International Congress on Environmental Modelling and Software

1st International Congress on Environmental Modelling and Software - Lugano, Switzerland - June 2002

Jul 1st, 12:00 AM

Ten years of experience in Designing and Building real Environmental Decision Support Systems. What have we learnt?

Manel Poch

Joaquim Comas

Ignasi Rodríguez-Roda

Miquel Sànchez-Marrè

Ulises Cortés

Ten years of experience in Designing and Building real Environmental Decision Support Systems. What have we learnt?

These matters that with myself I too much discuss
Too much explain
T.S. Elliot

Manel Poch[1], Joaquim Comas[1], Ignasi Rodriguez-Roda[1], Miquel Sànchez-Marrè[2], Ulises Cortés[2]

[1]*Laboratori d'Enginyeria Química i Ambiental, Univ. de Girona, Campus Montilivi, 17071, Girona, Spain. (manel@lequia.udg.es)*

[2] *Knowledge Engineering and Machine Learning Group, Technical University of Catalonia, c/ Jordi Girona 1&3, 08034 Barcelona, Spain.*

Abstract: The complexity of environmental problems make necessary the development and application of new tools capable of processing not only numerical aspects, but also experience from experts and wide public participation, all which are 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 ten years, in designing and building two real EDSSs, one for wastewater plant supervision, and one for the selection of wastewater disposal systems for communities of 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 about 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.

Keywords: Environmental Decision Support Systems, Artificial Intelligence, Wastewater Treatment

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 negatively the quality of water, air, and soil, 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 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 hand, 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, 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 a industrial process, this level of complexity is associated, for instance, with the environmental aspects of sewage 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 characterized by dynamics and interactions that do not allow for an easy division between social and biogeophysical phenomena. Much ecological theory has 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 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 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, Cortés *et al.* (unpublished results) found more than 600 references (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. 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 ten 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 about what constitutes an environmental decision support system. 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

Following Haagsma and Johanns [1994] and Cortés [2001], we define an EDSS as an intelligent information system that helps reduce the time in which decisions are made, and improves the consistency and quality of those decisions.

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 (figure 1)

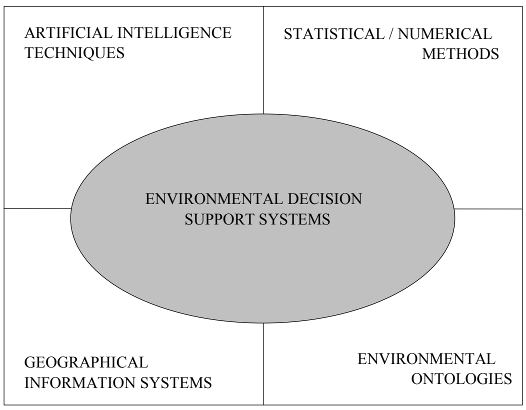

Figure 1. EDSS conceptual components

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, 2000].

This confers EDSSs the ability to confront complex problems, in which the experience of experts provide 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 can learn something new, and, last but not least, for its overall performance. Figure 2 shows schematically the methodology used in the two study cases presented here.

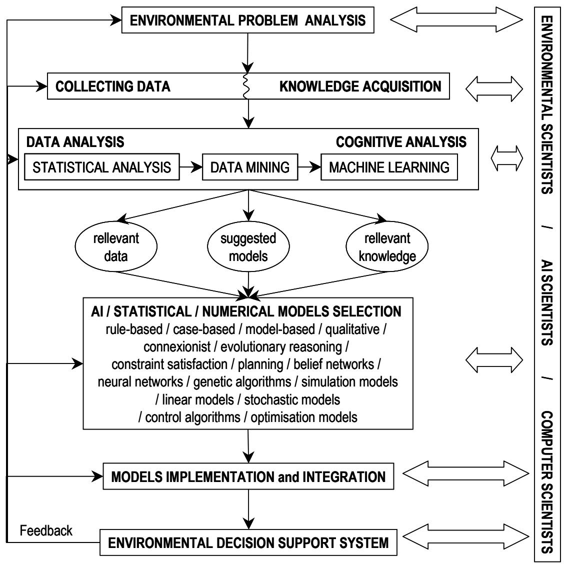

Figure 2 Flow diagram for development of an EDSS

We propose an EDSS architecture based on five levels:

- The first level of the EDSS encompasses the tasks involved in data gathering and registration into databases. Original raw data are often defective, requiring a number of preprocessing procedures before they can be registered in an understandable and interpretable way.
- The second 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 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.
- 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 a value that has to be accepted by the decision maker. This is the last level in the architecture that closes the loop.

