ELSEVIER

Designing and building real environmental decision support systems

Manel Poch [a,*], Joaquim Comas [a], Ignasi Rodríguez-Roda [a], Miquel Sànchez-Marrè [b], Ulises Cortés [b]

[a] *Laboratori d'Enginyeria Química i Ambiental (LEQUIA), Universitat de Girona, Campus de Montilivi s/n, Girona, Catalonia 17071Spain*
[b] *Knowledge Engineering and Machine Learning Group, Universitat Politècnica de Catalunya, Campus Nord, Edifici C5, Jordi Girona 1-3, 08034 Barcelona, Spain*

Abstract

The complexity of environmental problems makes necessary the development and application of new tools capable of processing not only numerical aspects, but also experience from experts and wide public participation, which are all needed in decision-making processes. Environmental decision support systems (EDSSs) are among the most promising approaches to confront this complexity. The fact that different tools (artificial intelligence techniques, statistical/numerical methods, geographical information systems, and environmental ontologies) can be integrated under different architectures confers EDSSs the ability to confront complex problems, and the capability to support learning and decision-making processes. In this paper, we present our experience, obtained over the last 10 years, in designing and building two real EDSSs, one for wastewater plant supervision, and one for the selection of wastewater treatment systems for communities with less than 2000 inhabitants. The flow diagram followed to build the EDSS is presented for each of the systems, together with a discussion of the tasks involved in each step (problem analysis, data collection and knowledge acquisition, model selection, model implementation, and EDSS validation). In addition, the architecture used is presented, showing how the five levels on which it is based (data gathering, diagnosis, decision support, plans, and actions) have been implemented. Finally, we present our opinion on the research issues that need to be addressed in order to improve the ability of EDSSs to cope with complexity in environmental problems (integration of data and knowledge, improvement of knowledge acquisition methods, new protocols to share and reuse knowledge, development of benchmarks, involvement of end-users), thus increasing our understanding of the environment and contributing to the sustainable development of society.
© 2003 Elsevier Ltd. All rights reserved.

Keywords: Environmental decision support systems; Artificial intelligence; Wastewater treatment; Knowledge management

1. Introduction

1.1. Statement of the problem

The increasing rhythm of industrialisation, urbanisation and population growth that our planet has faced for the last few hundred years has forced society to consider whether human beings are changing the very conditions that are essential to life on Earth. Environmental pollution affects the quality of water, air, and soil negatively, and hence plant, animal and human life (Sydow et al., 1998; El-Swaify and Yakowitz, 1998).

Whenever we attempt to tackle these issues, we are immediately confronted with complexity. There are at least two important reasons for this:

– *Uncertainty, or approximate knowledge*. Some of the sources of this uncertainty can be tamed with additional data or further investigation. Such is the case of uncertainty arising from random processes or from deficiencies in knowledge (lack of data, unsuitable datasets, etc.). But in other cases, uncertainty is insurmountable. This is the case for chaotic behaviour, or for self-organisation processes. It is also typical of

* Corresponding author. Tel.: +34-972-418804; fax: +34-972-418150.
E-mail addresses: manel@lequia.udg.es (M. Poch); quim@lequia.udg.es (J. Comas); ignasi@lequia.udg.es (I. Rodríguez-Roda); miquel@lsi.upc.es (M. Sànchez-Marrè); ia@lsi.upc.es (U. Cortés).

socio-ecological systems, which involve numerous players, each with their own goals.

– *Multiplicity of scales.* Environmental problems have been associated traditionally with distinct spatial scales (i.e. local, national, global), each associated with specific timescales. However, interactions among these scales are becoming increasingly clear. Therefore, advocating a single perspective that encompasses everything in a system is becoming increasingly difficult—plus ineffective.

The consensus is developing that, in order to account for these caveats, environmental issues must be considered in terms of complex systems. But not all environmental systems present the same level of complexity in terms of both the degree of uncertainty and the risk associated with decisions. If the degree of complexity is represented as a function of uncertainty, on one hand, and the magnitude or importance of the decision, on the other, then we might distinguish three levels of complexity (Funtowicz and Ravetz, 1993, 1999):

– The first level of complexity would correspond to simple, low uncertainty systems where the issue at hand has limited scope. A single perspective and simple models would suffice to provide satisfactory descriptions of the system. With regard to water issues, this level corresponds, for example, to the evolution of oxygen in a pristine stream after a pulse input of assimilable organic matter. In the context of industrial processes, an example is the design of a single treatment operation where the input is perfectly defined. In these cases, the information arising from analysis may be used for more wide-reaching purposes beyond the scope of the particular researcher.
– The second level would correspond to systems with enough uncertainty that simple models, applicable to different situations and manageable by any competent practitioner, can no longer provide satisfactory descriptions. Acquired experience becomes then more and more important, and the need to involve experts in problem solving becomes advisable. In the case of water issues, this level would correspond to a general model of water quality, where the need arises to establish which factors are the most important. In the case of an industrial process, this level would correspond to the installation of a wastewater treatment plant (WWTP), where goals for the quality of the output are well established but these can be reached through different schemes, and it is the responsibility of the designer to choose the most appropriate configuration.
– The third level would correspond to truly complex systems, where much epistemological or ethical uncertainty exists, where uncertainty is not necessarily associated with a higher number of elements or relationships within the system, and where the issues

at stake reflect conflicting goals. It is then crucial to consider the need to account for a plurality of views or perspectives. In the case of water issues, an example would be the problem of water quality in a stream catchment. Here, a variety of factors (economical, technical, ecological, etc.) are at play, and associated with each factor is a different set of goals. Thus, different kinds of expertise need to be taken into account. In the case of an industrial process, this level of complexity is associated, for instance, with the environmental aspects of wastewater treatments, which are discussed at the level of the company's policy. Thus, the problem is not the design of end of pipe installations for the treatment of specific outputs, but a more global view on the problem that would contemplate, for example, the installation of cleaner technologies in the production process itself.

In this sense, it is important to realise that environmental problems are characterised by dynamics and interactions that do not allow for an easy division between social and biogeophysical phenomena. Many ecological theories have been developed in systems where humans were absent, or in systems where humans were considered an exogenous, simple, and detrimental disturbance. The intricate ways in which humans interact with ecological systems have been rarely considered (Kinzig, 2001). Embracing a socio-economical perspective implies accepting that all decisions related to environmental management are characterised by multiple, usually conflicting objectives, and by multiple criteria (Ostrom, 1991). Thus, in addition to the role of experts, it becomes increasingly important to consider the role of wide public participation in the decision-making processes. Experts are consulted by policy makers, the media, and the public at large to explain and advise on numerous issues. Nonetheless, many recent cases have shown, rather paradoxically, that while expertise is increasingly sought after, it is also increasingly contested (Ludwig, 2001).

In our opinion, this third level cannot be tackled with the traditional tools of mathematical modelling. To confront this complexity, a new paradigm is needed. Adopting it will require that we deal with new intellectual challenges.

1.2. New tools for a new paradigm

In the last few decades, mathematical/statistical models, numerical algorithms and computer simulations have been used as the appropriate means to gain insight into environmental management problems and provide useful information to decision-makers. To this end, a wide set of scientific techniques have been applied to environmental management problems for a long time and with good results.

But most of these efforts were focused on problems that we could assign to the first level of complexity. Consequently, many complex environmental problems have not been effectively addressed by the scientific community. Recently, however, the effort to integrate new tools to deal with more complex systems has led to the development of the so-called environmental decision support systems (EDSSs) (Guariso and Werthner, 1989; Rizzoli and Young, 1997).

EDSSs have generated high expectations as a tool to tackle problems belonging to the second and third levels of complexity. Thus, in a recent review of the relevant literature in the topic, more than 600 references were found (including journal articles, conference papers, and technical reports) during the 90s, with only 10 references in 1992 and more than 150 references per year towards the end of the decade (Cortés et al., 2002). The range of environmental problems to which EDSSs have been applied is wide and varied, with water management at the top (25% of references), followed by aspects of risk assessment (11.5%) and forest management (11.0%). Equally varied are the tasks to which EDSSs have been applied, ranging from monitoring and data storage to prediction, decision analysis, control planning, remediation, management, and communication with society. It is not surprising then that three of the top 10 most downloaded articles published in *Environmental Modelling and Software* in January–December 2001 deal with EDSSs.

This review, together with the work of other authors, also revealed that there is a wide range of opinions on what constitutes an EDSS. The fact that this approach is relatively recent and integrates multiple tools means that there is not a single, consensual definition of EDSS. However, even though one may argue that a database management system could be used as a decision support system, today's consensus is that EDSSs must adopt a knowledge-based approach, which includes the steps of knowledge acquisition, representation, and management.

The fact that different tools can be integrated under different architectures makes EDSSs difficult to define. It also means that different approaches to design and implementation coexist.

