents an agent and each of the rhombuses specific actions and/or processes occurring in the system.

A partial Agent Model is shown in Figure 2, where roles have been assigned to the agent types. In Table 1 and 2 the WasteWaterProducer (WWP), WasteWaterTreatment (WWT)

and RiverProtection (RP) roles are, respectively, fully described, in order to show the detailed design (note that *activities* are tasks an agent performs without interacting with other agents, thus it will be not necessary to describe a protocol for them).

From the list of protocols, activities, responsibilities and liveness properties of roles described in Table 1 and 2, it is possible to more easily describe the functions of the agents. In the following section a description of those services introduced in the previous tables is given.

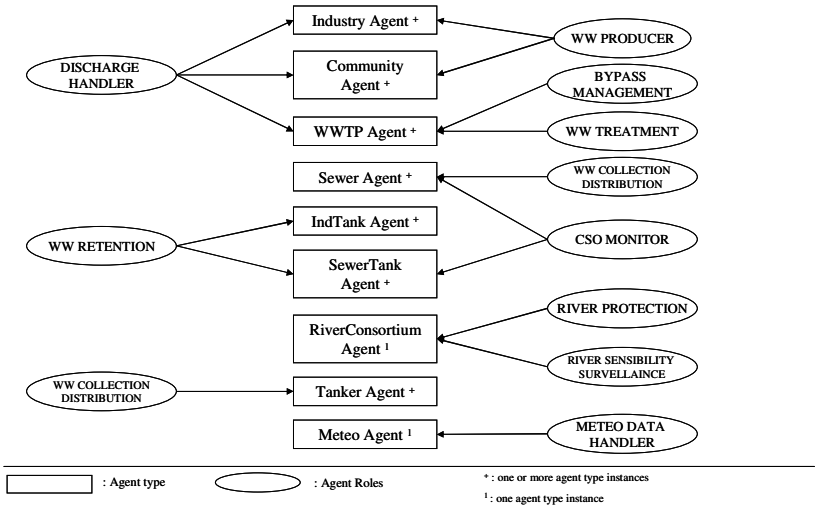


Figure 2. Agent Model for the wastewater management scenario

Table 1. Schema for the WasteWaterProducer and WasteWaterTreatment role

Role:	WasteWaterProducer (WWP)	WasteWaterTreatment (WWT)
Description:	Keeps track of individuals - industries, communities - quantity and quality of wastewater produced as consequence of their activity. Informs the receiver of the discharge about it. Calculate the cost of discharge.	Keeps track of wastewater flow that arrives at WWTP (according to treatment capacity – design parameters and yields – and hydraulic capacity) and controls the influent; supervise and control the treatment process, informing about the WWTP state if asked and giving alarms when a problem occurs; manages WWTP control set points.
Protocols & Activities:	CheckFlow, CheckPollutants (BOD, COD, TSS, nutrients -N and P -), CheckToxicity (heavy metals and other inhibitory substances), UpdateDB, InformDischargeCharacteristics, CalculateDischargeCost	CheckInfluent (flow, pollution, toxicity), SuperviseControl (ATL[1]), InformState, GiveAlarms, OperateSetPoints
Permissions:	Reads: chemical sensors, flow meters, toxicity indicators, tests Updates: DataBases (IndDB, ComDB, wwtpEffDB)	Reads: entrance sensors, on-line and off-line data, Writes: WWTP state, alarms Executes: control commands
Responsibilities		
Liveness:	WWP=(CheckFlow.CheckPollutants. CheckToxicity)$^{\infty}$ ‖ CalculateDischargeCosts $^{\infty}$. UpdateDB$^{\infty}$. InformDischargeCharacteristics	WWT=(CheckInfluent . SuperviseControl)$^{\infty}$. InformState* . GiveAlarm* . OperateSetPoints*
Safety:	True	True

[1] ATL makes reference to the supervisory system to support WWTP operation. For more information see Rodriguez-Roda *et al*. [2002], Comas *et al*. [2003] and Poch *et al*. [2004].

Table 2. Schema for the RiverProtection role

Role:	RiverProtection (RP)
Description:	Updates the ecological status of the river and its dilution capacity (flow, stational variations, maintenance flow…) and informs about it. Collects toxicity opinions and studies from different sources about the possible different source and type of pollution that might arrive to the river. Updates the database available about the different causes-effects of toxicity into rivers.
Protocols & Activities:	CheckRiverQuality, CheckDilutionCapacity, UpdateInfoDB, InformRiverState
Permissions:	Reads: sensors from automatic stations, river quality model, maintenance flows Writes: state of river
Responsibilities	
Liveness:	RP = CheckRiverQuality$^{\infty}$. CheckDilutionCapacity$^{\infty}$. UpdateInfoDB* . InformRiverState*
Safety:	RiverQuality ∈ quality goals: very good, good or moderate categories DilutionCapacity = Discharged Flow (annual average) / RM Flow (maintenance flow) < 50 %

4 THE SERVICES MODEL FOR THE WASTEWATER SCENARIO

Table 3 describes the service model. Take notice that for each service there is a protocol described in terms of a role(s) responsible for starting the interaction (*initiator*), a responder role(s) with which the initiator interacts (*partner*), the information used by the initiator role while enacting the protocol (*inputs*) and the information supplied by the protocol responder during interaction (*outputs*).

Taula 3. Service model for the wastewater management scenario

SERVICE	INITIATOR	PARTNER	INPUT	OUTPUT
InformDischarge Characteristics	WWP	WWCD WWT	Results of CheckFlow, CheckPollutants and CheckToxicity activities	Values for (Flow, Pollutants, Toxicity) characteristics
InformState	WWT	RP	ATL data acquisition and knowledge management modules	State of the plant
GiveAlarms	WWT	RP, WWP WWCD	ATL simulation module	Alert if a problem occurs into to the plant
InformRiverState	RP	WWT	River Quality data acquisition	Value for ecological status Value for dilution capacity effect

Notice that, in Table 3, some of the services' inputs are indeed the results of some activities performed by a specific role. Although activities do not require a protocol (there is not an initiator – partner tuple), it might be interesting to describe them in order to give fully comprehension to the services model. Thus, the activities on Table 4 are those of the WWP, WWT and RP roles. For example, the results of the activities *CheckFlow*, *CheckPollutants* and *CheckToxicity* performed by the WasteWaterProducer are the inputs for the protocol *InformDischargeCharacteristics*, since the initiator role will transfer them to the partner. So, in this case, there is a separation in between the fact of checking and acquiring on-line and of-line data from the fact to pass and give this information.

Table 4. Activities performed by the WWP, WWTP and RP roles

Activity *(role)*	**INPUT**	**OUTPUT**
CheckFlow *(WWP)*	On-line flow sensors, flow databases (flow records) from Industry, Community, WWTP and Tanker Agents. Tanker capacity.	Wastewater flow rates (m^3/hour, m^3/day, m^3/year). Volume in the tank.
CheckPollutants *(WWP)*	On-line data, off-line data (from tests)	Values for the next parameters: BOD, COD, TSS, TN, TP, KjN, $N_NH_4^+$, Temperature, pH
CheckToxicity *(WWP)*	Toxic tests	Values for equitox and/or specific pollutants

Activity (role)	*INPUT*	*OUTPUT*
CalculateDischargeCosts *(WWP)*	The price is a function of the quantity and quality wastewater parameters	Price (€/m^3 or total €)
CheckInfluent *(WWT)*	On-line flow sensors	Characteristics of influent water at wwtp (flow, pollution, toxicity)
OperateSetPoints *(WWT)*	ATL[1] control module	Change setpoint (DO, bypass, Qrec, Q_{WAS}_primary, Q_{WAS}_secondary)
CheckRiverQuality *(RP)*	Model for the river	Value for the ecological status of the river
CheckDilutionCapacity *(RP)*	Discharged flow and maintenance flow	Value for dilution capacity effect (annual average of discharged flow/maintenance flow of receiving media)

Note: BOD: Biological Oxygen Demand, COD: Chemical Oxygen Demand, TSS: Total Suspended Solids, TN: Total Nitrogen, TP: Total Phosphorous, KjN: Kjehdal Nitrogen, $N_NH_4^+$: ammonia nitrogen, Temp: temperature, DO: Dissolved Oxygen, Qrec: Recirculated Flow, Q_{WAS}_primary: waste activated sludge form the primary clarifiers, Q_{WAS}_secondary: waste activated sludge from the secondary clarifiers.

5 THE ACQUAINTANCE MODEL FOR THE WASTEWATER SCENARIO

In the previous section the type of information and the result of the communication protocols was described, whereas inhere the focus is on the flow of information, that is who can communicate with whom. Notice that the information provided in this section is consistent with the protocols described previously (in the sense that the interchanged information between roles when performing the communication protocols is coherent with the communication that can be established in between the specific agents that play such roles). To describe the interrelationships between the agents a parallelism with the real water flux is helpful. So in this case, we infer that the agents representing the origin of wastewater (*i.e.* industries, communities) communicate with the possible agents representing the reception of this wastewater (*i.e.* sewer, tanker or WWTP); the same way, this second level of agents is controlled by an upper level who can observe all the levels above them (*i.e.* the River Consortium, an entity in charge of sanitation infrastructures of a particular catchment and representing the interests of the river provided by other interested stakeholders such as ecologists or other environmental entities) (see Figure 3).