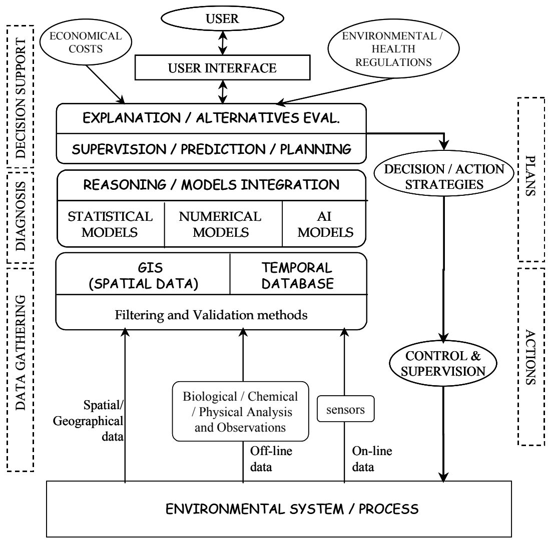

Figure 3. EDSS Architecture

2. TWO CASE STUDIES

In this section we present two case studies where the proposed methodology has been applied. The two case studies correspond to two different situations with different forms of complexity.

The first case corresponds to the application of an EDSS to the supervision of a WasteWater Treatment Plant (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 disposal and treatment 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

Problem analysis

A typical wastewater treatment plant usually includes a physical and/or chemical primary treatment and a biological secondary treatment to remove organic matter and suspended solids from wastewater. 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 multi-disciplinary 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 characterized 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.

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, 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 with numerical and conventional computational techniques (statistical methods, advanced and robust control algorithms and system identification techniques).

The wastewater treatment plant selected to develop and apply our proposed Supervisory System prototype is located in Granollers, in the Besòs river basin (Catalonia, NE of Spain). 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.

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 either supervised (mainly inductive learning techniques, CN2, C4.5 and *k*-NN]) or unsupervised [Rodríguez-Roda et al, 2001]. Figure 4 illustrates the main sources and methods that were used to acquire both kinds of knowledge on the wastewater treatment processes.

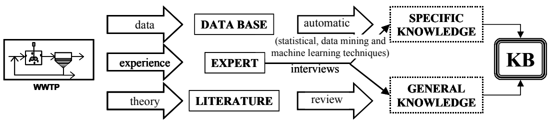

Figure 4 Methods used to acquire knowledge

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

Expert systems offer a number of advantages 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 ES offers potentially optimal answers because action plans are systematised for each problematic situation; and, finally, expert systems make possible the acquisition of a large general knowledge base, with flexible use for any WWTP.

The proposed Case-Based 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 about how to achieve the reasoner's goals.

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]. 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, comprised 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.

CBSs require a library of cases to broadly cover the set of potential problems. These cases are indexed in memory so as to be retrievable 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 technical books 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.

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 in its actual operating environment, where it worked as a real-time decision support system for more than 10 months. During this period of exhaustive validation, the

EDSS successfully identified 123 different problem situations and suggested suitable action strategies. Nowadays, the EDSS is used as a complementary tool of diagnosis for the everyday management of the activated sludge process [Rodríguez-Roda, 2002].

2.1.2 EDSS operation

Data gathering

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.

Every time a supervisory cycle is launched, the main task to be performed is *gathering data* and *updating* current data for the inference process. 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.

Diagnosis

Once the information has been collected, it is 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 similar case is modified so as to adapt it to the new situation.

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*). The expert can use a dynamic model implemented to support the selection process of an action plan by simulating the possible consequences of applying different alternatives.

Plans and actions

The EDSS suggests an action plan resulting from the *supervision* and *prediction* tasks, and integrating the expert recommendations sent by the ES and the experience retrieved by the CBS, while it evaluates any possible conflict. The *evaluation* of the results of the application of the action plan to solve the problem in the process allows the system *to close* the CBS cycle, to learn from successful and failed past experiences, and to upgrade the case-library. These features can be detected by the EDSS itself (unless a manual operation is carried out), but it is essential 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 base by acquiring new knowledge from new sources.