In this context, we present our experience with the design and implementation of two EDSSs in the domain of water management. We explicitly describe their development and the architecture used for the applications.

1.3. EDSS development

According to Fox and Das (2000), a decision support system is a computer system that assists decision-makers in choosing between alternative beliefs or actions by applying knowledge about the decision domain to arrive at recommendations for the various options. It incorporates an explicit decision procedure based on a set of theoretical principles that justify the "rationality" of this procedure.

Thus, an EDSS is an intelligent information system that reduces the time in which decisions are made in an environmental domain, and improves the consistency and quality of those decisions (Haagsma and Johanns, 1994; Cortés et al., 2001). Decisions are made when a deviation from an expected, desired state of a system is observed or predicted. This implies a problem awareness that in turn must be based on information, experience and knowledge about the process. Those systems are built by integrating several artificial intelligence methods, geographical information system components, mathematical or statistical techniques, and environmental ontologies (Fig. 1).

How a particular EDSS is constructed will vary depending on the type of environmental problem and the type of information and knowledge that can be acquired. With these constraints in mind, and after an analysis of the available information, a set of tools can be selected. This applies not only to numerical models, but also to artificial intelligence (AI) methodologies, such as knowledge management tools. The use of AI tools and models provides direct access to expertise, and their flexibility makes them capable of supporting learning and decision-making processes. Their integration with numerical and/or statistical models in a single system provides higher accuracy, reliability and utility (Cortés et al., 2000).

This confers EDSSs the ability to confront complex problems, in which the experience of experts provides valuable help for finding a solution to the problem. It also provides ways to accelerate identification of the problem and to focus the attention of decision-makers on its evaluation. Once implemented, an EDSS, like any knowledge-based system, has to be evaluated for what it knows, for how it uses what it knows, for how fast it
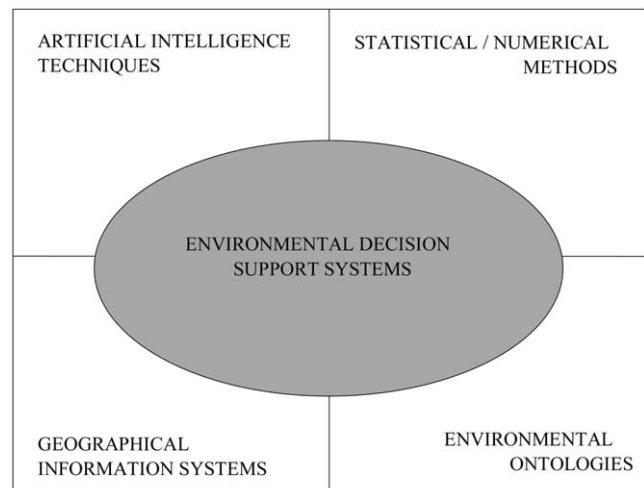
Fig. 1. EDSS conceptual components.

can learn something new, and, last but not least, for its overall performance. Fig. 2 shows schematically the methodology used to develop the two study cases presented here.

Both the proposed EDSS development procedure and the EDSS architecture are general enough to be intended to cope with any kind of EDSS deployment.

We propose an EDSS architecture based on five levels (Fig. 3):

- The first level of the EDSS (*data gathering*) encompasses the tasks involved in data gathering and registration into databases. Original raw data are often defective, requiring a number of pre-processing procedures before they can be registered in an understandable and interpretable way. Missing data and uncertainty must be also considered in this level.
- The second level, *diagnosis level*, includes the reasoning models that are used to infer the state of the process so that a reasonable proposal of actuation can be reached. This is accomplished with the help of statistical, numerical and artificial intelligence models.
- The third level, *decision support level*, establishes a

supervisory task that entails gathering and merging the conclusions derived from knowledge-based and numerical techniques. This level also raises the interaction of the users with the computer system through an interactive and graphical user–machine interface. When a clear and single conclusion cannot be reached, a set of decisions ordered by their probability should be presented to the user.

- In the fourth level, plans are formulated and presented to managers as a list of general actions suggested to solve a specific problem.
- The set of actions to be performed to solve problems in the domain considered are in the fifth level. The system recommends not only the action, or a sequence of actions (a plan), but also a value that has to be accepted by the decision-maker. This is the last level in the architecture that closes the loop.

2. Two case studies

In this section, two case studies are presented where the proposed methodology has been applied. The two
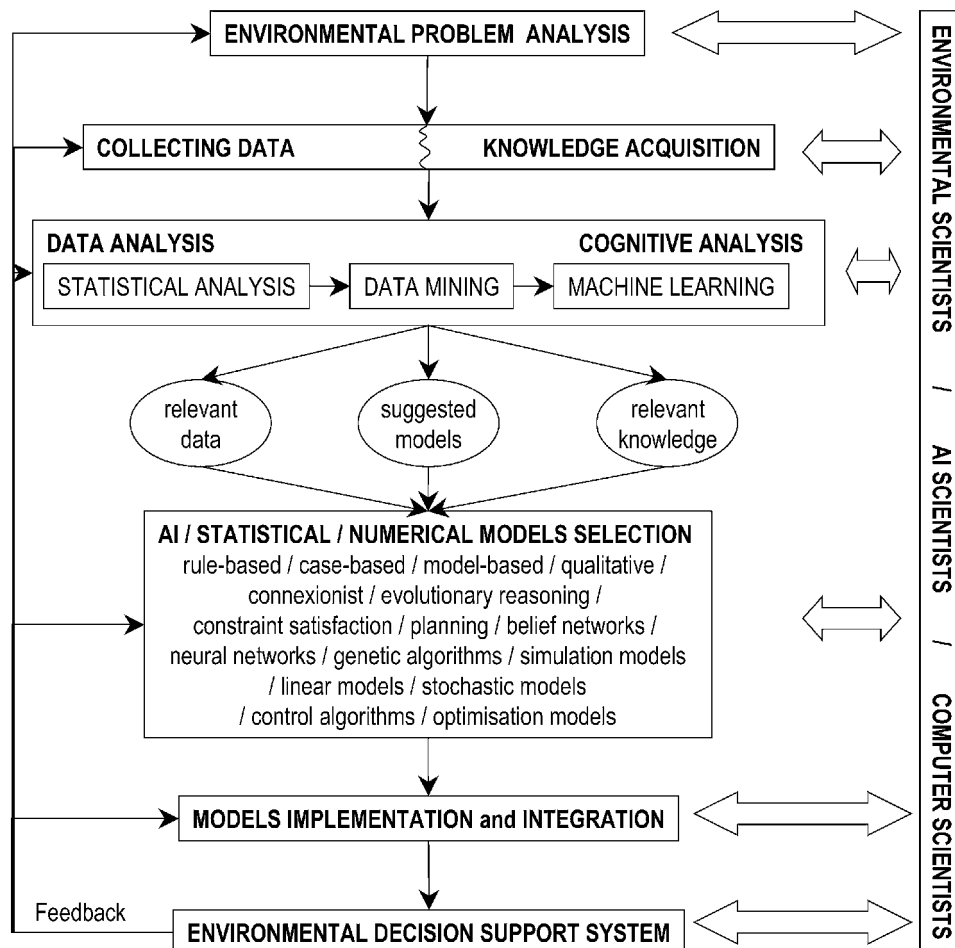

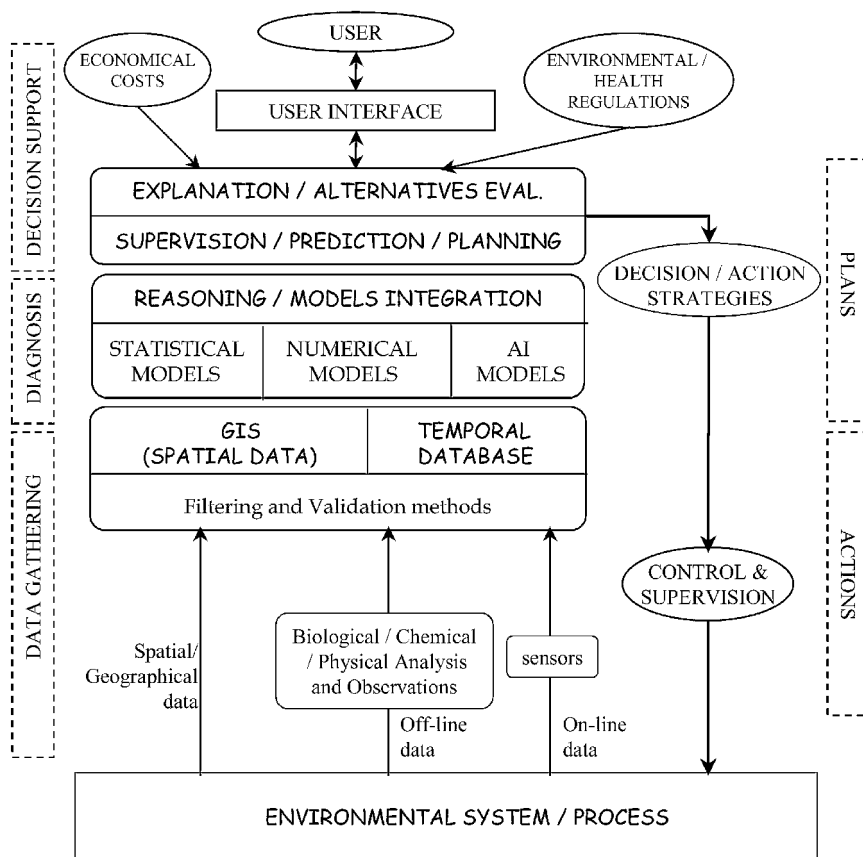Fig. 2.   Flow diagram for development of an EDSS.