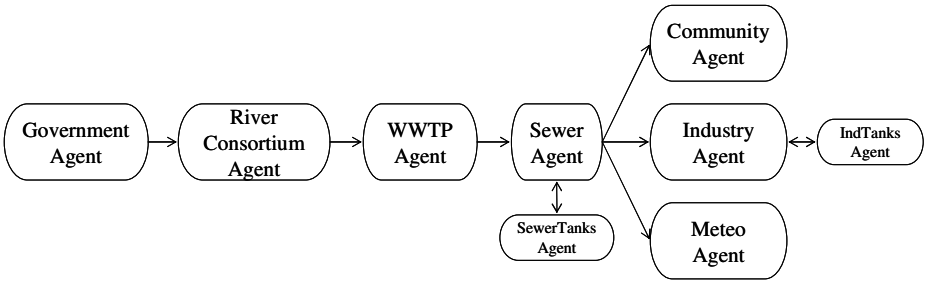


Figure 3. The acquaintance model for the wastewater system. Arrows show the possible relationships between agents; their direction indicates that agents can receive information – maintain communication – with agents in the above levels.

At catchment scale, the agents' metaphor works and it is clear that it helps to depict all the relevant elements of this scenario. This is a challenging, that of combined multirisk planning, entails a demand for high level of collaboration and inter-operation among the involved actors and information entities. In this sense we are in line with some EU-funded efforts, like ORCHESTRA [Denzer *et al.*, 2005], that are open architectures to support interoperability in environmental risk management.

6 CONCLUSIONS AND CHALLENGES

UWS activities involve multiple organizations at various administrative levels, each one having their own systems, services and interests. In many cases, the capacity (and will) to share relevant information between organizations is limited. In the best of the cases is forced by law. This limits the chances of preventing the impact of human activities in the river. The use of Agents is meant to challenge the problems related with information sharing and to improve the interoperability among actors in order to support better coordination and more informed decision-making.

The results of applying the agent-oriented design for the UWS are three different but interrelated models: the agent model, the service model and the acquaintance model for the wastewater scenario. Such design, in terms of agents, adds several useful aspects to the modelling tasks *w.r.t.* having simply objects or elements. Some of the advantages encountered while conceptualizing the urban wastewater system using an agent-based oriented design are the following:

1. It provides a way to better integrate data and information from heterogeneous sources and to better distribute the data.
2. It provides the possibility to establish a more direct and natural communication and coordination between the elements and the possibility to keep the history of the course of the interaction between them. However, although communication between the agents can multiply their effectiveness [Wooldridge 2001; Müller 1996], for each specific implementation it should be evaluated to what extent and for what tasks inter-agent communication significantly improve the desirable performance in comparison with other local methods for conflict resolution.
3. It provides a major abstraction and consequently a higher adaptability to the environment and to changing conditions, thanks to the capacity of multi-agent systems to accept new elements (*i.e.* new industries entering the system could be easily modelled according the abstractions provided here for industry agent and roles).
4. Moreover, the possibility to coordinate their actions by working in a cooperative fashion adds a greater value than the one that can be obtained from any individual or even integrated mechanistic model. For instance, in the urban wastewater domain, some activities need to be coordinated because of shared resources (*i.e.* the WWTP), or because some activities depends upon others activities (*i.e.* industrial discharges require a permission from water authorities) or just because, intuitively, working proactively agent's self-utility increase.

We have shown that agent-methodologies are a useful tool to analyse the system and better explain *how* it should work. Also, are important to raise awareness of the advantages and problems of open systems. The development of agents, independently of their complexity, help to more accurately describe the activities and processes occurring into the system. In special, those related with: Accessibility to information, Interoperability and Coordination. However, work still need to be done in order to fully implement the system in the urban wastewater management scenario.

Evaluation phase will be based on the execution of existing and well-documented cases of discharges that will be used as gold standards. In the long run the evaluation will be performed on a real basin.

Apart from using the models depicted in this paper to evaluate, through agent-based simulation, some (behavioural) aspects of the UWS agents, some of the future challenges of this research include the construction of agent reasoning patterns in order to introduce argumentation over a proposed action among the involved agents.

ACKNOWLEDGEMENTS

Authors would like to acknowledge support from the EC funded project SHARE-it: Supported Human Autonomy for Recovery and Enhancement of cognitive and motor

abilities using information technologies (FP6-IST-045088). The views expressed in this paper are not necessarily those of the SHARE-it consortium. The authors want to thank as well Dr. Ioannis Athanasiadis and Dr. Andrea Rizzoli (researchers at IDSIA, Switzerland) for the useful discussions about the topic.

REFERENCES

Athanasiadis I.N. A review of agent-based systems applied in environmental informatics, *MODSIM 2005 Int'll Congress on Modelling and Simulation*, Melbourne, Australia, 1574-1580, 2005.

Biswas P.K. Towards an agent-oriented approach to conceptualization. *Applied Soft Computing* 8(1):127-139, 2008.

Butler D. and M. Schütze. Integrating simulation models with a view to optimal control of urban wastewater systems, *Environmental Modelling & Software* 20(4): 415-426, 2005.

Comas J., I. Rodríguez-Roda, M. Sànchez-Marrè, U. Cortés, A. Freixò, J. Arràez and M. Poch. A knowledge-based approach to the defloculation problem: integrating on-line, off-line, and heuristic information. *Water research*, 37(10):2377-2387, 2003.

Denzer R., R. Güttler, G. Schimak, T. Usländer, M. Atkinson. ORCHESTRA – Development of an Open Architecture for Risk Management in Europe. *International Symposium on Environmental Software Systems (ISESS 2005)* Sesimbra, Portugal, http://www.isess.org/, 2005.

Fu G., D. Butler and S.T. Khu. Multiple objective optimal control of integrated urban wastewater systems. *Environmental Modelling & Software* 23(2):225-234, 2008.

Jennings N.R., K. Sycara and M. Wooldridge. A Roadmap of Agent Research and Development. *Autonomous Agents and Multi-Agent Systems*, 1(1):7-38, 1998.

Luck M., P. McBurney, O. Shehory and S. Willmott. Agent Technology: Computing as Interaction: A Roadmap for Agent Based Computing. University of Southampton. AgentLink III, ISBN 085432 845 9, 2005.

Müller J.P. The Design of Intelligent Agents. A layered Approach. *Lecture Notes in Computer Science*, *Vol. 1177: Lecture Notes in Artificial Intelligence*. Springer-Verlag, Germany, 1996.

Numi Tran Q.-N. and G. Low. MOBMAS: A methodology for ontology-based multi-agent systems development. *Information and Software Technology*, In Press, 2007.

Nwana H.S. Software agents: an overview, *Knowledge Engineering Review*, 11(3):205-244, 1996.

Pleau M., H. Colas, P. Lavallée, G. Pelletier and R. Bonin, Global optimal real-time control of the Quebec urban drainage system, *Environmental Modelling & Software* 20(4):401-413, 2005.

Poch M., J. Comas, I. Rodríguez-Roda, M. Sànchez-Marré and U. Cortés, Designing and building real environmental decision support systems, *Environmental Modelling & Software* 19(9):857-873, 2004.

Rizzoli A. and W.J. Young, Delivering environmental decision support systems: software tools and techniques. *Environmental Modelling & Software* 12(2-3):237-249, 1997.

Rodriguez-Roda I., M. Sànchez-Marrè, J. Comas, J. Baeza, J. Colprim, J. Lafuente, U. Cortés and M. Poch, A hybrid supervisory system to support WWTP operation: implementation and validation. *Water science and technology*, 45(4):289-297, 2002.

Russell S. and P. Norvig. Artificial Intelligence: A Modern Approach. Prentice Hall Series in Artificial Intelligence. Pearson Education, Inc., Upper Saddle River, NJ, USA, second edition, 2003.

Schütze M., A. Campisano, H. Colas, W. Schilling and P. Vanrolleghem. Real time control of urban wastewater systems – where do we stand today? *Journal of Hydrology* 299:335-348, 2004.

Vanrolleghem P.A., L. Benedetti L. and J. Meirlaen. Modelling a real-time control, of the integrated urban wastewater system. *Environmental Modelling & Software* 20(4):427-442, 2005.

Wooldridge M., An Introduction to MultiAgent Systems. John Wiley & Sons, Inc. New York, NY, USA, 2001.

