# A hybrid supervisory system to support WWTP operation: implementation and validation

**I.Rodriguez-Roda*, M. Sànchez-Marrè**, J. Comas*, J. Baeza***, J. Colprim*, J. Lafuente***, U. Cortés** and M. Poch***

* Laboratori d'Enginyeria Química i Ambiental, Universitat de Girona, Campus Montilivi, E-17071 Girona, Catalonia, Spain

** Dept. de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Edifici C5-Campus Nord, Jordi Girona 1-3, E-08034 Barcelona, Catalonia, Spain

*** Departament d'Enginyeria Química, Universitat Autònoma de Barcelona, Campus Bellaterra, E-08193 Cerdanyola, Catalonia, Spain

**Abstract** Integrated operation of Wastewater Treatment Plants is still far from being solved. A reasonable proposal should link advanced and robust control algorithms to some knowledge-based techniques, allocating the detailed engineering to numerical computations, while delegating the logical analysis and reasoning to supervisory intelligent systems. This paper describes the development and implementation of a knowledge-based Hybrid Supervisory System to support the operation of a real Wastewater Treatment Plant. The system integrates different reasoning modules, overcoming the limitations in the use of each single technique, while providing an agent based architecture with additional modularity and independence. It is structured into three separated levels: data gathering, diagnosis, and decision support. The different tasks of the system are performed in a seven-step cycle: data gathering and update, diagnosis, supervision, prediction, communication, actuation, and evaluation phase. In spite of certain reservations of the scientific community about the use of these techniques, the system is successfully performing real-time support to the operation of the Granollers facility since September 1999. Results of the first four-month validation period are shown and discussed. An example of the system behavior is also shown in the paper. The conclusions indicate the key steps which are necessary to transfer the system to another facility.

**Keywords** Activated sludge; case-based reasoning; decision support; diagnosis; expert system; knowledge; modelling

Water Science and Technology **Vol 45 No 4–5 pp 289–297** © IWA Publishing 2002

## Introduction

Environmental systems possess several inherent characteristics which make their understanding and control difficult: they evolve over time, involve processes which take place in a 3-dimensional space, are complex, involve interactions between physical–chemical and biological processes, are stochastic, and, very often, are periodic in time (Guariso and Werthner, 1989). The complexity and the multi-faceted nature of many environmental problems, suggests that their suitable management cannot be based on a single technique. Wastewater Treatment Plants (WWTP) are clear examples of complex processes which meet all these distinctive features of environmental systems.

Although progress in control engineering, computer technology, and process sensors has enabled automatic control improvement, integrated operation of WWTP is still far from being solved. The number of measured variables in a WWTP is increasing and the need and possibility to control the process is becoming greater. With this increasing instrumentation there is certainly more information available, but it must be reminded that data rich is not the same as information rich (Rosen and Olsson, 1998). It is not an easy task for operators and process engineers to acquire, to integrate and to understand all this day-to-day increasing amount of information. The solution could arrive with the development of knowledge-based decision support systems that handle the particular characteristics of the process, using this available but incomplete information to guarantee the quality of the discharged water. Although knowledge-based systems came into picture in the 1980s, some authors suggest that they never succeeded for two reasons: they were too complex, and the available knowledge could not be captured in reliable models and advisory systems (Olsson *et al*., 1998). A support system cannot only be based on mathematical modelling, but must also take advantage of heuristic knowledge from literature and experts, while including specific experiences accumulated through years of experience in the own facility. A reasonable proposal should link advanced and robust control algorithms to some knowledge-based techniques, allocating the detailed engineering to numerical computations, while delegating the logical analysis and reasoning to supervisory intelligent systems (Stephanopoulos and Han, 1996).



There have been some approaches to improve WWTP operation using single knowledge-based techniques, such as Baeza *et al.* (2000), R.-Roda *et al.* (1999a), Bow (1998), Wang *et al.* (1997) and Zhu and Simpson (1996). However, an effective decision support system to supervise the operation of the process should be described as a hierarchical multi-level structure that integrates different concurrent modules, overcoming the limitations in the use of each single technique, and providing higher accuracy, reliability and usefulness (Cortés *et al*., 2000). The authors of the paper propose the development of a three level Hybrid Supervisory System (HSS) to support WWTP operation.