Fig. 3.   EDSS architecture.

case studies correspond to two different situations with different forms of complexity. Although specific details of the corresponding databases cannot be published (proprietary information of Consorci per a la Defensa de la Conca del riu Besòs and Agència Catalana de l'Aigua), the available information included in the examples is clear enough to follow the process of designing and building the EDSS.

The first case corresponds to the application of an EDSS to the supervision of a WWTP. Here, both quantitative information (obtained on-line and off-line) and qualitative information are used, with the important participation of experts. While discrepancies among experts may arise, there are no conflicts of interest. Of the four conceptual components of the EDSS as stated above, the geographic component is not relevant in this case, while the numeric component is the only one traditionally used for tackling the problem. In the scheme of complexity and risk, it lies between the second and third levels.

The second case corresponds to the selection of wastewater treatment and disposal systems in Catalonia. It is a planning problem in which the temporal component has little relevance, since on-line responses are not needed. The importance of numeric methods is lower than in the first case, while the importance of the GIS and expert experience components increases. Conflicts

of interest among experts may arise, and the interactions between social and biogeophysical phenomena become relevant. In the scheme of complexity, it would lie in the third level.

2.1. Wastewater treatment plant supervision

2.1.1. EDSS building
2.1.1.1. Problem analysis   A typical WWTP usually includes a physical and/or chemical primary treatment and a biological secondary treatment to remove organic matter and suspended solids from wastewater. Primary treatment is designed to physically remove solid material from the incoming wastewater. The wastewater flowing to the secondary treatment is called the primary effluent. Secondary treatment usually consists of a biological conversion of dissolved and colloidal organic compounds into stabilised, low-energy compounds and new biomass cells, caused by a very diversified group of microorganisms, in the presence of oxygen. The mixture of microorganisms and particles has the ability to settle and separate from treated water in the clarifier. A biological reactor followed by a secondary settler or clarifier constitutes the activated sludge process, which is the most well-known process of secondary treatment because it is also the most widely used.

Like other environmental and biotechnological processes, WWTPs are complex systems involving many interactions between physical, chemical and biological processes, e.g. chemical and biological reactions, kinetics, catalysis, transport phenomena, separations, etc. The successful management of these systems requires multidisciplinary approaches and expertise from different scientific fields. Some of the special and problematic features of these processes are:

- Intrinsic instability: most of the chemical and physical properties as well as the size and species diversity of the population of microorganisms involved in environmental processes do not remain constant over time.
- Many of the facts and principles underlying the domain cannot be characterised precisely solely in terms of a mathematical theory or a deterministic model with clearly understood properties.
- Uncertainty and imprecision of data or approximate knowledge and vagueness: these processes generate a considerable amount of qualitative information.
- Huge quantity of data/information: the application of current computer technology to the control and supervision of these environmental systems has led to a significant increase in the amount of data acquired.
- Heterogeneity and scale: because the media in which environmental processes take place are not homogeneous and cannot easily be characterised by measurable parameters, data are often heterogeneous. Also, the different scale times inherent to different measures in the process have to be properly integrated and managed.

Due to the complexity of wastewater treatment process control, even the most advanced conventional *hard* control systems have encountered limitations when dealing with problem situations that require qualitative information and heuristic reasoning for their resolution (Olsson et al., 1998). Indeed, to describe these qualitative phenomena or to evaluate circumstances that might call for a change in the control action, some kind of linguistic representation built on the concepts and methods of human reasoning, such as intelligent systems, has been necessary. This is also the reason why human operators have, until now, constituted the final step in closed-loop plant control. A deeper approach is necessary to overcome the limited capabilities of conventional automatic control techniques when dealing with *abnormal* situations in complex systems, and to provide the level and quality of control necessary to consistently meet environmental specifications.

For these reasons, the use of EDSSs began to look promising in terms of solutions to these problems. A reasonable, distributed proposal outlines the scope for the integration of AI tools such as pattern recognition, knowledge-based systems, fuzzy logic, artificial neural networks, case-based systems, or inductive decision trees, which handle the particular characteristics of complex processes with numerical and conventional computational techniques, such as statistical methods, advanced and robust control algorithms and system identification techniques.

The WWTP selected to develop and apply our proposed Supervisory System prototype is located in Granollers, in the Besòs river basin (Catalonia). Nowadays, this facility provides preliminary, primary and secondary treatment to remove the organic matter, suspended solids and, under some conditions, nitrogen contained in the raw water of about 130,000 inhabitant-equivalents. The Granollers WWTP has several particular characteristics that increase the potential advantages of the development and application of an intelligent supervisory system to control and supervise the wastewater treatment process. Among these characteristics, we would like to emphasise the following:

- Availability of a significant amount of historical records describing plant operation.
- This plant has a high level of automation centralised in a computer that collects on-line data and controls most of the plant operations.
- The Granollers WWTP is a highly variable system. A wide range of different situations take place throughout the year causing significant changes in the influent characteristics (storms, overloading, nitrification in hot periods, uncontrolled industrial spills etc.), which affect standard process operation.
- High level of specialisation of plant experts who have been working in the plant from the beginning of its operation. They are perfectly acquainted with all sorts of details that make up the heuristic knowledge of the plant.

2.1.1.2. Data collection and knowledge acquisition A variety of methods were used for the development of a knowledge base for this study. Conventional knowledge acquisition methods (literature review, interviews, etc.) were used first. To overcome the limitations of conventional methods, these were supplemented with the use of different automatic knowledge acquisition methods. These latter methods can be supervised, mainly inductive learning techniques such as CN2, C4.5 and k-NN or unsupervised, such as some clustering method (R.-Roda et al., 2001). Fig. 4 illustrates the main sources and methods that were used to acquire both general and specific knowledge on the wastewater treatment processes.

2.1.1.3. Model selection Two types of models were selected: rule-based reasoning models (expert system) and case-based reasoning models.

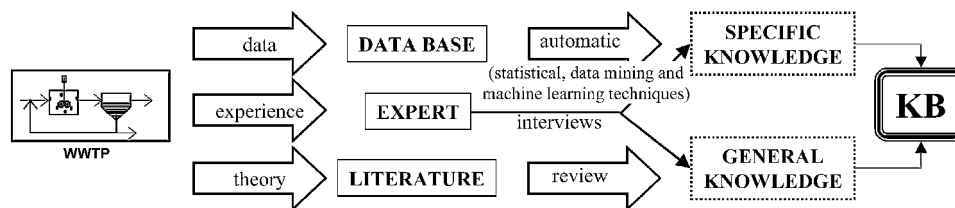Rule-based systems (RBS) offer a number of advan-

Fig. 4.   Methods used to acquire knowledge.

tages that overcome some of the limitations of other techniques: they facilitate the inclusion and retention of heuristic knowledge from experts and allow the processing of qualitative information; knowledge is represented in an easily understandable form (rules); a well-validated expert system offers potentially optimal answers because action plans are systematised for each problematic situation; and, finally, expert systems make the acquisition of a large general knowledge base possible with somewhat rigid use for any WWTP.

A Case-Based Reasoning System (CBS) is based on the idea that solving a problem for the second time is usually easier than solving it for the first time because we remember and repeat the previous solution or recall our mistakes and try to avoid them. The basic idea is to adapt solutions applied in the past to particular problems affecting process performance and apply them to new problems that are similar in nature with less effort than with other methods that start from scratch. A case is described as a conceptualised piece of knowledge representing an experience that teaches a fundamental lesson on how to achieve the reasoner's goals. In the wastewater treatment domain, the case is a codified description of a specific state or experience of the studied WWTP. Codification should be in computer storable form in order to be easily retrieved in the future. This paradigm supplies a flexible and dynamic model allowing the EDSS to be adapted and used for any WWTP with a similar technology.