Zambonelli F., N.R. Jennings and M. Wooldridge .Developing multiagent systems: the GAIA methodology. *ACM Transactions on Software Engineering and Methodology* 12(3):317–370, 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

Improving waste water treatment quality through an auction-based management of discharges

Javier Murillo[a], Dídac Busquets[a], Jordi Dalmau[b], Beatriz López[a] and Víctor Muñoz[a]

[a]Institut d'Informàtica i Aplicacions,Universitat de Girona, Campus Montilivi, 17071 Girona, Spain, {jmurillo, busquets, blopez, vmunozs}@eia.udg.edu

[b] Laboratory of Chemical and Environmental Engineering, Universitat de Girona, Campus Montilivi, 17071 Girona,Spain, jordi@lequia.udg.edu

Abstract: This paper proposes the use of an auction process in which the capacity of a Waste Water Treatment Plant (WWTP) is sold to coordinate the industrial discharges. The main goal of coordination is to manage the wastewater inflow rate and pollutants to improve the WWTP operation. The system is modeled as a multi-agent system where each industry is represented by an agent, another agent represents the influent coming from the domestic use and one more agent represents the WWTP. When the maximum level of the flow or the maximum concentrations of some components exceed the plant's capacity, an auction starts. In the auction, the WWTP agent is the auctioneer that sells its capacity (resources) and the industry agents are the bidders that want to buy the resources. In the auction process the bidders send their bids to the auctioneer and the auctioneer decides which are the winners. The winners will discharge to the sewage system and the losers will have to wait for the next opportunity. After the coordination process, the resulting wastewater discharge schedules of the industries have been analyzed using the IWA/COST simulation benchmark as a case study. The results obtained through this simulation protocol show that the auction-based coordination mechanism using both pollution and hydraulic capacity constraints accomplishes the goal of improving the effluent quality, achieving a reduction in the impact of industrial discharges up to 20,99%.

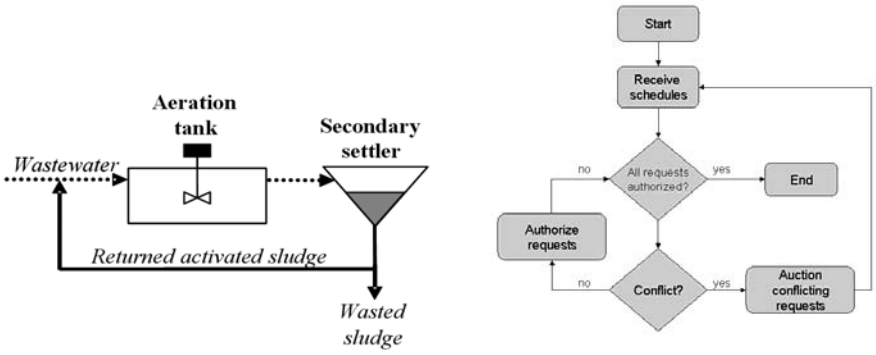
Keywords: Discharge coordination; Integrated management; Water quality; Auction mechanisms; Benchmark; BSM.

1. INTRODUCTION

A Waste Water Treatment Plant (WWTP) receives the polluted wastewater discharges coming from different industries. Nowadays the most common wastewater treatment is the activated sludge process. The system consists in an aeration tank in which the microorganisms responsible for treatment (i.e. removal of carbon, nitrogen and phosphorous) are kept in suspension and aerated followed by a liquid-solids separation, usually called secondary settler. Finally a recycle system for returning a fraction of solids removed from the liquid-solids separation unit back to the reactor, whereas the other fraction is wasted from the system (Metcalf and Eddy [2003]; Figure 1(a)).

The treatment capacity of the plant is limited, therefore all pollutants arriving at the WWTP should be under certain limits; otherwise, the wastewater could not be fully treated and the river would be polluted. Currently, there exist regulations intended to achieve this goal by assigning a fixed amount of authorized discharges to each industry. However, they are not sufficient to guarantee the proper treatment of the wastewater. The problem is that, although these regulations enforce industries to respect the WWTP capacity thresholds, they do not

take into account that simultaneous discharges by different industries may exceed the WWTP's thresholds. In such a case, no industry would be breaking the rules, but the effect would be to exceed the WWTP capacity. Besides, industry discharges add complexity to the waste water treatment system, given that the high variability in influent pollutants composition hampers the WWTP operation because they must discharge under certain limits.



(a) (b)

Figure 1. Activated sludge process (a). Coordination system (b).

As an alternative and more flexible regulation mechanism, we propose the use of an auction process in which the capacity of the WWTP is sold. Auctions are a popular mechanism in economy, usually used to distribute shared resources among different agents (Chevaleyre et al. [2006]). Auctions are currently being used in several industrial scenarios (Bichler et al. [2006]), as the electricity market in which different kinds of energies are auctioned in order to favour the use of non pollutant sources of energy. Recently, auctions have been also considered to deal with natural resources, as CO_2 emissions. In these models, each industry bids for CO_2 emission credits, in such a way that high pollutant industries pay for a lot of emission credits (unless they install some kind of filters in their factories), while industries with non pollutant processes do not need any emission credit, keeping their manufacturing process at a lower cost. Moreover, this approach is in line with a more integrated management of the river basin (Butler and Schütze [2005]) taking into account not only the plant, but also the rest of the components of the treatment system, such as the industries and their discharges. In this context, it seems suitable to raise the possibility of using auctions to deal with wastewater resources, as in a WWTP. This paper presents a first approach to this possibility.

The paper is organized as follows. In Section 2 we present the coordination system, describing in detail each of the involved steps. The implementation of a first prototype is described in Section 3. In Section 4 we discuss the experimental results obtained through simulation and Section 5 derives some conclusions.

2. AUCTION-BASED MANAGEMENT

The wastewater treatment problem could be solved using a centralized approach, where given all the planned discharges from the industries, a new schedule for each of them would be generated, in a way that the capacities of the plant are not exceeded at any time. Centralized approaches imply that a central scheduler would make all the decisions. However, such decisions should be made by each of the industries, since they may not be willing to disclose private information related with the production process upon which their decisions are based. Thus, in order to preserve privacy, other coordination mechanisms should be considered.

In the waste water treatment scenario there is one central element, the treatment plant, who assumes the role of coordinating the discharges of the industries. Then, the plant's capacities are modeled as individual resources, shared by all the industries. Each time a conflict in a

resource occurs (i.e. the capacity is violated) an auction is started in order to determine which of the conflicting discharges will be authorized to discharge and which will have to be delayed. We have chosen auctions as they are a well-known mechanism to distribute shared resources among competing agents when information privacy is a concern.

We assume that each industry has a retention tank of a given capacity, where it can store a discharge whenever it is not authorized, and empty it later on. We also assume that each industry can estimate in advance the discharges that it will generate according to the production process. Although this estimation may differ from the real discharges, they help in the process of coordinating all the discharges and so, adjust properly the WWTP.

In the next two sections we explain in more detail how we have modeled the WWTP scenario as a multi-agent system, as well as the coordination process based on auction mechanisms.

2.1 Multi-agent modeling

Our system, modeled as a multi-agent system, reflects the physical separation between the participants (the plant and the different industries) and also supports privacy in the decision making process of each of the involved agents. Multi-agent systems allow the implementation of complex interactions among the different agents through an appropriate coordination mechanism. In our WWTP scenario, the WWTP agent is the agent who owns the resources (hydraulic and pollution capacities) and the industry agents want to use them.

The process for coordinating the different discharges coming from the industries is depicted in Figure 1(b). Firstly, the industries inform the treatment plant about their scheduled discharges. These schedules contain the set of discharges that they plan to perform in a given period of time, and for each discharge the information about its starting time, duration, flow and contaminant levels is also included. Hence, a schedule from an industry k is described as $S_k=\{d_1,...,d_n\}$ where n is the number of discharges contained in the schedule and each discharge d_i is defined as $d_i = \{s_i, t_i, \bar{q_i}\}$, where s_i stands for the start time, t_i is the duration and $\bar{q_i}$ is a vector containing the flow and contaminant levels of the discharge. The start time of discharges can be modified depending on the industries location: when a difference between the discharge time and the discharge arrival to the WWTP exists this delay should be added to s_i.

The WWTP agent, upon reception of all the industries' discharge schedules for a given day (or any different predefined period of time), starts checking for conflicts. A conflict arises when the discharges planned to be performed at a given time violate any constraints (see Section 3.2 for their definition). Whenever a conflict is detected, the involved industries (the industries whose discharges are scheduled at the time of the conflict) are informed about it, and an auction is started to solve it, forcing industries to modify their schedules. The resolution is done in a sequential way, treating one conflict at a time in chronological order. This process is repeated until all the discharges have been authorized, and the result is that each industry has a new schedule, and these resulting schedules do not produce any conflict.

The unauthorized discharges should not cause problems in the production processes of the industries. In case an industry agent has to reschedule its discharges, its behavior is the following: it first tries to store the rejected discharge into the tank; the discharge of the tank is then scheduled as the first activity of the agent once the current conflict has finished. Conversely, if the industry has its tank already full, the discharge will be performed anyway[1]. However, the influent coming from the domestic use does not have any retention tank and, consequently, its discharges cannot be modified.