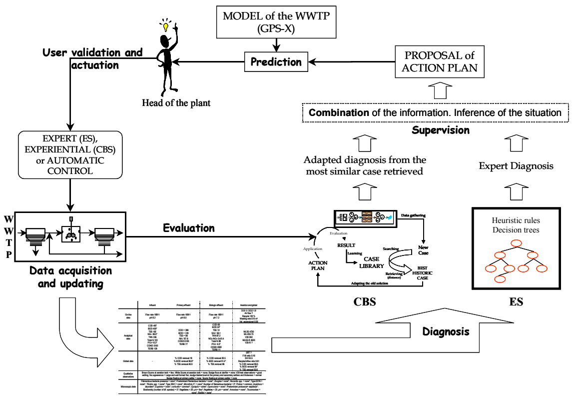

Figure 5. Supervisory cycle

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

2.2.1 EDSS building

Problem analysis.

Wastewater Treatment Plants (WWTPs), and especially biological plants based on several variants of the activated sludge process with free biomass, are currently the predominant system for urban sewage disposal in Catalonia. In accordance with the goals established in the First Urban Sewage Treatment Programme (PSARU I), WWTPs have been built to serve every town in Catalonia with more that 2000 inhabitants. In communities with less than 2000 inhabitants, however, the situation is different. Few of them have wastewater treatment systems in place today, but all should have them by 2005.

While an European Water Directive specifies the type of treatment to implement in towns with more than 2000 inhabitant-equivalents, for smaller communities the Directive only states that the type of treatment should be appropriate. This changes significantly 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 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 recruit 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 rural community itself. This is an aspect of evident importance given the large number of rural communities in our country (more than 3000) 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, 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 Second Urban Sewage Treatment Plan (PSARU II) is implemented. Protection of the receiving environment is of the highest importance, as endorsed by the recent Directive 2000/60. In order to improve on the current state, an assessment of the current ecological quality of the site is needed.

(3) The sewage 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 scenarios in support of the decision-making process that will have taken into account not just aspects of technical optimisation, but also environmental, economic 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 Environmental Decision Support System. 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 recognize 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 Environmental Decision Support System was chosen because it integrates expert knowledge and because it 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 allows the management and analysis of large volumes of data.

Collecting data and knowledge acquisition
Three different sources of knowledge were pooled together to build a knowledge base as comprehensive as possible. These sources were:

- Knowledge extracted from interviews with experts in water management and sewage treatment, as well as experts in the quality of the receiving environment.
- Knowledge drawn from the scientific and technical literature as well as from visits to regions where this type of wastewater disposal programme has already been implemented.

- Knowledge derived from the analysis of historical data for the receiving environment as well as data on the rural communities themselves.

The first source of knowledge we turned to was a group of experts in the sewage treatment process. 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 specialized journals and books, and from notes taken during the filed visits. Finally, the analysis of historical databases of permanent and temporary streams as well as of targeted wastewater treatment plants (where these existed) formed the third source of knowledge.

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

1. A knowledge base for the (quantitative) assessment of the characteristics and state of the **receiving environment**. This knowledge, acquired through conversations with experts from ACA, allowed us to determine the minimum treatment level for each case consistent with the current state of the receiving environment.
2. A knowledge base for the identification of **disposal sites** for each community. 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 about 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. A knowledge base about the **treatment alternatives**, with information about performance, space requirements, climatic constraints, installation and operation cost, etc.

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

Model implementation
Once the knowledge acquisition process was completed, we proceeded to structure the acquired knowledge or, in other words, to transform it into a graphical representation easy to understand and amend by experts. For instance, the knowledge base about the receiving environment was organized and documented in the form of decision trees as a prior step to developing the expert system for the selection of wastewater treatment as a function of the receiving environment.

In addition, building a knowledge base for treatment alternatives allowed us to construct two useful matrices. One allows the qualitative comparison of alternative treatments based on economic, environmental, technological, and other criteria. The other matrix associates the levels 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 expert system.

The function of this table is to assess the value of four key variables for the selection of treatment and propose one or more alternatives for sewage treatment for each of the communities. In addition to these four key variables, we organized the remaining considerations for each type of treatment as a series of 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.
- *Unfavouring rules* evaluate criteria that lower the value of certain treatments in certain situations.

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 began to be applied to 3482 different small communities comprised in the *Small Communities Wastewater Treatment Plan of Catalonia*, grouped according to river basins. At the moment of writing, nine of the twenty river basins of Catalonia have already been processed and the corresponding 951 small communities (27%) already have a proposal for the most suitable treatment. Some of the WWTPs proposed by the EDSS are already under the building's project step.