2.1.1.4. Model implementation   Among the different possibilities (tables, decision trees, or knowledge diagrams and frames) for the representation of the elicited knowledge, decision trees were selected as the most suitable representation (Sànchez-Marrè et al., 1996; Comas et al., 2003). All the symptoms, facts, procedures and relationships used for problem diagnosis can be cast into a set of decision trees. The translation of the knowledge contained in a branch of decision trees into a production rule is direct. The resulting trees, which avoid contradictions and redundancies, comprise, in our case, diagnosis, cause identification, and action strategies for a wide range of problems in WWTP operation. Logic trees serve as a record of the expert's step-by-step information processing and decision-making activity. Some branches are specific and contain peculiarities of the plant, while others are more general and can be applied to any plant.

The set of specific cases is stored in a structured memory in a case-base (the case-library) and initialised with a set of typical cases in the plant. The CBS development includes the case definition, the case-library structure definition, and the selection of the initial seed. CBSs require a case-library to broadly cover the set of potential problems. These cases are indexed in memory so as to be retrieved whenever the experiences they encapsulate can contribute to achieving the goals of the process. Both successes and failures must be included in the case-library. It is advisable to initialise the library with a set of common situations (or cases) obtained from real data or provided by experts on the process. Thus, the CBS will be from the very start ready to propose solutions to problems that are similar to those considered in the initial "seed". The initial seed at Granollers included 74 real cases from the historical database, which covered a broad range of situations covering the main problems in the process as well as normal situations. The library is updated with new cases as the knowledge about the process progresses; so the CBS evolves into a better reasoner and system accuracy benefits from these new acquisitions. However, because large amounts of information can overcrowd the library, only the most relevant cases are included.

Fig. 5 shows our intelligent approach to improve the supervision of WWTPs. The integration of different AI technologies (RBS, CBS and Artificial Neural Networks) with numerical methods (classical control systems or models) leads to a *hybrid* knowledge-based system capable of overcoming the limitations found when solving complex problems with a sole classical or AI technique. Thus, this multidisciplinary approach appears to be the most optimal solution to guarantee the successful control of complex processes like WWTP.

2.1.1.5. Validation of the EDSS   Field testing was considered to be the most effective validity test. The main objective of field validation was to test the use of the overall EDSS in situ with real cases. We wanted to test the system in its real environment and identify needs for further modifications. The system performance was tested for more than 10 months in its actual operating environment, where it is working as a real-time decision support system. During the period of exhaustive validation, the EDSS coped with 123 different problem situ-
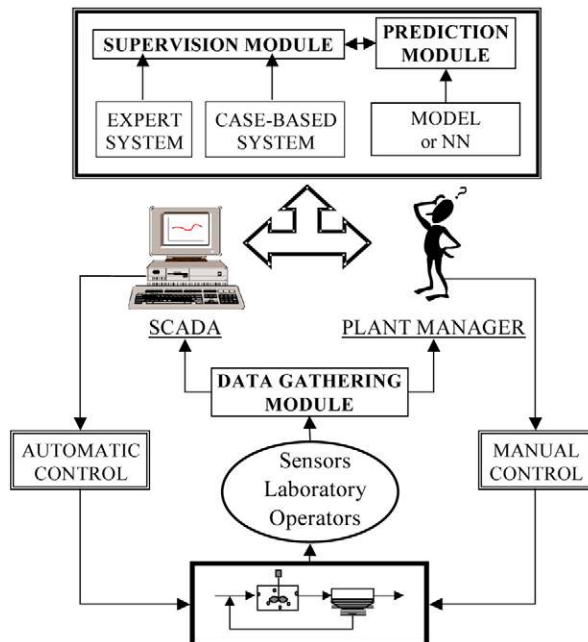
ENVIRONMENTAL DECISION SUPPORT SYSTEM

Fig. 5.   Our hybrid intelligent approach to supervise WWTP.

Table 1
Situations successfully identified during the period of validation

Situations	Number of situations detected in advance	Number of situations 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

ations and suggested suitable action strategies, in most of the cases.

The WWTP problem situations detected included foaming, rising, filamentous bulking, underloading, overloading, deflocculation (including possible toxic shock), hydraulic shock, mechanical fault, poor primary settler performance, non-biological origin problem on clarifier (sudden oscillation 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 (98 situations, about one-third in advance and two-third the same day), and 8.1% were wrong identified (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 whether they were detected in advance or they were detected the same day. Nowadays, the EDSS is used as a complementary tool of diagnosis for the everyday management of the activated sludge process (Rodríguez-Roda et al., 2002).

2.1.2. EDSS operation

The different tasks of the five levels of the EDSS for WWTP control and supervision are performed cyclically, using a supervisory cycle. Fig. 6 outlines the EDSS supervisory cycle.

Each cycle is composed of five steps and several tasks: data gathering (with the data acquisition and updating tasks), diagnosis, decision support (with the user-validation and action tasks) and plans and actions (with the supervision, prediction and evaluation tasks).

The operation of the EDSS supervisory cycle can be summarised as follows.

2.1.2.1. Data gathering   Every time the supervisory cycle is launched, the main task to be performed is data gathering and updating current data for the inference process. Data gathering is accomplished through on-line data acquisition systems (sensors and equipment) and off-line data acquisition systems (biological, chemical and physical water and sludge analyses and other qualitative observations of the process). Moreover, this level of operation implements data filtering, validation and management processes on the temporally evolving (real-time) database where on-line data, off-line data and data calculated by the system are stored.

According to the manager of the plant, there is a minimum set of variables—the basic information—that must be updated in order to make a reliable diagnosis of the current state of the process. In the Granollers WWTP, these are the influent flow rate and the chemical oxygen demand of the biological influent. However, the diagnosis mechanisms cannot rely on these two values, and the same conclusion can be inferred from different variables, avoiding the interruption of the process because of the lack of information.

2.1.2.2. Diagnosis   Once the data have been collected, they are sent to the *diagnosis* module where the knowledge-based systems (ES and CBS) are executed concurrently without any kind of interaction between them. Thus, the current state of the process will be diagnosed through a reasoning task based on both the expert rules and the most similar cases retrieved. If a problem is detected or suspected, the diagnosis module will also try to identify the specific cause. The solution to the most
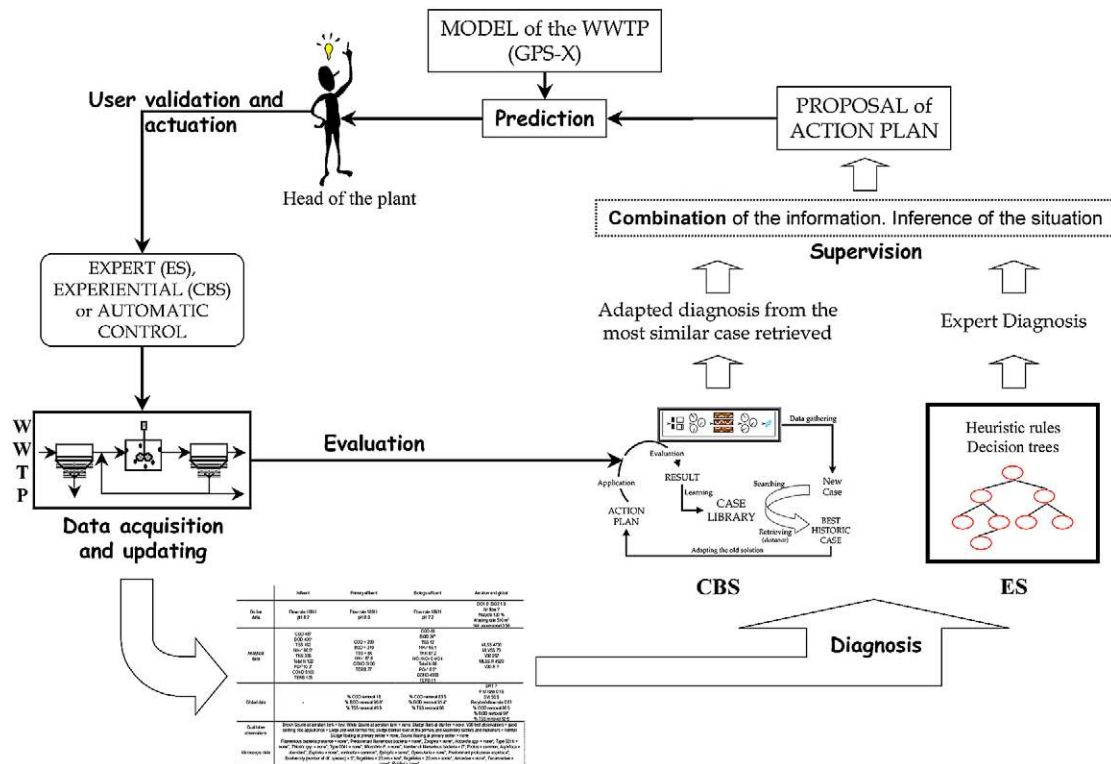
Fig. 6. Supervisory cycle.

similar case is modified so as to adapt it to the new situation.