Note that it is not necessary to know in real-time the industrial discharges, since the coordination process is done offline, for example one day before.

[1] It is possible to minimize these situations in the auction mechanism, following for example Muñoz et al. [2007]

2.2 Auction mechanism

Once the involved discharges in a conflict have been detected, their corresponding agents (industries) are informed about the conflict and the auction process begins. The WWTP agent assumes the *auctioneer* role who is in charge of selling its flow and pollution capacities resources and the industry agents assume the *bidders* role. The goal of the auction is to select a subset of industries, which will be authorized to perform their discharges, while the remaining should have to be delayed (stored in the tank). The selection criterion is based on the bids submitted by the agents. These bids represent the urgency that each of them has to perform the discharge. A high bid indicates that the agent really needs (or wants) to perform the discharge, while a low bid indicates that the agent could delay the discharge and therefore it can miss the opportunity to perform it at the auctioned time. For example the bid value v_i of the industry agent i could be calculated dividing tank occupation of industry by the total tank capacity of industry.

Note that the auction process allows industries to express their interest of discharging at a given time through the bidding policy, conversely to other centralized approaches that forces industries to discharge at a given time. Even that we have used the tank capacity for bidding generation, other policies can be implemented according to the industries strategies (prices, etc.). Then, the auctioneer clears the auction (i.e. determines which discharges to authorize) by solving the Winner Determination Problem (WDP) (Kalagnaman and Parkes [2005]). Particularly, since the auctioneer offers multiple (but limited) units of different items, and bidders submit bids for a certain number of units of each item, we are dealing with a multi-unit combinatorial auction whose WDP is modeled according to the Equation 1.

$$\max \sum_{i=1}^{NC} x_i \cdot v_i$$
$$s.t. \sum_{i=1}^{NC} x_i \cdot q_{i,j} \leq Q_j \qquad \forall_j \in C \tag{1}$$

where NC is the number of conflicting discharges, $x_i \in \{0,1\}$ represents whether discharge i is denied or authorized, $v_i \in \Re^+$ is the bid value for discharge i, $q_{i,j}$ is the capacity requirement of the resource j for the discharge i, Q_j is the resource j capacity and C is the set of resources. In our case, the items represent the flow and the contamination levels whose available units are determined by the capacities of the plant. This problem is similar to the multi-dimensional Knapsack problem (Kelly [2005]).

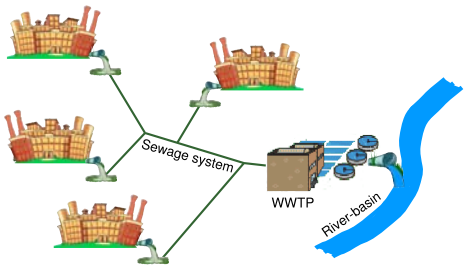


Figure 2. Case study.

3. CASE STUDY

Our case study is a typical wastewater treatment system with 4 industries (depicted in Figure 2). The industries discharge their wastewater to the sewage system, which directs it to the WWTP. The plant, once the wastewater has been treated, puts it back to the river. The first industry is pharmaceutical, which is increasing its discharge flow during the week and does not discharge during the weekend. The second one is a slaughterhouse that discharges a constant flow, except at the end of the day when it increases. The third one is a paper

industry that discharges a constant flow during the seven days of the week. The fourth one is a textile industry, whose discharges flow oscillates during the day. Industries discharges are added to the effluent of the city which is fixed and cannot be coordinated. Altogether represent the WWTP influent. Thus, the output of the auction system is a set of coordinated discharges that is entered as input for the WWTP.

3.1 IWA/COST Simulation Benchmark

As a model for the WWTP, the IWA/COST Simulation Benchmark has been used. This simulation protocol has resulted in more than 100 publications worldwide (Jeppsson et al. [2006]) and provides several tools such as control evaluation, prediction, estimation of biomass activities and effluent quality parameters. The Benchmark Simulation Model N1 (BSM1) layout contains the Activated Sludge Model N1, ASM1 (Henze et al. [1987]) for two anoxic and three aerobic tank reactors, followed by the Takács ten-layer model for the secondary settler (Takács et al. [1991]). Model influent files include 14-days weather disturbances (i.e., dry, rain and storm weather) with 15-minutes sampling. In our case the dry BSM1 influent has been selected and modified with a set of real data provided by the Laboratory of Chemical and Environmental Engineering (LEQUIA) to represent industries discharges. Among the outputs the model provides the Effluent Quality Index (EQI) and the Influent Quality Index (IQI), both are a weighted calculation of the amount of pollutants (i.e. carbon and nitrogen in different forms) present in both the influent and the effluent of the model. They are used for the plant performance evaluation based on the total kilograms of pollutants present in the effluent and the influent (Copp [2002]). We have essentially considered the EQI and IQI values for the evaluation of the auction system. Likewise, the outcome of the benchmark permits us to observe the effects of the coordination mechanism on the quality of the treated wastewater.

3.2 Constraints

In order to coordinate the industrial discharges and improve the effluent quality, the following constraints have been defined:

- **Hydraulic capacity constraint**. This constraint ensures that the total flow arriving to the WWTP (from the influent and industrial discharges) does not exceed a certain threshold at any time. This threshold is called the Maximum Flow (MF).

$$\forall t \in [0,T]: \sum_{i=0}^{N} flow_{i,t} \leq MF \tag{2}$$

  where T is the final time of the simulation, N is the set of industries (including the influent) and $flow_{i,t}$ is the flow discharged by industry (or influent) i at time t.

- **Pollution constraints**. These constraints ensure that the concentrations of certain components arriving to the WWTP do not exceed their respective thresholds at any time. There are 4 constraints, for the following components: COD (Chemical Oxygen Demand), BOD (Biological Oxygen Demand), TKN (Total Kjeldhal Nitrogen) and TSS (Total Suspended Solids). Each constraint is defined as follows:

$$\forall t \in [0,T]: \forall c \in C: \frac{\sum_{i=0}^{N} conc_{i,c,t} \cdot flow_{i,t}}{\sum_{i=0}^{N} flow_{i,t}} \leq MC_c \tag{3}$$

  where C is the set of components (COD, BOD, TKN and TSS) and $conc_{i,c,t}$ is the concentration of component c produced by the industry i or influent at time t. MC_c is the maximum concentration threshold for component c.

4. EXPERIMENTAL SETUP

To test the coordination mechanism we have implemented a prototype of the system, using Repast[2], a free open source software framework for creating agent based simulations using Java language. The simulation reproduces the process and the communication between the plant and the industries performing discharges. We have created one agent to represent the plant and another for each of the industries. The free linear programming kit GLPK[3] has been used to solve the winner determination problem related to each auction.

In order to compare the results obtained with and without auction-based management, three different scenarios have been defined.

- The first scenario considers the influent (dry) without industrial discharges. This scenario shows the wastewater IQI and EQI when simulating with only domestic wastewater (i.e. BSM1 default influent file).
- The second scenario adds to the influent the industrial discharges, without using the auction-based management mechanism. This scenario is useful to calculate the impact of industrial discharges in the WWTP effluent. None of the discharges violate current legislation but there is a deterioration of water quality (increase of IQI and therefore EQI) due to industrial discharges. We have measured such deterioration.
- The third scenario is the same as the second one but using the auction-based management mechanism. This scenario is useful to determine the benefits of the proposed system in terms of EQI.

5. AUCTION-BASED MANAGEMENT MECHANISM PERFORMANCE

The experimental results have been obtained with the simulation protocol BSM1 in the three different scenarios previously described. In order to evaluate the system the effluent quality has been considered.

		IQI $\left(\frac{Kg.poll \cdot unit}{day}\right)$	***EQI*** $\left(\frac{Kg.poll \cdot unit}{day}\right)$	***Increment*** $\left(\frac{Kg.poll \cdot unit}{day}\right)$	***Reduction***	***Constraints***
Influent (scenario 1)		42042,81 (20195,42)	7556,54 (2219,93)	-	-	-
Influent and industries	***w/o auction based mng. (scenario 2)***	59092,21 (20080,39)	9127,37 (2303,71)	1570,83	-	-
	With auction based mng. (scenario 3)	59163,90 (16527,12)	8958,64 (1935,64)	1402,11	10,74%	R1
		59139,67 (16225,36)	8901,23 (1931,82)	1344,70	14,40%	R2
		59157,73 (15854,09)	8900,13 (2073,70)	1343,60	14,47%	R3
		59126,16 (16086,98)	8886,15 (1906,97)	1329,61	15,36%	R4
		59158,62 (18489,36)	9043,95 (2162,55)	1487,42	5,31%	R5
		59132,39 (13378,48)	8797,63 (1729,65)	1241,10	20,99%	R6

Table 1. Results obtained with the benchmark BSM1 in different scenarios.