The theoretical approach of the HSS was previously discussed in R.-Roda *et al.* (1999b). There is a lower level to gather and structure into a database all the available information generated in the process. These data feed the classical control (automatic or manual), which can entail either a simple control loop or an algorithm based on a mechanistic model of the process. Simultaneously, the modules included in the second level of the HSS have access to the database to infer the possible operating state of the plant. This level has two main knowledge-based reasoning modules: an Expert System (ES) including the heuristic knowledge of the process, and a Case-Based reasoning System (CBS) that reuses the experience from previous particular situations. The partial conclusions of this reasoning phase are sent to the upper level, where the suggested actuation strategies can be evaluated by means of a predictive model. The final conclusion(s) are sent to the operator through the computer interface or, when possible, the strategy can be applied automatically to the plant, modifying the set points of the automatic control device. This HSS acts in supervisory cycles.

This paper describes the implementation of the proposed HSS into a real WWTP. The HSS is performing real-time support to process operation since September 1999. Results of the first four-month validation period are also shown and discussed in the paper. The facility selected to implement the HSS is located in Granollers, in the Besòs river basin (Spain). The water treatment line encompasses preliminary, primary and secondary treatment to remove the organic matter, the suspended solids and, under some conditions, the nitrogen contained in the raw water of about 130,000 equivalent-inhabitants. The sludge treatment line encompasses thickening, anaerobic digestion and dewatering. The raw influent comes from a combined sewer.

## Implementation

The HSS is structured into three separated levels: data gathering, diagnosis, and decision support, providing an agent based architecture with additional modularity and independence (a key factor to guarantee the re-design and transferability of the system to another facility). Figure 1 shows the multi-layer architecture of the HSS, connecting the user (e.g. the head of the plant) with the WWTP. The different tasks of the HSS are performed in a seven-step cycle: data gathering and update, diagnosis, supervision, prediction,

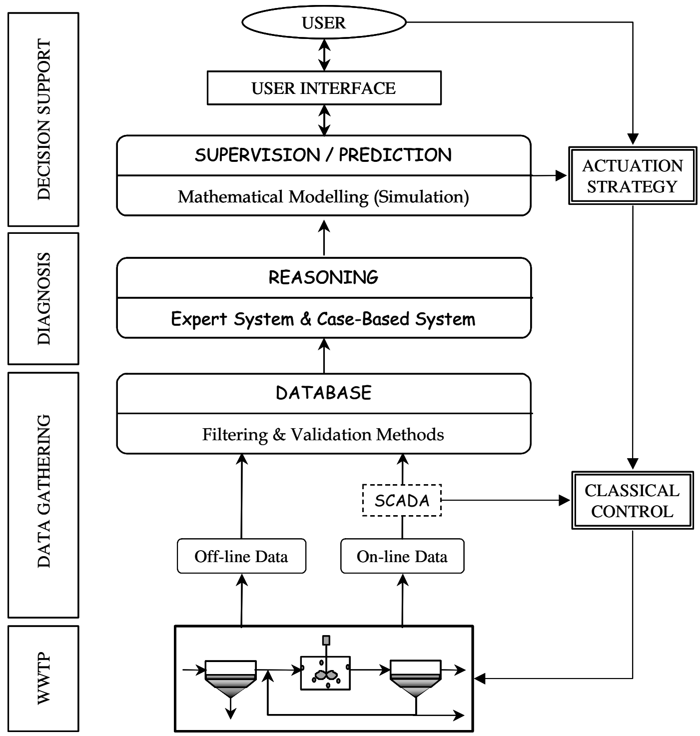


**Figure 1** Three-level architecture of the Hybrid Supervisory System



communication, actuation, and evaluation phase. The cycle is routinely executed once a day, although it can also be started manually at any time, and it is fired up whenever any alarm symptom is fulfilled.

#### First level

The first and lower layer of the HSS encompasses the tasks involved in data gathering and registration into the real-time database. A data acquisition system gathers all kind of data collected in the Granollers WWTP. On-line data, previously acquired by means of an SCADA system (Supervisory Control And Data Acquisition), include measurements from sensors and equipment (210 digital and 20 analogical signals, e.g. flow rate, pH, pump status, etc.). Off-line data include both numerical and qualitative information (i.e. analytical determinations in the laboratory and operator observations of the activated sludge, e.g. chemical oxygen demand, suspended solids, food to micro-organism ratio, protozoa biodiversity, presence of bubbles in the secondary settler, etc., up to 158 variables).