2.1.2.3. Decision support The conclusions reached in the diagnosis phase are sent to the decision support module. This upper module infers the global situation of the WWTP. The final result is sent through the computer interface to the operator, who will finally decide on the action to be taken (user-validation and action) (Fig. 7). The expert can use a dynamic mechanistic or a black-box model implemented to support the selection process of an action plan by simulating the possible consequences of applying different alternatives (Belanche et al., 1999, 2000).

2.1.2.4. Plans and actions The EDSS suggests an action plan resulting from the supervision and prediction tasks, and integrating the expert recommendations sent by the RBS and the experience retrieved by the CBS. In case of conflict, the user acts as the expert of the process, evaluating the suggestions of both systems, checking their validity and deciding which is the best strategy to deal with the situation. The evaluation of the results of the application of the action plan to solve the problem allows the system to close the CBS working cycle (see Fig. 8), to learn from successful and failed past experiences, and to update the case-library.

These features can be detected by the EDSS itself (unless a manual operation is carried out), but it is essen-

tial that confirmation be provided by the plant manager, who will have the opportunity to change misleading information or add missing information. In addition, the EDSS can extend the knowledge bases by acquiring new knowledge from new sources.

2.2. Selection of wastewater treatment and disposal systems for communities with less than 2000 inhabitants

2.2.1. EDSS building
2.2.1.1. Problem analysis WWTPs, and especially biological plants, based on several variants of the activated sludge process (suspended growth biomass), are currently the predominant system for urban wastewater treatment and disposal in Catalonia. In accordance with the goals established in the First Urban Wastewater Treatment Programme of Catalonia (PSARU I), WWTPs have been built to serve every town in Catalonia with more than 2000 inhabitants. In communities with less than 2000 inhabitants, however, the situation is different. Few of them have wastewater treatment systems in place today (2001), but all should have them by 2005. The number of communities lower than 2000 inhabitant-equivalents (i.e.) involved in the Small Communities Wastewater Treatment Plan of Catalonia (PSARU II) is about 3500 agglomerations, affecting to more than 800 municipalities. It means planning the wastewater treatment for 200,000 inhabitants (currently censed),
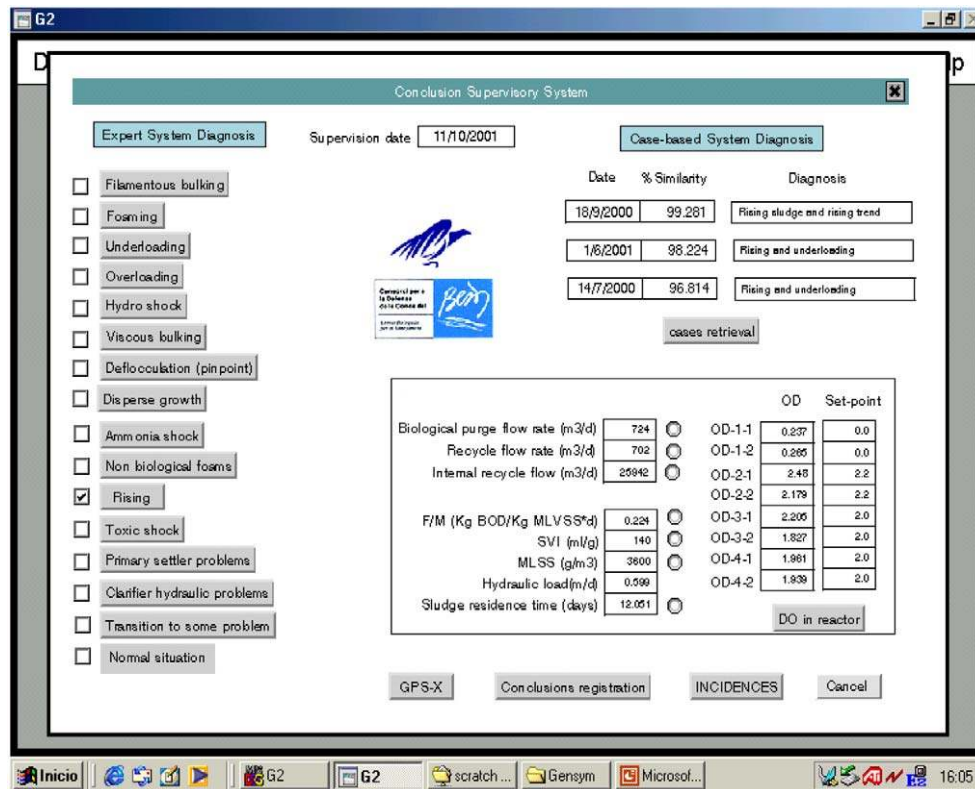
Fig. 7.   User-interface showing a summary of the EDSS for WWTP supervision.
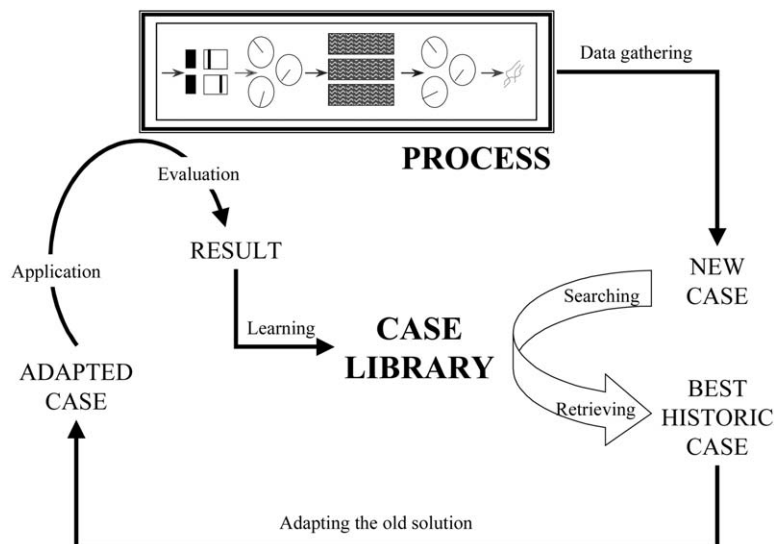

Fig. 8.   Case-Based Reasoning System working cycle.

approximately 5% of the Catalan population. The rest of the 95% were already considered within the plan already developed to treat wastewater from populations higher than 2000 i.e. The distribution of population lower than 2000 i.e. shows that the main part of this population lives in rural communities (73.8% of the agglomerations has lower than 200 i.e.).

Other important remarks of the small communities in Catalonia concern the relevant contribution of the percentage of seasonal population (45.1%) with respect to the permanent population and the proportion of industrial wastewater in some villages (globally, 27.2%), especially when it cannot be assimilable to urban wastewater.

While the European Water Directive 91/271/EEC specifies the type of treatment to implement in towns

with more than 2000 inhabitant-equivalents, for smaller communities this directive only states that the type of treatment should be appropriate. This significantly changes the decision process of selecting the optimal treatment. In this context, 'appropriate treatment' is defined as one that fulfils the quality standards set for the receiving waters. This suggests new dimensions of analysis where the landscape and the affected environment together with the characteristics of the wastewater treatment technologies for small communities are to be taken into account. Thus, in order to make sound recommendations based on the available technologies and the characteristics of the receiving environment and the landscape, it becomes necessary to acquire and integrate expertise from diverse disciplines. This change of perspective on the problem with respect to PSARU I suggests that we should move towards a paradigm that allows dealing with complexity.

In view of this complexity, three dimensions of analysis must be taken into account during the decision-making process:

1. The characteristics of the small community itself. This is an aspect of evident importance given the large number of rural communities in Spain and the variety of climatic, geomorphologic and population dynamics conditions that should be taken into consideration when selecting the best option. It is also important to consider that, unlike larger communities and towns, rural communities directly experience the implementation of the sewage treatment system, with respect to both perceived benefits and perceived impacts on their environment.
2. The receiving environment which should improve significantly once the Small Communities Wastewater Treatment Plan of Catalonia is implemented. Protection of the receiving environment is of the highest importance, as endorsed by the recent 2000/60/EC Directive. In order to improve on the current state, an assessment of the current ecological quality of the site is needed. Significantly, many of the sites where treatment systems are to be implemented are in protected areas or in rural areas with high actual or potential ecological quality that deserves to be preserved or restored.
3. The wastewater treatment systems appropriate for small communities. These differ broadly in terms of both technology and operation, and need to be accommodated to each particular situation. Thus all the advantages, disadvantages, and any factors that might affect the final decision must be taken into consideration.