Table 1 shows the results obtained with the different simulations. The first column is the Influent Quality Index (*IQI*) measured with BSM1, integrating the last seven days of

[2] REPAST Agent Simulation Toolkit, http://repast.sourceforge.net
[3] GLPK Gnu Linear Programming Kit, http://www.gnu.org/software/glpk

weather simulation (Copp [2002]) with the standard deviation in brackets. The second column corresponds to the EQI measured by the benchmark with the standard deviation too. The third column (*Increment*) represents the difference in the *EQI* value between the first and the other scenarios; this value represents the impact of the industrial discharges in the wastewater. The fourth column shows the reduction in percentage on the value *Increment* when using the coordination mechanism. Finally, the fifth column (*Constraints*) indicates the set of constraints thresholds used for the coordination.

Six different constraints have been considered. *R1* is related to the maximum flow (MF), setting it to 25000 m^3/day; *R2* sets the Maximum TSS to 275 mg/l; *R3* dictates a Maximum TKN of 55 mg/l; *R4* states BOD = 234 mg/l; *R5* states Maximum COD = 575 mg/l and finally *R6* contains these constraints: Maximum flow = 32500 m^3/day, Maximum TSS = 275 mg/l, Maximum TKN = 50 mg/l, Maximum BOD = 260 mg/l and Maximum COD = 100 mg/l (Copp [2002]).

The results show that the value of *EQI* obtained in the first scenario is 7556,54 Kg poll·unit/day. When the industrial discharges are added to the influent in scenario 2 the value of *EQI* becomes 9127,37 Kg poll·unit/day, therefore the industries are causing an increase of 1570,83 Kg poll·unit/day. The other data of the table corresponds to the executions of the benchmark in the third scenario with coordinated data and using different sets of constraints, showing that when the auction-based management mechanism is used the *EQI* is reduced and consequently, the impact of industrial discharges (*Increment*) is smaller. In the best case (set of constraints *R6*) the impact is reduced up to 20,99%. According to *t-test*, the values of the mean and standard deviation of *EQI* in the second scenario and the best *EQI* obtained in the third scenario (set of constraints *R6*) are statistically extremely significant.

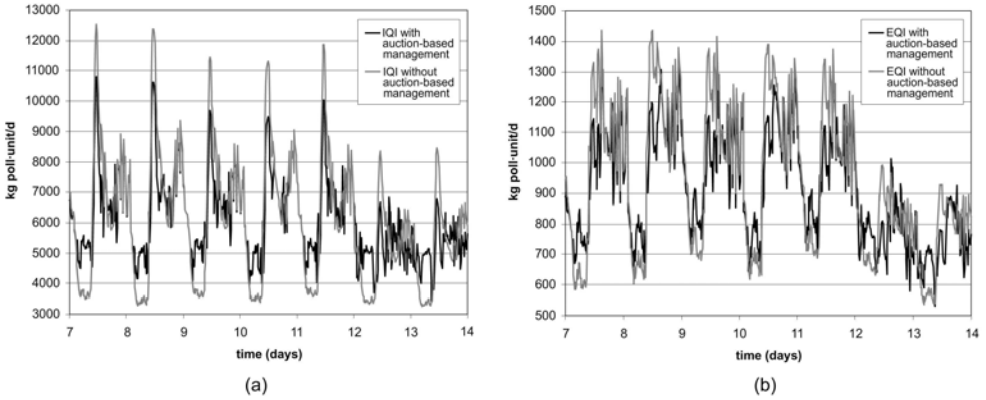


Figure 3. IQI during the 7 days (a), EQI during the 7 days (b).

Figure 3 (a) shows the values of *IQI* during the seven days with and without auction-based management. This picture shows that the auction-based management has lowered the upper values and raised the lower values. The steadiness in the *IQI* profile is important as it improves wastewater treatment by reducing the variability of the influent composition. This circumstance allows the WWTP to process more efficiently the pollutants and consequently achieve better results regarding *EQI*. Figure 3 (b) shows the comparison between the *EQI* values during the seven days with and without auction-based mechanism. Analogously to the Influent Quality Index (*IQI*), the values have also been homogenized.

6. CONCLUSIONS

In this work, the use of an auction-based management mechanism has been proposed in order to coordinate the industrial discharges. In this auction the WWTP assumes the role of the auctioneer that is selling its capacity as a resource, and the industries assume the role of bidders that want to buy the WWTP capacity. The auction determines which industries are going to be allowed to discharge to the sewage system and which are not. This process is repeated each times the hydraulic capacity constraint or the pollution constraints would be violated by the discharges in a given time (if they were not coordinated). The results obtained with the IWA/COST simulation shows that the auction-based management mechanism using pollution and hydraulic capacity constraint reduces the impact of industrial discharges up to 20,99%. This fact has been possible due to the IQI variability reduction, since it has made the WWTP able to process the pollutants more efficiently. Although the simulations do not completely match the treatment system in reality, it is a good starting point for showing to the public authorities and to the industries that the auction-based management approach could help improving both the WWTP operation and the water quality.

ACKNOWLEDGEMENTS

Many thanks to the LEQUIA research group for providing the opportunity to work with them in this problem. This research project has been done with the support of the Commissioner for Universities and Research of the Department of Innovation, Universities and Enterprises of Generalitat of Catalonia and of the European Social Fund and DURSI AGAUR SGR 00296 (AEDS).

REFERENCES

Bichler, M., A. Davenport, G. Hohner, and J. Kalagnanam. Combinatorial Auctions, chapter Industrial Procurement Auctions, pages 593–612. MIT Press, 2006.

Butler, D. and M. Schütze. Integrating simulation models with a view to optimal control of urban wastewater systems. Environmental Modelling and Software. Vol. 20, No. 4, pages 415-426, Elsevier, 2005.

Chevaleyre, Y., P. Dunne, U. Endriss, J. Lang, M. Lemaître, N. Maudet, J. Padget, S. Phelps, J. Rodríguez-Aguilar, and P. Sousa. Issues in multiagent resource allocation. Informatica, 30: 3–31, 2006.

Copp, J. The COST simulation benchmark. Description and simulator manual. Office for Official Publications of the European Communities, 2002.

Henze, M., C. J. Grady,W. Gujer, G. Marais, and T. Matsuo. Activated sludge model no. 1. Technical report, IAWPRC Scientific and Technical Report No. 1, IAWQ: London, UK, 1987.

Jeppsson, U., C. Rosen, J. Alex, J. Copp, K. Gernaey, M.-N. Pons, and P. A. Vanroellegem. Towards a benchmark simulation model for plant-wide control strategy performance evaluation of wwtps. Water Science and Technology, 53(1):287–295, 2006.

Kalagnanam, J. and D. Parkes. Handbook of Supply Chain Analysis in the E-Business Era, chapter Auctions, bidding, and exchange design. Kluwer Academic Publishers, 2005.

Kelly, T. Generalized Knapsack Solvers for Multi-unit Combinatorial Auctions: Analysis and Application to Computational Resource Allocation. LNAI, 3435:73–86, 2005.

Metcalf and Eddy. Wastewater Engineering. Treatment and Reuse. 4th edition. McGraw-Hill, New York, 2003.

Muñoz, V., J. Murillo, D. Busquets, and B. López. Improving water quality by coordinating industries schedules and treatment plants. In AAMAS workshop on Coordinating Agents Plans ans Schedules (CAPS), 2007.

Takács, I., G. Patry, and D. Nolasco. A dynamic model of the clarification thickening process. Water Research, 25(10):1263–1271, 1991.

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

Measuring Predictability of Daily Streamflow Processes Based on Univariate Time Series Model

Wen Wang[a] , Pieter H.A.J.M. Van Gelder[b], and J. K. Vrijling[b]

[a] *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, 210098, China*

[b] *Faculty of Civil Engineering & Geosciences, Section of Hydraulic Engineering, Delft University of Technology, 2628 CN Delft, Netherlands*

Abstract: Predictability is an important aspect of the dynamics of hydrological processes. The predictability of streamflow processes can be estimated based on the multivariate approach, which takes multiple explanatory variables into consideration, or, based on a univariate time series approach, which measures the predictability on the basis of univariate streamflow itself. In this study we investigate the predictability of 31 daily average discharge series with different drainage areas observed in 8 river basins in Europe and northern America on the basis of univariate time series approach. The results show that, although the existence of long memory is detected in the daily streamflow processes, the predictability of the streamflow process is more dominated by short-range autocorrelations than by the existence of long-memory in the streamflow process; the predictability is positively related to watershed scale, that is, the larger the watershed scale, the better the predictability of the streamflow process, and this kind of relationship mainly stems from the positive relationship between autocorrelation and basin scale. Because of the impacts of many factors, the predictability is dynamic rather than invariable, and there are many uncertainties present in the estimation of streamflow predictability.

Keywords: predictability; time series; long memory; ARFIMA model; AR model.