Original raw data are often defective, requiring a number of pre-processing procedures before being registered into the database in an understandable and interpretable way. The first step is the data validation, which includes filtering the correctness of the values (outliers), the noise, and the missing values. Then, all these filtered values are integrated into an homogeneous unit and time-scale. Since WWTP are dynamic systems, the optimal monitoring of the process state must also include information about the evolution of the main variables. Thus, the database calculates and analyses the derivatives of some specific variables to detect sudden deviations, trends, and periodicities. Moreover, some on-line data coming from sensors are reformatted in order to obtain different time scale averaged values and some daily accumulated values.

**Second level**

The second level of the HSS includes the reasoning modules to infer the state of the process, in order to reach a reasonable proposal of actuation to support the whole plant supervision. The result of this layer is based on the combination between the heuristic knowledge of the ES and analogous experiences retrieved by the CBS. Both modules deal with qualitative information, and can interact with classical methods based on modelling and computing.

*Expert system*. The ES was codified into G2 (Gensym, 2000), a commercial shell that includes the inference engine. The Knowledge Base of the ES reflects the problem-solving strategies of the process. The knowledge, codified by means of heuristic rules, was acquired from literature, from experts, and automatically from the database of the facility (R.-Roda *et al.*, 2001).

The structure of the Knowledge Base is modular (see Figure 2), using meta-rules and decision rules as the most appropriately knowledge representation scheme and reasoning strategy to handle sub-goals (a meta-rule is a rule that determines which other module/rules are to be used to solve a problem). Each module has a specific task and consists of different sets of rules, methods and/or procedures.

Once the ES is launched, it receives any requested data from the database. Then, the Data Abstraction module carries out a qualitative abstraction of the whole quantitative data (e.g. IF Suspended-Solids-Effluent > 35 $g \cdot m^{-3}$ THEN Suspended-Solids-Effluent is high). The boundaries among qualitative modalities like "low/normal/high" and "increasing/maintaining/decreasing" were first established by means of a statistical study of each variable. Afterwards, these ranges were submitted to judgement to the head of the plant, who finally adjusted them. Based on this qualitative abstraction, the Meta-Diagnosis module determines which tree and diagnosis paths (i.e. heuristic rules or procedures of the Knowledge Base) must be explored to infer the situation and to suggest an actuation strategy. The Knowledge Base of the system includes more than 700 rules and 200 procedures, which are structured into three main sub-modules: Fault Detection, Operational Problems and Transition States.

- Fault Detection module includes all the knowledge related to 8 operational faults due to mechanical equipment or electrical failures: e.g. damaged or clogged pumps and pipes, electrical fault detection, air system failure, sludge removal system break, etc.
- Operational Problems module includes the knowledge to diagnose 7 primary treatment and 17 secondary treatment problems, e.g. old sludge, storm, overloading, filamentous bulking, low pre-treatment efficiency, foaming, etc. These problems are also divided into biological and non-biological nature depending on the cause of the dysfunction.

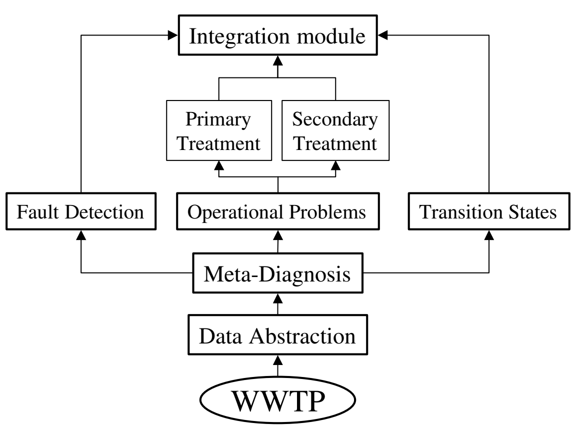




**Figure 2** Modular structure of the Expert System

- Transition States module contains all the knowledge necessary to cope with transient states that can evolve in some of the undesirable problems contained in the previous modules.

Whenever a situation is identified, the diagnosis task of the ES is reinforced with the detection of the cause of the problem. If the right cause of the problem is determined, a specific actuation is recommended. If the cause is not successfully determined, a non-specific actuation must be proposed to soften the effects of the trouble without tackling the real cause.



*Case-based system.* The CBS was codified in Common-Lisp (Allegro CL, 1998). In our approach each case is the codified description of a 24-hour period state or experience within the facility in a storable, easily retrievable way. Each experience is described through the most relevant measurements carried out in the process. The set of specific cases is stored in a structured fashion in a hierarchical case library. The 16 variables selected for our application include water and sludge flow rates, organic matter and nutrient concentration in different locations of the plant, biomass characteristics as settleability, biodiversity, or predominant species, and physical observations as presence of foam in the aeration tank. This selection was done according to the criteria of experts.