For each particular case, the integration of these three types of information—the rural community, the receiving environment, and the type of treatment—will suggest optimal and multidisciplinary scenarios to support the decision-making process, since they will have taken into account not just technical aspects of treatment optimisation, but also environmental, economical and social factors. Reflecting the will to face the problem in all its complexity, the Catalan Water Agency (ACA, "Agència Catalana de l'Aigua") decided to design an EDSS. A consortium formed by four universities (Universitat de Girona, Universitat Politècnica de Catalunya, Universitat de Barcelona and Universitat Autònoma de Barcelona) and the Spanish Scientific Council (CSIC) was commissioned to develop a system that would attempt to reproduce the reasoning process followed by a group of experts facing the highly complex situation at stake. The goal of embracing complexity implied that we should not limit ourselves to 'formal' knowledge, but should attempt to incorporate 'non-formal' knowledge. The latter derives both from the 'subjective' reasoning processes of experts in different disciplines and from the knowledge accumulated by persons or social groups sensitive to the problems and involved in finding solutions to them. This allowed us to recognise the multiple views and interests that are involved in the decision-making process: financial cost, social and environmental benefits, technical criteria, and so on.

An EDSS was chosen as the most suitable tool to support the identification of the appropriate wastewater treatment for small communities because it integrates expert knowledge and encourages a multidisciplinary approach—with respect to the affected land and environment—that incorporates knowledge from affected persons and social groups. It is then possible to reach a consensus among disparate views that approaches an optimal solution. Furthermore, since an EDSS is a computer system, it not only allows the management and analysis of large volumes of numerical data but also that of symbolic and, sometimes, uncertain and inexact data.

2.2.1.2. Collecting data and knowledge acquisition

Three different sources of knowledge were pooled together to build a knowledge base as comprehensive as possible (Fig. 9). These knowledge sources were:

- Interviews with experts in water management and wastewater treatment, as well as with experts in the quality of the receiving environment.
- Reviews from scientific and technical literature as well as knowledge drawn from visits to relevant regions where this type of wastewater treatment programme has already been implemented.
- Analysis of the available historical data for the receiving environment as well as from data on the small communities themselves.

The first source of knowledge we turned to was a group of experts in the wastewater treatment process.
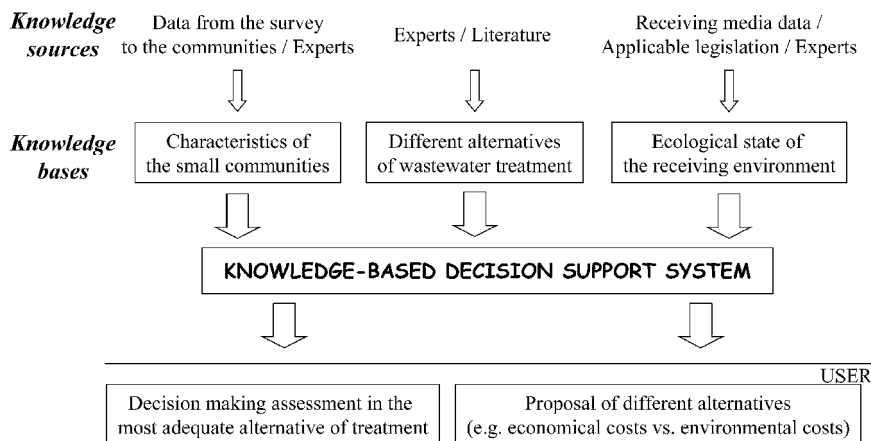
Fig. 9.   Knowledge sources and knowledge bases involved in the knowledge acquisition phase.

The knowledge we were seeking was extracted from a series of interviews or conversations. Specifically, we set up a series of interviews with experts in wastewater management and environmental experts from the administration, research centres, and engineering consulting firms. From this series of interviews, we gathered heuristic knowledge specific to Catalonia. This knowledge, accrued during years of work and experience in the same field, is essential for the successful development of the EDSS. This was supplemented with knowledge derived from specialised journals and books, and from notes taken during the field visits. Finally, the analysis of historical databases of permanent and temporary streams as well as of targeted WWTPs (where these existed) and data from the small communities obtained through a detailed survey distributed to each one of the municipalities, formed the third source of knowledge.

In order to produce a knowledge base as comprehensive and accurate as possible, we organised the knowledge acquired from the three sources described above into three distinct type of knowledge:

1. Knowledge for the (quantitative) assessment of the characteristics and state of the receiving environment, such as quantity of water in the stream, presence of aquifers, sensible zones, groundwater nitrate pollution vulnerability, and protected areas. This knowledge, acquired through conversations with experts from the Catalan Water Agency (ACA), allowed us to determine the minimum treatment level for each case consistent with the current state of the receiving environment, such as primary treatment, secondary treatment—only carbon removal—, secondary with nitrification, secondary with N/D, nutrient removal or nutrient removal plus disinfection.
2. Knowledge for the identification of disposal sites and characteristics for each community, such as number of inhabitants, surface available, climatic, geological and hydrological conditions, future prospects, benefits

and impacts of the new WWTP, and other economic, social and environmental aspects. This was obtained through a survey of municipalities conducted by an engineering firm. One caveat of this type of survey is that the answers given to the questionnaire by municipal officers may be subjective, and hence qualitative and vague. It is nonetheless a valuable tool since it provides information on the territory and the environment that can be obtained only from local knowledge. Moreover, the views of local officers on the selection of treatment often differ from those of experts, and should be included in the decision-making process.
3. Knowledge about the treatment alternatives for small communities, with information about removal efficiencies, space requirements, climatic constraints, geological and hydrographical features, such as altitude, slope, presence of aquifers, groundwater nitrate pollution vulnerability, investment and operating costs, manpower, social aspects and any advantage and disadvantage that must be considered to be implemented.

2.2.1.3. Model selection   Among the several types of knowledge-based systems, we chose a rule-based system (RBS) because it allowed the best representation of the knowledge needed to select the optimal wastewater treatment system, with due consideration to the receiving environment and to the characteristics of the rural community. We developed the rule-based expert system in two main parts. In the first one, the RBS assists in the selection of the treatment level adequate to fulfil the target quality standards for the receiving environment. In the second one, the RBS is subsequently used to select the specific type of treatment.

2.2.1.4. Model implementation   Once the knowledge acquisition process was completed, we proceeded to structure the acquired knowledge or, in other words,

transform it into a graphical representation that is easy to understand and amend by experts. For instance, the knowledge about the receiving environment was organised and documented in the form of decision trees as a prior step to developing the part of the rule-based system for the selection of the level of wastewater treatment as a function of the receiving environment (Fig. 10).

In addition, the knowledge acquired about treatment alternatives allowed us to construct two useful matrices. One allows the qualitative comparison of the alternative treatments based on economic, environmental, technological, and other criteria. The other matrix associates the level of treatment established for the receiving environments with the optimal treatment system for each case. These two matrices formed the basis for a *hierarchical discriminant table*, which, after many modifications aimed at removing redundancies and contradictions, became the core of the rule-based system to support the specific type of treatment.

The function of this table is to assess the value of four key variables (inhabitant-equivalents, level of treatment required, water flow of the receiving media and available surface of the site to build the WWTP) for the selection of treatment and propose one or more alternatives for wastewater treatment for each of the communities. In addition to these four key variables, we organised the remaining considerations for each type of treatment as a series of the so-called *safety* rules:

- *Discarding rules* include criteria for discarding a particular treatment proposed as an alternative.

- *Favouring rules* evaluate criteria for favouring certain treatments.
- Disadvantaging *rules* evaluate criteria that lower the value of certain treatments in certain situations.

The integration of the RBS with geographical information and a numerical model, for economical estimations of each alternative, was accomplished with an EDSS built on a hierarchical multilevel architecture (data gathering, diagnosis, decision support, plans and actions levels). In addition, the EDSS developed provides easy connectivity with external applications, including database operations. Finally, a menu-based interface provides a simple and transparent way to communicate with end-users.

2.2.1.5. EDSS validation   The execution of a series of experiments with preliminary real data collected from the receiving media and small communities enabled us to validate the accuracy, correctness, consistency, and usability of the acquired knowledge. When necessary, the knowledge base was confronted against experts and the rules were refined, adjusted, corrected and/or extended.