1. INTRODUCTION

Predictability refers to the ability to make predictions of future events based on either past information or a theoretically complete knowledge of the physical system. The predictability of streamflow processes can be estimated based on the multivariate approach, which takes multiple explanatory variables into consideration, or, based on a univariate time series based approach, which measures the predictability on the basis of univariate streamflow itself. The efforts in evaluating predictability of real-world processes so far are in fact concentrating on improving long-term forecasting accuracy by finding better predictors, which is essentially a multivariate approach. Although making forecasts is normally an inevitable step in the procedure of assessing predictability, the issue of forecast accuracy should not be confused with the issue of predictability, and efforts to improve forecasts should not be confused with approaches to understanding the predictability, because predictability is a physical system property that depends on intrinsic dynamics. It is believed [Bloschl and Zehe, 2005] that to learn how to separate the predictable and the unpredictable would be an exciting research field in hydrology in the coming years. It would be also interesting to investigate how each factor, such as El Nin˜o-Southern Oscillation (ENSO) and the Arctic Oscillation (AO) [Hastenrath, 1990; Maurer and Lettenmaier, 2003; Berg and Mulroy, 2006], contributes to the overall predictability of streamflow processes. On the other hand, because river runoff gives an integral measure of

the hydrometeorological conditions in a catchment as a whole, it would therefore be interesting to investigate how the signal of previous streamflow, a basic predictable component of hydrologic systems at watershed scale, contributes to streamflow predictability. However, that is an issue not well recognized.

In addition, the evidences of the presence of long memory property has been revealed in many observed hydrological time series [e.g., Mandelbrot and Wallis, 1969; Hosking, 1984; Wang et al., 2007; Mudelsee, 2007]. The calculations based on comparing the innovation variance and unconditional variance of stationary series suggest that long-memory fractionally integrated autoregressive moving average (ARFIMA) processes often have quite long predictable memory and that fractional integration extends the prediction memory of an ARMA (integrated autoregressive moving average) process [Andersson, 2000]. Hence, a special concern in the present study is that whether we can see long predictability in an observed hydrological series even though we can detect the existence of long-memory.

Because the streamflow flow series is often treated as a univariate time series and can be modelled conventionally with ARMA type models, we therefore take a univariate time series based approach [Wang, et al., 2004] to measure the predictability of univariate streamflow time series in the present study. The paper is organized as follows. In Section 2, the theoretical predictability of ARMA process and ARFIMA process will be analyzed as a basis for comparing with the univariate streamlfow processes. In Section 3, the predictability of 31 univariate streamlfow series at various sites over the world will be analyzed. Some discussions will be given in Section 4 and the results are concluded in Section 5.

2. MEASURING PREDICTABILITY OF A TIME SERIES

2.1 Definition of time series predictability

Clements and Hendry [1998] define a random variable x_t to be unpredictable with respect to an information set Ω_{t-1} if the conditional distribution $Fx_t(x_t\ |\Omega_{t-1})$ and the unconditional distribution $Fx_t(x_t)$ of x_t coincide, i.e. if

$$Fx_t(x_t\ |\Omega_{t-1}) = Fx_t(x_t). \quad (1)$$

This notion of unpredictability implies that the information set Ω_{t-1} does not improve the prediction of x_t. If Ω_{t-1} is restricted to past realizations of x_t, then (1) implies that past realizations do not help to predict x_t.

A weaker condition would only require that the conditional variance of the residual series $\{e_t\}$ equals the constant unconditional variance σ_x^2, i.e.,

$$\mathrm{Var}(e_t|\Omega_{t-1}) = \mathrm{Var}(e_t) = \sigma_x^2. \quad (2)$$

for all t. The mean value as well as the volatility of a time series is said to be predictable if (2) does not hold [Raunig, 2006].

2.2 Theoretical predictability for ARMA process or ARFIMA process with known formula

To see if random variable x_t is predictable, we need to know the conditional variance $\mathrm{Var}(e_t|\Omega_{t-1})$, or simply $\mathrm{Var}(e_t)$, in (2), and the constant unconditional variance σ_x^2. Theoretical formulae are available for calculating both of them for the ARFIMA model and its reduced version ARMA model. Consequently, we refer to the predictability obtained in this way as the theoretical predictability.

The general form of ARFIMA(p,d,q) model is given by

$$(1-\phi_1 B-\mathrm{L}\ \phi_p B^p)(1-B)^d x_t = (1+\theta_1 B+\mathrm{L}\ \theta_q B^q)a_t \quad (3)$$

where $|d| < 0.5$, B is the back shift operator, i.e., $Bx_t = x_{t-1}$ and a_t is i.i.d. with mean zero and variance σ_a^2. When $d = 0$, the ARFIMA model is reduced to an ARMA(p,q) model. If $q = 0$, the ARMA(p,q) model if further reduced to an AR(p) model.

According to the Wold decomposition theorem, under stationarity, the process variance σ_x^2 and h-step ahead optimal forecast error variance $\xi(h)$ of (3) are given as:

$$\begin{cases} \sigma_x^2 = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2, \ \psi_0 = 1 \\ \xi(h) = \sigma_a^2 \sum_{j=0}^{h-1} \psi_j^2, \ \psi_0 = 1 \end{cases} \tag{4}$$

where ψ_j are given by:

$$(1+\psi_1 B+\psi_2 B^2+\mathrm{L}\ ) = \frac{(1+\theta_1 B+\mathrm{L}\ \theta_q B^q)}{(1-\phi_1 B-\mathrm{L}\ \phi_p B^p)(1-B)^d}$$

Granger and Newbold [1986, p. 310] proposed a measure of predictability for covariance stationary series, as the ratio of the variance of the forecast error to the variance of the original time series. According to the definition of Granger and Newbold [1986], predictability is given by $R_{h|x}^2 = 1-\xi(h)/\sigma_x^2$. With the increase of *h*, $\xi(h)$ and σ_x^2 will get close. At a certain lead time *H*, they get sufficiently close so that $\xi(H)\big/\sigma_x^2 < c$, where *c* is a constant close to 1 (say, 0.95), whereas for lead times larger than *H* (i.e., *H*+), $\xi(H+)\big/\sigma_x^2 \geq c$, i.e., $R_{H+|x}^2 \leq 1-c$. Therefore, given a c, we could get a *H*, which satisfies

$$\xi(H) < c\sigma_x^2 \leq \xi(H+1). \tag{5}$$

The estimate given by (5) means that the model forecasts are not more accurate than the mean value of the process after *H* steps at a given level 1–*c*. Consequently, instead of using the general definition in (1) or (2), we can define the predictability of a stationary process more precisely as the predictable horizon *H* after which the prediction is no better than the mean value. Because the predictability defined in this way is based on theoretical formula, we refer it to as theoretical predictability (TP).

2.3 Sample predictability for a process with unknown formula

As mentioned above, TP is calculated based on the known model of covariance stationary series, for which the formula are available for calculating the variance of the process and the variance of multi-step forecast error. However, the true model of a given time series is rarely known, especially for real-world processes. Alternatively, we may take a forecast error based approach, which measures the predictability based on forecast errors. Correspondingly, such type of predictability is referred to as sample predictability (SP), because it is estimated from forecast error samples.

To measure the SP, we may use the coefficient of efficiency (*CE*) proposed by Nash and Sutcliffe [1970] which is widely adopted as a model performance measure in the hydrology community, given by

$$CE = 1 - \frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{n}(Q_i - \bar{Q})^2}, \tag{6}$$

where *n* is the data size, Q_i is the observed value, $\hat{Q}_i$ is the predicted value, $\bar{Q}$ is the mean value of the observed data. When measuring predictability for a observed processes, we should split the entire data into two parts: one calibration data set which is used to build a model, and another validation data set (or out-of-sample data) that is used for calculating *CE* at different lead times. Notice that, in calculating *CE* for measuring predictability, the mean value $\bar{Q}$ should be calculated on the basis of calibration dataset rather than validation dataset, because by definition, the predictability is the degree that the past can be used to predict the future. This is what differs between measuring predictability from evaluating hydrological model performance, in which $\bar{Q}$ is calculated based on the validation dataset.

In fact, the predictability measure $R_{h|x}^2$ proposed by Granger and Newbold [1986] is essentially the same as *CE*. The difference between $R_{h|x}^2$ and *CE* is that the former is

calculated based on theoretical formula whereas the latter is calculated based on forecast errors for sample data. With *CE*, the predictability of a stationary process is re-defined as the predictable horizon *H* after which the prediction is no better than the mean value for the process at a given level $CE = CE_H$, where CE_H is a small value less than 1 (say, 0.1 or 0.5), which is related to *c* by $c = 1\text{-}CE_H$.

3. PREDICTABILITY OF UNIVARIATE STREAMFLOW PROCESSES

3.1 Daily streamflow data used

Daily average discharge series recorded at 31 gauging stations in eight basins in Europe, Canada and USA are analyzed in the present study. The data come from Global Runoff Data Centre (GRDC) (http://grdc.bafg.de), US Geological Survey Water Watch (http://water.usgs.gov/waterwatch), and Water Survey of Canada (http://www.wsc.ec.gc.ca). We generally have the following three rules to select gauging stations in each basin:

(1) The basins where the stations are located covers different geographical and climatic regions;

(2) The drainage area of each station is basically at 5 different watershed scales, namely, > 10^6 km^2; 10^6 ~ 10^5 km^2; 10^5 ~ 10^4 km^2; 10^4 ~ 10^3 km^2; < 10^3 km^2;

(3) The stations are located in the main river channel of the river if possible. When stations at the main channel are not available, stations at major tributaries are used.