Our CBS proposal is based on a working cycle (R.-Roda *et al.*, 1999a) that consists of the following steps: i) gathering and processing data from the process to define the current case, ii) searching the case library and retrieving the case that best fits the current one, iii) adapting the solution if the retrieved case does not perfectly match the current case, iv) applying the adapted solution to the process, v) evaluating its consequences, and vi) learning details about the new experience. To optimise retrieval and learning phases, the case library was organized in a hierarchical discretized tree (Sànchez-Marrè *et al.*, 1997).

Among the whole historical days stored in the data base of the plant, twenty-five days representing typical situations in the Granollers WWTP were selected as initial cases to seed the case library (e.g. storm, filamentous bulking, rising, foaming, overloading, and normal situation). After the 4-month validation period, more than 100 new relevant experiences were learned and stored in the case library.

#### Third level

The third and upper level of the HSS establishes a supervisory and predictive task over the WWTP. The supervision task entails the gathering and combination of the conclusions from both knowledge-based techniques (ES and CBS modules), in order to identify whether there is a problem or not. The final diagnosis, together with the suggested actuation strategy, is sent to the user through the message-board of the computer interface. The expert evaluates the suggestions, checking its validity and deciding which is the best strategy to deal with the situation.

A mechanistic model of the treatment process of the Granollers facility was developed using GPS-X, a commercial multipurpose modelling environment for the simulation of large-scale WWTP (Hydromantis, 1999). The biological reactor was modelled as four Continuous Stirred Tank Reactors, while primary and secondary settlers were modelled as one-dimensional tank with 10 layer of solids flux without biological reaction. The calibration procedure is still being carried out to adjust the kinetic, stoichiometric and settling parameters of the ASM1 model (Henze *et al.*, 1987) and the settler model (Tákacs *et al.*, 1991). The standard values for these parameters given by the GPS-X were used in the first simulations except for those more relevant, which were changed manually.

The model enables the HSS to simulate several off-line scenarios with different operational conditions, changes in the influent characteristics (underloading, overloading, storms, etc), and alternative actuation strategies proposed by the HSS. In spite of these capabilities, these kind of models show some limitations when dealing with problematic situations of biological origin (filamentous bulking, foaming, rising, etc.) as well as with situations for which it has not been calibrated. In this sense, the utilisation of soft-computing techniques to build a non-mechanistic model to simulate the behaviour of the plant at any situation is also being studied to be included as a new module into the HSS (Belanche *et al.*, 2000).



According to the predictive results, the head of the plant performs a final validation and decides which kind of control should be carried out. He/she (acting as an expert) can maintain, modify or deactivate the automatic control over the plant (i.e. a closed loop for controlling the dissolved oxygen level in the aeration tank), supported by the actuation strategy that the HSS suggested (expert, based on the reasoning procedure that the plant manager would do, or experiential, based on historical cases (real experiences) occurred in the plant).

This module also raises the interaction of the users with the computer system throughout an interactive and graphical user–machine interface. Moreover, at any moment and until a new supervisory cycle starts, the user can consult the conclusions of the HSS as well as any quantitative or qualitative variable to know the state and trend of the plant.

### Validation

The HSS is performing real-time support to process operation in the Granollers WWTP since September 1999. During the first four months the system was validated. The main objectives of the validation process were to guarantee the right performance of the HSS prototype, while checking for compliance with user requirement specifications. The methodology to validate the system was carried out through a two-stage validation procedure.

- Laboratory testing, which involves the execution of series of experiments to validate the correctness, consistency, and usability of each module. Face validation and historical cases techniques were combined to discover inaccuracy and inconsistency of the reasoning modules.
- Field testing of the overall HSS, which was faced up to real cases to detect module integration errors, and to assure the system could deal with real qualitative variables and missing information. The flow of information throughout the system was strictly followed to detect weak reasoning. When necessary, the Knowledge Base and the Case library were refined, adjusted, corrected and/or extended.

The ES module was tested in the laboratory with historical real cases. The standard reference was the expert criteria (i.e. the answers of the ES were compared to the solutions given by different WWTP experts). The reasoning paths were followed, compared, and discussed with the assistance provided by the expert of the plant. This comparison enabled us to discover which of the rules and procedures had been correctly fired, checking for unnecessary IF conditions, and redundant, conflicting, subsumed, circular, dead-end, unreachable and missing rules. As a result, during this period, new rules, procedures or facts were added, whilst others were modified or deleted (rule refinement, reformulation and revision).