Once validated, the EDSS was applied to all the different agglomerations comprised in the Small Communities Wastewater Treatment Plan of Catalonia. The 3482 small communities were grouped by river basins and processed by the EDSS to obtain a proposal with the most suitable treatment for each community. Alternatively, each small community can be processed individu-
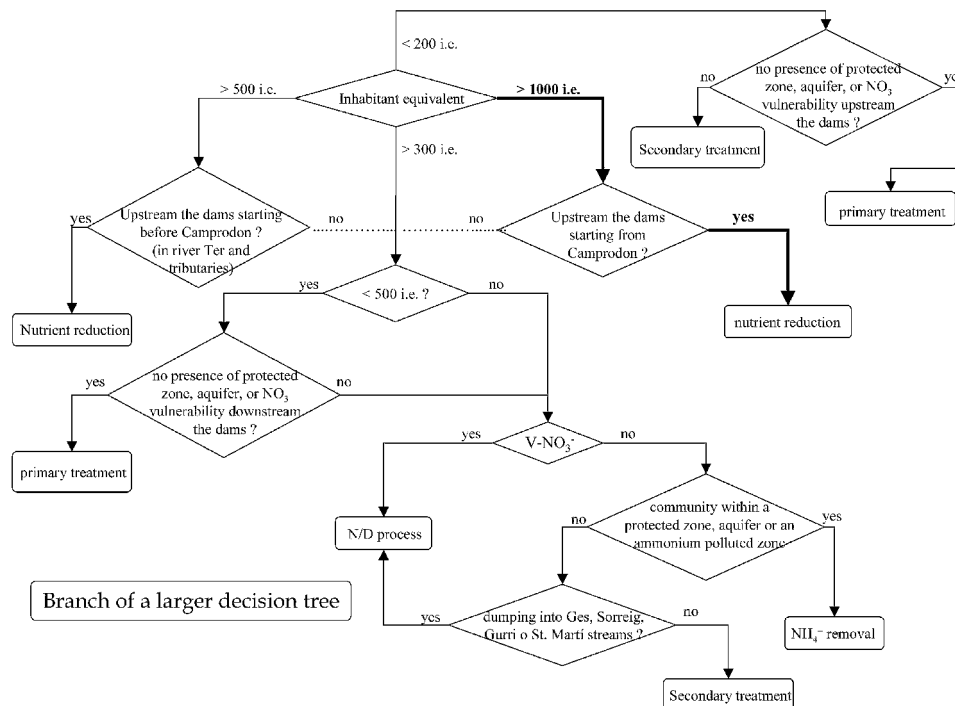

Fig. 10.   Decision tree for the selection of the level of treatment.

ally by the EDSS. This option is better if we want to identify the reasoning path followed by the EDSS. The results of the EDSS application to about 3500 real *cases* were proved satisfactory because the EDSS proposed the optimal solution according to the pool of experts. Some of the WWTPs proposed by the EDSS are already under the building's project step.

2.2.2. EDSS operation

Fig. 11 offers a schematic representation of how the EDSS proceeds to provide the optimal treatment alternative for a particular community (or for all the communities within a particular catchment). For each community, the level of treatment is first established in order to maintain or improve the current ecological state of the receiving media and then, according to the treatment and community features, a set of possible alternatives of treatment are proposed (not only one). Each one of these alternatives setting to the level of treatment can be later favoured, disadvantaged or discarded, according to the features of the community and receiving media. The steps followed during the EDSS operation may be summarised as follows.

2.2.2.1. Data gathering The user introduces the code of the system or catchment for which wastewater treatment alternatives are required. The EDSS then reads the

database that stores the information on the place gathered from the municipal survey or from GIS databases. The data contained in the knowledge base are subsequently filtered and abstracted. Filtering consists of a number of operations aimed at discarding erroneous, foreign or missing data. Abstraction transforms quantitative variables into qualitative variables.

2.2.2.2. Diagnosis The decision support system activates a set of rules that evaluates the number of inhabitant-equivalents, the level of treatment, the abundance of water in the environment, and the area of land available for treatment facilities. This step concludes with a shortlist of alternative treatments. Subsequently, the safety rules for the treatment alternatives included in the shortlist are activated. These rules may invoke other rules or procedures (subroutines of the expert system) until a final list of possible treatments is obtained. For each alternative, this list provides an economic evaluation of the investment and operational costs and indicates whether the alternative has been discarded, favoured or disadvantaged, and, if so, the reasons why.

2.2.2.3. Decision support For each community, the EDSS provides an economic evaluation of the cost of construction and operation of each of the alternatives as a function of the number of inhabitant-equivalents to be
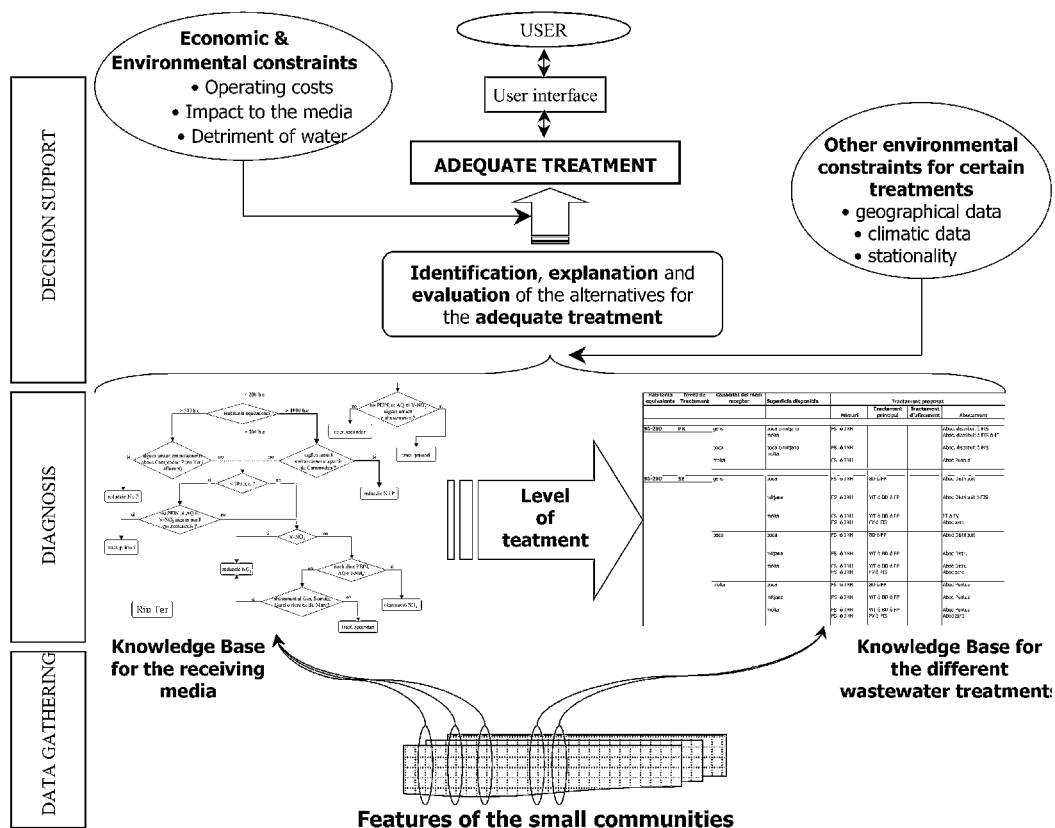

Fig. 11.   EDSS operation.

treated and the type of treatment selected. For each solved system (whether it is a community, a set of communities for a given catchment, or a set of neighbouring communities), a report is produced containing the following results:

- Characteristics of the community used in the reasoning process of the EDSS
- List of the selected treatment alternatives marking which have been discarded, favoured or disadvantaged.
- Environmental technical justification for the selected treatments and the reasons for discarding, favouring or disadvantaging it.
- Economic evaluation of each alternative.

Fig. 12 illustrates one of the interfaces (in Catalan) to show the results of the EDSS to the final users. In this case, six possible wastewater treatments are proposed. Some of these alternatives are favoured (√) and some are disadvantaged (!) for different circumstances. For example, waste stabilisation ponds (or lagoons) are favoured with respect to the other treatments because: (1) this community presents an important contribution of the seasonal population (seasonal population/permanent population = 4.118) and (2), this community belongs to a region with suitable climatic conditions for optimal performance of ponds. On the other hand, waste stabilisation ponds are disadvantaged with respect to the other treatments because the site where the WWTP will be

constructed is situated at less than 200 m away from the community, which means that, in windy days, odours from ponds could arrive to the population. The system can be requested to offer explanations about the conclusions reached and the deductive processes followed. In this example, none of the possible treatments is discarded (X).

2.2.2.4. Plans and actions   In order to make a final decision on the optimal treatment alternative for a given community, or on the optimal wastewater treatment alternative for a system composed by more than one community (e.g. to decide implementing separate or combined treatment systems for a set of communities, or connect them to an already existing or planned WWTP), a consensual function was developed among experts in wastewater treatments, those on the receiving environment, the administration, and engineering firms.