For each station, we select a segment of historical daily streamflow records of mostly 30 years long. However, because of data limitation, the shortest series covers a period of only 14 years. The segments are chosen with following criteria:

(1) The series should be approximately stationary, as least by visual inspection. We have stationarity as our primary criterion because, when certain types of nonstationarity are present, many long-memory parameter estimators may fail.

(2) The recording period of the data should be as early in time as possible, assuming that the influence of human intervention would be less intensive in early period (in early 20th century or even late 19th century) than in later period.

(3) The temporal spans of streamflow series at different locations in one basin should be as close as possible, so as to mitigate possible impacts of regional low-frequency climatic variations.

The selected stations and corresponding daily streamflow series are listed in Table 1.

3.2 Calculation of sample predictability of daily streamflow processes

Because *CE* is a measure comparing the predicted values with the overall mean value, it is not effective enough to evaluate the predictions for those series whose mean values change with seasons, which is almost always the case for hydrological processes like streamflow processes. Therefore, a seasonally-adjusted coefficient of efficiency (*SACE*) [Wang et al., 2004] is used here for evaluating the predictability. *SACE* is calculated by

$$SACE = 1 - \frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{n}(Q_i - \bar{Q}_m)^2}, \tag{7}$$

where *m* is the "*season*", $m = i \bmod S$ (mod is the operator calculating the remainder), ranging from 0 to *S*-1; and *S* is the total number of "*season*" (Note that, a "*season*" here is not a real season. It may be a month or a day depending on the timescale of the time series. For daily streamflow series, one season is one day over the year.); $\bar{Q}_m$ is the mean value of season *m*.

In the present study, we use *SACE* to measure SP of daily streamflow processes. Each daily streamflow data series are split into two parts: a calibration data set which is used to build a model, and a validation data set (or out-of-sample data) that is not used for model fitting. $\bar{Q}_m$ for each day in (7) is calculated based on the calibration dataset. We estimate the predictability at two *SACE* levels 0.1 and 0.5, corresponding to $c = 0.9$ and 0.5 in (5).

To measure SP of a process, we need a suitable forecasting model together with the model performance measure. In the present study, we use autoregressive (AR) model for all streamflow series. Each AR model is built based on the calibration dataset, then applied to the validation dataset for make predictions of different lead times.

For each out of the 31 streamflow series, the long-memory parameter (or fractional differencing parameter) *d* is estimated [Wang et al., 2007] with the method implemented in S-Plus. For a stationary long-memory series, $0 < d < 0.5$. The larger *d*, the longer memory the stationary process has.

The results of both the long-memory parameter estimates and predictability estimates that are represented by predictable horizon are reported in Table 1. The following facts could be revealed from Table 1:

Table 1 Selected daily streamflow time series and the estimates of their predictability

Basin	Location of gauging stations	Area (km^2)	Period	ACF(1)	*d*	Predictability	
						SACE = 0.1	SACE = 0.1
Colorado	Colorado River At Lees Ferry	289,400	1922-1951	0.9738	0.4478	98	8
	Colorado River Near Cisco	62,390	1923-1952	0.9627	0.4506	18	5
	Colorado River Near Kremmling	6,167	1904-1918	0.9431	0.4863	17	5
	Williams Fork Near Parshall	476	1904-1924	0.9549	0	34	11
Columbia	Columbia River At The Dalles	613,565	1880-1909	0.991	0.4615	45	15
	Columbia River at Trail	88,100	1914-1936	0.9966	0.4187	26	13
	Columbia River at Nicholson	6,660	1933-1962	0.9778	0.4392	21	9
	Columbia River Near Fairmont Hot Springs	891	1946-1975	0.9676	0.4213	20	9
Danube	Danube river at Orsova	576232.	1901-1930	0.9931	0.2634	55	17
	Danube river at Achleiten	76653.	1901-1930	0.9577	0.3598	26	4
	Inn river at Martinsbruck	1945.	1904-1933	0.9326	0.4059	9	2
Fraser	Fraser River at Hope	217,000	1913-1942	0.9772	0.3878	32	8
	Fraser River at Shelley	32,400	1950-1979	0.9734	0.3529	12	3
	Fraser River at Mcbride	6,890	1959-1988	0.9582	0.1886	9	3
	Canoe River below Kimmel Creek	298	1972-1994	0.9294	0.3100	5	1
Mississippi	Mississippi River At Vicksburg	2,962,974	1932-1961	0.9961	0.3909	73	15
	Mississippi River at Clinton	221,608	1874-1903	0.9921	0.4001	33	13
	Minnesota River At Mankato	38,574	1943-1972	0.9917	0.4847	17	7
	Minnesota River At Ortonville	3,003	1943-1972	0.9563	0	49	10
Missouri	Missouri River at Hermann	1,353,000	1929-1958	0.9711	0.4238	28	4
	Missouri River at Bismarck,	482,776	1929-1953	0.9805	0.4124	15	2
	Missouri River at Fort Benton	64,070	1891-1920	0.9165	0	38	7
	Madison River near McAllister	5,659	1943-1972	0.9522	0	45	7
Ohio	Ohio River At Metropolis	525,500	1943-1972	0.9723	0.2983	15	7
	Ohio River at Sewickley	50,480	1943-1972	0.9547	0.2581	8	2
	Tygart Valley River At Colfax	3,529	1940-1969	0.9291	0.2263	7	2
	Tygart Valley River Near Dailey	479	1940-1969	0.8985	0.3324	1	0
Rhine	Rhine at Lobith	160,800	1911-1940	0.9897	0.4254	46	8
	Rhine at Rheinfelden	34,550	1931-1960	0.9715	0	43	9
	Rhine at Domat/Ems	3,229	1911-1940	0.9048	0.4176	7	1
	Emme River at Emmenmatt	443	1915-1944	0.8739	0.3447	7	1

Note: The estimates of long memory parameter *d* are adopted from Wang et al. [2007].

(1) The predictability is closely related to autocorrelations at lag 1, ACF(1). An exponential relationship between ACF(1) and the predictability at different *CE* levels can be seen by visual inspection at Figure 1.

(2) The estimates of *d* with the *S-MLE* method versus the predictability at two level (SACE = 0.1 and 0.5) are plotted in Figure2 (Note that the zero estimates of *d* are removed in the plot). We can discern a exponential relationship between the two, but the relationship is less clearer than that between the predictability and ACF(1).

(3) There is a positive relation between the predictability and the watershed scale. It is shown in Figure 3a, and Figure 3b that, the larger the watershed scale, the better the predictability. Because it has been found that the relationship between the long-memory parameter d and the basin scale is very weak [Wang et al., 2007], whereas ACF(1) has a good relationship with the basin scale (shown in Figure 4), therefore, it seems that the increase of predictability with the increase of basin scale mainly results

from the increase of autocorrelation with the increase of basin scale.

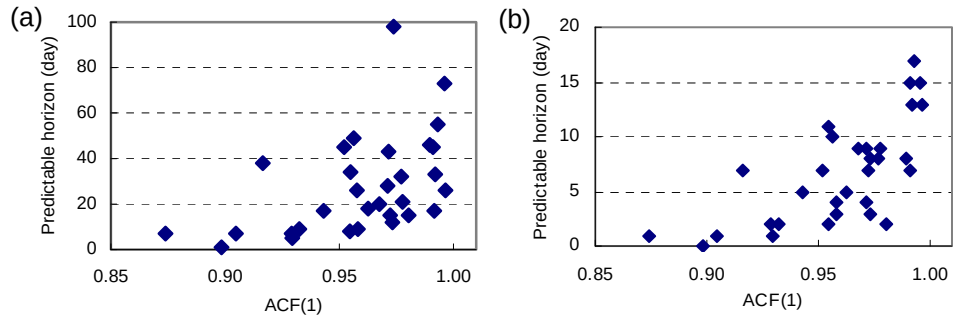


Figure1 Predictability versus ACF(1) for streamflow processes (a) at SACE=0.1 level and (b) SACE = 0.5 level

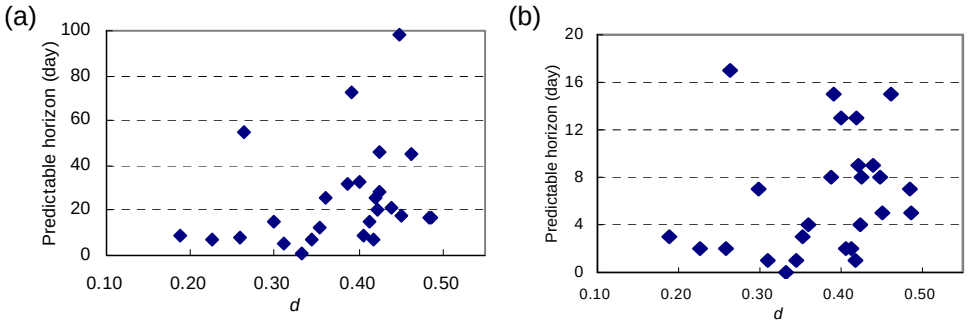


Figure 2 Predictability versus long memory parameter for streamflow processes (a) at SACE=0.1 level and (b) SACE = 0.5 level