On the other hand, the CBS module was validated according accuracy (set of acceptable responses) and adequacy to the domain knowledge covered. Expert experiences and historical cases with known solutions were used as standard reference. Several experiments were carried out to validate the hierarchical structure of the case library, the similarity measurement used, and the accuracy of using meta-libraries in the retrieval phase. If the CBS response differed from the expected output (i.e. the retrieved case suggested a diagnosis that was not similar enough to the standard reference diagnosis), details of the CBS module were revised (e.g. the relevant variables, their weight, the similarity criteria, the structure of the case library, etc.).

The second phase of the validation phase was carried out by means of field tests in the real facility. The main objective was to test the system within its real environment and to identify the need for further modifications. It was necessary to refine some modules of the HSS, still uncovering unexpected errors or dealing with new cases not provided by the system.

During this 4-month period of validation, the HSS was able to identify 123 different problem situations, including foaming, rising, filamentous bulking, underloading, overloading, deflocculation (including possible toxic), hydraulic shocks, mechanical faults, poor primary settler performance, non-biological origin problems on clarifier (sudden oscillations of up-flow velocity and bad performance of the clarifier due to an excess of biomass concentration), and influent nitrogen/organic matter shock. From those, 79.7% were successfully identified (about one third in advance and two thirds the same day), and 8.1% were wrong identifications (10 situations), while 12.2% were not identified (15 situations). Table 1 lists the number of correct situations detected in this period of time, specifying the problem and if they were detected in advance or if they were detected the same day.

As an example of the HSS behaviour, we can compare the real state of the process during the last week of September 1999 with the HSS diagnosis. On September 26th and 27th, the HSS identified a low Food to Micro-organism ratio (F/M). Low F/M is a cause of foaming due to filamentous organism proliferation. On September 28th the HSS alerted to the transition to foaming due to low F/M ratio. On the following days, the HSS was already pointing to a clear situation of foaming caused by low F/M, combined with an underloaded influent, which could worsen the biomass settling conditions. On the other hand, the real state of the Granollers WWTP during the two first days (26th and 27th) was registered by the head of the plant as Normal Operation. It was not until September 28th that low F/M values were detected. On September 29th the process presented the first signs of foams caused by the filamentous bacteria *Microthix Parvicella*, a situation that became more severe during the following days.

Both the low F/M values and the transition to foaming were identified by the supervisory system at least 48 hours in advance in comparison with the Granollers WWTP. Within this period of time, the plant manager could have acted over the process to avoid the consequences of the disturbance. In the case of the example, since the cause was well known, the specific control action of increasing F/M could have been carried out. It involves an

**Table 1** Situations identified successfully during the 4-month period of validation

| Situations | # detected in advance | # detected the same day |
|---|---|---|
| Filamentous bulking | 5 | 4 |
| Foaming | 7 | 10 |
| Rising | – | 2 |
| Underloading | 8 | 8 |
| Organic overloading | 3 | 1 |
| High influent organic concentration | – | 3 |
| High influent nitrogen concentration | – | 7 |
| Hydraulic overloading | 3 | 7 |
| Deflocculation problems | 4 | 4 |
| Primary settler problems | – | 5 |
| Mechanical or electrical problems | – | 15 |
| Non-biological problems on clarifier | – | 2 |
| TOTAL | 30 | 68 |

increasing of the waste activated sludge flow rate to decrease the biomass concentration and the sludge residence time. When the foaming episode is severe, the corrective actions to solve the problem can be extended with any non-specific method (chlorination, chemical reagent addition or manipulation of the recycle activated sludge).



## Conclusions

A hybrid knowledge-based supervisory system was developed and implemented to support the operation of a real WWTP. The system, which links classical control and numerical modelling to intelligent techniques, was structured into an agent-based architecture with three different levels: i) data gathering; ii) reasoning and diagnosis; iii) integration and decision support. In spite of certain reservations of the scientific community about the use of these techniques, the system has been successfully supervising the operation of the Granollers WWTP since September 1999. An 80% of correct identification of the process situation (one third of them in advance) during the first four-month validation process corroborates this conclusion. Validation of the proposed actuation strategies could not be evaluated because the HSS still does not act directly over the process. A deeper revision of the whole system has been carried out during the last year, in order to optimise the reasoning mechanisms of the tool, which must be robust enough to act automatically over the process. Concerning the transferability of the HSS to another facility, both technical and human bottlenecks must be seriously considered, in order to minimise time and cost effort. Among the long list of tasks to be scheduled, the key steps are: acquisition of specific knowledge of the facility, definition of the case and CBS calibration, adaptation of the ES Knowledge Base, communication of the HSS to SCADA system and peripherals, validation of the performance, and delivery to the owners of the facility. Main human bottlenecks include skill and motivation of the operators, trust of the head of the plant in the HSS, and expert paradigm.