This function allows a numerical evaluation according to three factors: an energetic or economical factor (i.e. the operating costs), an ecological factor or impact to the receiving media factor (the ratio of wastewater and receiving media water flows corrected by the removal efficiency of each treatment) and an environmental factor considering the detrimental impact on the receiving stream (i.e. the length of sewers). Considering this selection function, the EDSS finally proposes a set of alternatives with a hierarchical order from the optimum to the less recommended option. The decision-maker should finally decide among these alternatives.

Fig. 12.   User-interface showing the results of the EDSS proposal for the *Malavella Parc* community.

3. Discussion and conclusions

Environmental problems are complex in the ecological domain, and usually controversial in the socio-economic domain. The *optimal* solution to those problems may be more easily found by tight cooperation among scientists from several research fields and decision-makers. EDSSs are increasingly used as a basis for better decision-making in many real applications. In this paper, a methodology and a possible architecture for EDSS development has been proposed. Also, the description of two real case studies within the water domain has been detailed to ensure the reliability of the approach.

From our experience in the development of EDSS during last 10 years, it can be foreseen that the future of EDSS research will be focused on the following issues.

3.1. Integration of several sources of data and knowledge

Integration of various sources of knowledge, intelligent techniques and numerical tools is the key step to develop successful EDSS for environmental problems. Intelligent decision-making requires, either implicitly or explicitly, a model of the world that embodies both prior knowledge and measured data. At the level of data and background information, numerous and often incompatible bits of information from disparate sources have to be brought together. At the level of tools, there are several levels of integration, ranging from simple file transfer between different methods and programs to fully integrated systems. Typical examples of different methods that lend themselves to integration include geographical information systems and models as well as rule-based systems, models and databases, algorithmic models and intelligent reasoning systems, simulation and optimisation models.

3.2. Improvement of knowledge acquisition methods

EDSSs use different knowledge sources and this usually implies different ways to represent, extract and combine information. The nature of the problems that EDSSs try to solve makes the knowledge acquisition step a crucial one. For most of the problems, there exist huge quantities of data on the process itself, but the information on the causal or dependence relations among variables is not well known. In many cases, AI tools are used to discover those relationships. Future work in this field should also include the integration of knowledge (about the problem domain) in the data mining task to increase meaningful knowledge extraction (supervised knowledge acquisition).

A possible solution to integrate and share information about knowledge structures is to build and use ontologies. This task is only starting to be generally recognised as a key issue in environmental fields. Therefore, this is the appropriate moment to define the relevant entities. Ontologies could be used to assess and evaluate the knowledge about a certain topic or situation with the goal of informing decision-makers. Ontologies can give answers to some of the following questions: What is known and with what degree of certainty? What is not known? What is the relevance of that knowledge to decision-makers? Construction of specific ontologies or equivalent paradigms could represent better the *know-how* and *know-what* in environmental systems.

3.3. Elaborate protocols to facilitate sharing and reuse of knowledge

Once an EDSS has acquired information on a complex environmental process, what are the available ways to share that information with other systems? If EDSSs are designed to be cooperative, under which conditions does this cooperation occur? What happens if cooperation fails? Who will assess the quality of the exchanged information? Who will harmonise indicators and exchange protocols?

Solutions for sharing knowledge in environmental processes are far from being fully developed, but one has to consider the great variety of data, and the strong dependencies of environmental processes to local constraints, such as weather conditions, climatic aspects, geographical positions, environmental or health law regulations, etc. If specific models are to be developed for environmental problems, greater generality, precision (when possible) and realism will be required.

3.4. Involvement of end-users in EDSS development

In general, the role of the user in EDSS development is still poorly defined. These systems are developed to support users' decision-making activities in highly complex problems.

The following questions are still to be answered: to what extent can an EDSS be modified directly by *any* user? Who should decide that an EDSS has to start a learning process? Who has to validate the results of such process? Why should an EDSS start a learning process? Who is legally responsible for the decisions made by an EDSS?

We propose, as a first approach, the creation of user profiles with different privileges and responsibilities in the interaction with the EDSS. This will lead to the definition of different levels of interaction between the user and the EDSS.

On the other hand, the users must be involved in the whole process of EDSS design and development to ensure the usability of the final system. The degree to which the users become involved in EDSS development

will determine their level of confidence in the final system. In the worst case, the system might remain unused.

3.5. Development of benchmarks for the validation of EDSS

In our opinion, one of the most promising research lines in EDSS development is the definition of benchmarks to assess and evaluate the performance of EDSSs in a set of well-defined circumstances, and their capacity to react to new situations. This will also allow the creation of a better framework for comparison between EDSSs.

We are aware of no attempt to do this. This validation of an EDSS in the appropriate context may simplify the tuning tasks and help to enhance the system's performance.

3.6. Final conclusion

Environmental issues belong to a set of critical domains where wrong management decisions may have disastrous social, economic and ecological consequences. Decision-making performed by EDSSs should be collaborative, not adversarial, and decision-makers must inform and involve those who must live with the decisions. What an EDSS contributes is not only an efficient mechanism to find an optimal or suboptimal solution, given any set of whimsical preferences, but also a mechanism to make the entire process more open and transparent. In this context, EDSSs can play a key role in the interaction of humans and ecosystems, as they are tools designed to cope with the multidisciplinary nature and high complexity of environmental problems.

References

Belanche, L., Valdés, J.J., Comas, J., R-Roda, I., Poch, M., 1999. Towards a model of input–output behaviour of wastewater treatment plants using soft computing techniques. Environmental Modelling and Software 14 (5), 409–419.

Belanche, L.I, Valdés, J.J., Comas, J., R-Roda, I., Poch, M., 2000. Prediction of the bulking phenomenon in wastewater treatment plants. Artificial Intelligence in Engineering 14 (4), 307–317.

Comas, J., Rodríguez-Roda, I., Sánchez-Marrè, M., Cortés, U., Freixó, A., Arráez, J., Poch, M., 2003. A knowldge-based approach to the deflocculation problem: integrating on-line, off-line, and heuristic information. Water Research 37 (10), 2377–2387.

Cortés, U., Sànchez-Marrè, M., Cecaronni, L., R.-Roda, I., Poch, M., 2000. Artificial intelligence and environmental decision support systems. Applied Intelligence 13 (1), 77–91.

Cortés, U., Sànchez-Marrè, M., Sangüesa, R., Comas, J., R-Roda, I., Poch, M., Riaño, D., 2001. Knowledge management in environmental decision support systems. AI Communications 14 (1), 3–12.

Cortés, U., Sànchez-Marrè, M., Comas, J., Rodríguez-Roda, I., Poch, M., 2002. Unpublished results.

El-Swaify, S.A., Yakowitz, D.S. (Eds.), 1998. Multiple Objective Decision Making for Land, Water, and Environmental Management. Lewis Publishers.

Fox, J., Das, S., 2000. Safe and sound. In Artificial Intelligence in Hazardous Applications. AAAI Press/The MIT Press.

Funtowicz, S.O., Ravetz, J.R., 1993. Science for the post-normal age. Futures 25 (7), 739–755.

Funtowicz, S.O., Ravetz, J.R., 1999. Post-normal science—an insight nom maturing. Futures 31 (7), 641–646.

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

Haagsma, I.G., Johanns, R.D., 1994. Decision support systems: an integrated approach. In: Zannetti, P. (Ed.), Environmental Systems, vol. II., pp. 205–212.

Kinzig, A., 2001. Bridging disciplinary divides to address environmental and intellectual challenges. Ecosystems 4, 709–715.

Ludwig, D., 2001. The era of management is over. Ecosystems 4, 758–764.

Olsson, G., Aspegren, H., Nielsen, M.K., 1998. Operation and control of wastewater treatment—a Scandinavian perspective over 20 years. Water Science and Technology 37 (12), 1–13.

Ostrom, E., 1991. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press.

Rizzoli, A.E., Young, W.Y., 1997. Delivering environmental decision support systems: software tools and techniques. Environmental Modelling and Software 12 (2-3), 237–249.

R.-Roda, I., Comas, J., Poch, M., Sànchez-Marrè, M., Cortés, U., 2001. Automatic knowledge acquisition from complex processes for the development of knowledge-based systems. Industrial Engineering Chemistry Research 40, 3353–3360.

Rodríguez-Roda, I., Comas, J., Colprim, J., Poch, M., Sànchez-Marrè, M., Cortés, U., Baeza, J., Lafuente, J., 2002. A hybrid supervisory system to support wastewater treatment plant operation: implementation and validation. Water Science and Technology 45 (4-5), 289–297.

Sànchez-Marrè, U., Cortés, J., Lafuente, I., R.-Roda, I., Poch, M., 1996. DAI-DEPUR: a distributed architecture for wastewater treatment plants supervision. Artificial Intelligence in Engineering 10 (3), 275–285.

Sydow, A., Rosé, H., Rufeger, W., 1998. Sustainable development and integrated assessment. Ercim News (32), 34.