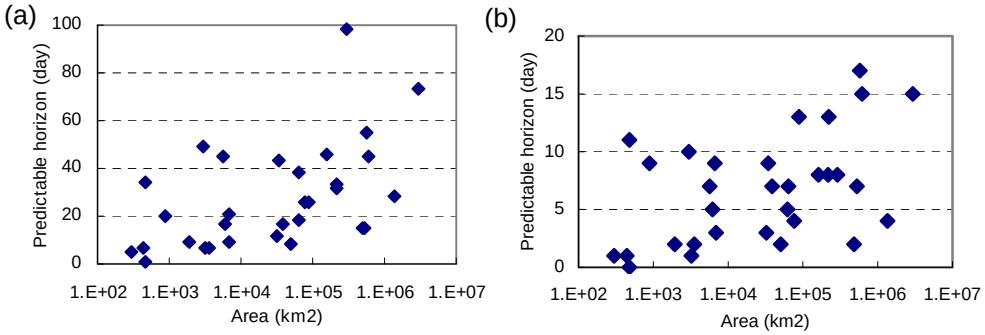


Figure 3 Predictability versus watershed scale for streamflow processes (a) at SACE = 0.1 level and (b) SACE = 0.5 level

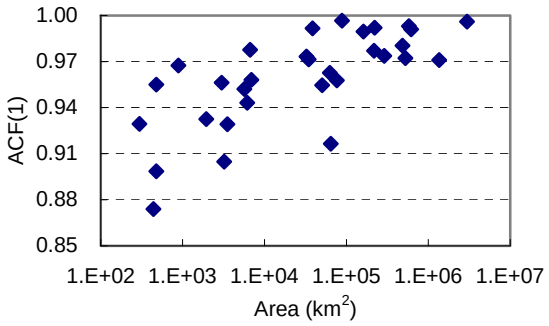


Figure 4 ACF(1) versus watershed scale for streamflow processes

4. UNCERTAINTIES IN THE ESTIMATION OF PREDICTABILITY

Predictability is impacted by many factors, including meteorological factors (such as how river flow is fed, temporal and spatial variability of precipitation processes), and basin characteristics (such as the size, topography, control structures, and drainage network of the basin, and land cover types). These factors may vary from event to event, from season to season, and from region to region, which make the predictability more or less dynamic, rather than an invariable value, albeit it is generally a stable physical feature of a streamflow process.

Apart from the temporal and spatial variability in various factors that impact the predictability, there are several sources of uncertainty in measuring SP of real-world processes, including:

(a) Uncertainty in model selection, especially when different mechanisms may act simultaneously underlying a time series. We suggest using AR model to estimate predictability for the purpose of a "fair" comparison, because of its easiness for using and simplicity for building.

(b) Uncertainty in model parameter estimation, especially when the data size is not big enough. Generally, the larger the data size, the better the parameter estimates because the availability of larger samples allows one to better inspect the asymptotical properties of model parameter.

(c) Uncertainty in data selection. Due to possible long-term variation in climate system and the change of basin characteristics mainly due to human activities, watershed system may exhibit long-term variation. In turn, the mean values of streamflow processes may change over time, thus resulting in the exaggeration of predictability if the mean value of the calibration dataset differs significantly from that of the validation dataset. Therefore, different data selection may give different predictability estimates.

(d) Uncertainty in data quality, which may results in slight exaggeration of predictability. For example, for some gauging stations, the gauged discharges are often the same for two weeks or even over a month continuously, especially when the discharge is very low. For instance, Minnesota River at Ortonville, MN, the discharge kept to be 58 cubic feet per second (cfs) for 17 days (1945.1.12 ~ 1945.1.28); 0.7 cfs for (December 28, 1964 ~ February 5, 1965); 3 cfs for 49 days (1968.1.6 ~ 1968.2.24), 11 cfs for 44 days (1971.1.11 ~ 1971.2.16). While in some cases, this may be true, but most probably this is due to the limited measurement accuracy or even error, which may lead to an slight exaggeration of predictability.

Above uncertainties make it impossible to make precise estimates of predictability, even for the specific approximately stationary period of time that we chose for each streamflow process in the present study.

Due to the seasonality in precipitation and vegetation coverage, streamflow processes usually exhibit strong seasonality, not only in the mean values and variances but also in the autocorrelation structures [see e.g., Wang et al., 2006]. Therefore, the presence of seasonality makes the estimation of predictability of streamflow processes more complicated, and would more or less result in seasonal variation of predictability of streamflow processes. But the seasonality of predictability is not treated in the present study.

5. CONCLUSIONS

In the present study, the predictability of 31 daily average discharge series observed in 8 river basins in Europe and northern America are estimated with univariate time series approach that is based on time series analysis techniques. The results show that, Although the existence of long memory is detected in the daily streamflow processes [Wang et al., 2007], the predictability of these streamflow processes is more similar to that of short-memory AR processes than to that of long-memory ARFIMA processes, indicating that the predictability of the streamflow process is more dominated by short-range autocorrelations than by the existence of long-memory. Possible explanation for that may be in that, the presence of more profound short range memory and seasonality dominate the estimation of predictability albeit the existence of long-memory, which make the effect of long-memory cannot show itself in the predictability measurement. It is also shown that the predictability

is positively related to watershed scale, that is, the larger the watershed scale, the better the predictability of the streamflow process, and this kind of relationship is mainly resulted from positive relationship between autocorrelation and basin scale.

Predictability is impacted by many factors, which make the predictability more or less dynamic, rather than a invariable value. In the present study, the data set were chosen elaborately so as to make the data series approximately stationary and avoid significant impacts of climate change and human interventions (e.g., reservoir impoundment). But many uncertainties still exist, which make it impossible to make precise estimates of predictability.

ACKNOWLEDGEMENTS

The financial supports from the National Science Foundation of China (Project No. 40771039) and the Research Fund for the Doctoral Program of Higher Education (Project No.20060294003) are gratefully acknowledged.

REFERENCES

Andersson, M. K. Do long-memory models have long memory? *International Journal of Forecasting* 16, 121–124, 2000

Berg, A.A., and Mulroy, K.A., Streamflow predictability in the Saskatchewan/Nelson River basin given macroscale estimates of the initial soil moisture status. *Hydrological Sciences Journal*, 51(4), 642-654, 2006.

Bloschl, G., and Zehe E., On hydrological predictability. *Hydrol. Process.* 19, 3923–3929, 2005.

Clements, M. P., and Hendry, D. F., *Forecasting economic time series*. Cambridge University Press, 1998.

Granger, C.W.J., and Newbold, P., *Forecasting Economic Time Series (2nd)*. Academic Press, Orlando, FL., 1986.

Hastenrath, S. Predictability of Anomalous River Discharge in Guyana. Nature, 345: 53-54, 1990.

Hosking, J. R. M., Modeling persistence in hydrological time series using fractional differencing, *Water Resour. Res.*, 20(12), 1898-1908, 1984.

Mandelbrot, B. B., and Wallis J.R., Some long-run properties of geophysical records, *Water Resour. Res.*, 5, 321–340, 1969.

Maurer, E.P., and Lettenmaier, D. P. Predictability of seasonal runoff in the Mississippi River basin. *J. Geophys. Res.*, 108(D16), 8607, doi: 10.1029/2002JD002555, 2003.

Mudelsee M., Long memory of rivers from spatial aggregation. *Water Resour. Res.*, 43, W01202, doi:10.1029/2006WR005721, 2007.

Nash, J. E., and Sutcliffe, J.V., River flow forecasting through conceptual models, I, A discussion of principles, *J. Hydrol.*, 10, 282–290, 1970.

Raunig, B., The longer-horizon predictability of German stock market volatility. *International Journal of Forecasting*, 22(2), 363-372, 2006.

Wang, W., van Gelder, P.H.A.J.M., Vrijling, J.K., and Ma, J., Predictability of streamflow processes of the Yellow River. *In: Proceedings of the 6th International Conference on Hydroinformatics*. Singapore : World Scientific, pp.1261-1268, 2004.

Wang, W., van Gelder P.H.A.J.M., Vrijling J.K., and Ma J., Forecasting Daily Streamflow Using Hybrid ANN Models. *Journal of Hydrology*, 324, 383-399, 2006.

Wang, W., van Gelder P.H.A.J.M., Vrijling J.K., and Chen X., Detecting long-memory: Monte Carlo simulations and application to daily streamflow processes. *Hydrol. Earth Syst. Sci*., 11, 851-862, 2007.