## Acknowledgements

This research has been partially supported by the Spanish CICyT project TIC2000-1011. The authors also wish to acknowledge the co-operation of *Consorci per a la Defensa de la Conca del riu Besòs*, and especially of Àngel Freixó, its manager assistant.

## References

Allegro Common-Lisp (1998). Professional Edition 5.0, Franz Inc., Berkeley, USA.

Baeza, J., Ferreira, E.C. and Lafuente, J. (2000). Knowledge-based supervision and control of wastewater treatment plants: a real-time implementation. *Water Science & Technology*, **41**(12), 129–137.

Belanche, Ll., Valdés, J.J., Comas, J., R-Roda, I. and Poch, M. (2000). A Soft Computing Techniques Study in Wastewater Treatment Plants. *Artificial Intelligence in Engineering*, **14**(4), 307–317.

Bow-C. Chen (1998). Back-propagation Neural Network Adaptive Control of a Continuous Wastewater Treatment Process. *Industrial & Engineering Chemistry Research*, **37**(9), 3625.

Cortés, U., Sànchez-Marrè, M., Cecaronni, L., R.-Roda, I. and Poch, M. (2000). Artificial Intelligence and Environmental Decision Support Systems. *Applied Intelligence*, **13**(1), 77–91.

Gensym (2000). G2 Classic 5.1 Rev. 7, Gensym Corporation, Cambridge, USA.

Guariso, G. and Werthner, H. (1989). *Environmental Decision Support Systems*. Ellis Horwood-Wiley, New York.

Hydromantis (1999). GPS-X Version 3.0, Hydromantis Inc., Ontario, Canada.

Henze, M., Grady, Jr. C.P.L., Gujer, W., Marais, G.v.R. and Matsuo, T. (1987). *Activated Sludge Model No. 1. IAWPRC Scientific and Technical Reports No. 1*, IAWQ, London, UK.

Olsson, G., Aspegren, H. and Nielsen, M.K. (1998). Operation and Control of Wastewater Treatment – A Scandinavian Perspective over 20 years. *Water Science & Technology*, **37**(12), 1–13.

Rosen, C. and Olsson, G. (1998). Disturbance detection in wastewater treatment plants. *Water Science & Technology*, **37**(12), 197–205.

R.-Roda, I., Sànchez-Marrè, M., Cortés, U., Lafuente, J and Poch, M. (1999a). Consider a Case-Based System for Control of Complex Processes. *Chemical Engineering Progress*, **95**(6), 39–45.

R.-Roda, I., Sànchez-Marrè, M., Comas, J., Cortés, U., Lafuente, J. and Poch, M. (1999b). An Intelligent Integration to improve the Control and Supervision of large WWTP. *Proceedings of the 8th IAWQ Conference on Design, Operation and Economics of Large Wastewater Treatment Plants*, Budapest (Hungary), 566–570.

R.-Roda, I., Comas, J., Poch, M., Sànchez-Marrè, M. and Cortés, U. (2001). Automatic Knowledge Acquisition from Complex Processes for the development of Knowledge-Based Systems. *Industrial & Engineering Chemistry Research*, **40**, 3353–3360.

Sànchez-Marrè, M., Cortés, U., R.-Roda, I., Poch, M. and Lafuente, J. (1997). Learning and adaptation in Wastewater Treatment Plants through case-based reasoning. *Microcomputers in Civil Engineering* (Special issue: Machine Learning), **12**(4), 251–266.

Stephanopoulos, G. and Han, C. (1996). Intelligent Systems in process Engineering: a review. *Computers & Chemical Engineering*, **20**(6/7), 743–791.

Takács, I., Patry, G.G. and Nolasco, D. (1991). A dynamic model of the clarification-thickening process. *Water Research*, **25**(10), 1263–1271.

Wang, X.Z., Chen, B.H., Yang, S.H., Mc Greavy, C. and Lu, M.L. (1997). Fuzzy Rule Generation from Data for Process Operational Decision Support. *Computers & Chemical Engineering*, **21**(1001), S661–S666.

Zhu, X.X. and Simpson, A.R. (1996). Expert system for water treatment plant operation. *Journal of Environmental Engineering*, 822–829.