2.2.2 EDSS operation

Figure 6 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). The steps followed by the EDSS may be summarized as follows:

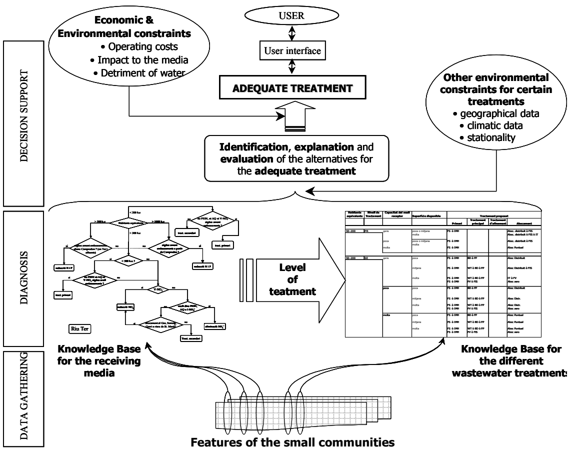

Figure 6. EDSS operation

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 about the place gathered from the municipal survey or from GIS databases (elevation and groundwater nitrate pollution vulnerability). 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.

Diagnosis
The decision support systems 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 amount 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 o procedures (subroutines of the expert system) until a final list of possible treatments is obtained. For each alternative, this list provides an economic evaluation and indicates whether the alternative has been discarded, favoured or unfavoured, and, if so, the reasons why.

Decision support
For each community, the EDSS can provide 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 treated and the type of treatment selected. For each solved system (whether it is a community, a set of catchments for a given community, 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 DSS
- List of the selected treatment alternatives marking which have been discarded, favoured or depreciated.
- Justification for the selected treatments and the reasons for discarding, favouring or unfavouring it.

Plans and actions
In order to make a final decision on the optimal treatment alternative for a given community, or on the optimal wastewater disposal alternative for a system (e.g., implement separate or combined treatment systems for a set of communities or catchments, or connect them to an existing or planned WWTP), a consensual formula was developed among experts in wastewater treatments, experts on the receiving environment, the administration, and engineering firms. This formula allows the evaluation of the energetic and environmental impact (including the potentially detrimental impact on the receiving stream) of each treatment system.

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. It can be foreseen that the future of EDSS research will be focused on the following issues:

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 expert systems, models and databases, algorithmic models and expert systems, simulation and optimisation models.

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 about the process itself, but the information about the causal or dependence relations among variables is not well known.

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?

Elaborate protocols to facilitate sharing and reuse of knowledge
Once an EDSS has acquired information about a complex environmental process, which 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.

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, 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 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.

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.

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 sub-optimal 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.

4 ACKNOWLEDGEMENTS

We wish to thank Josep Bou and Marta Lacambra (Agència Catalana de l'Aigua) and Angel Freixó and Josep Arraez (Consorci per a la Defensa de la Conca del riu Besòs) for their confidence in the possibilities of EDSS. We would also like to acknowledge the IST-2000-28385 Agentcities.NET and the IST-1999-10176 A-TEAM projects. The views expressed in this paper are not necessarily those of Agentcities.NET and A-TEAM consortia. We also benefited from the partial support of CICyT TIC2000-1011 and CICyT DPI2000-0665-C02-01 grants.

5 REFERENCES

Cortés, U., Sànchez-Marrè, M., Sangüesa, R., Comas, J., R-Roda, I., Poch, M. & Riaño, D. "Knowledge Management in Environmental Decision Support Systems". *Ai Communications*. 14(1):3-12 2001.

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

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

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

Funtowicz, S.O. and Ravetz, J.R. Post-Normal Science – an insight nom maturing, *Futures* 31(7), 641-646, 1999

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

Haagsma, I.G. and Johanns, R.D. Decision support systems: An integrated approach. In Environmental Systems Vol. II. Zannetti, P. (Ed.), pp. 205-212, 1994.

Kinzig, A., Bridging Disciplinary Divides to Address Environmental and Intellectual Challenges, *Ecosystems*, 4, 709-715, 2001

Ludwig D., The Era of Management Is Over, *Ecosystems*,4, 758-764, 2001

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

Ostrom, E. Governing the commons: the evolution of institutions for collective action. Cambridge University Press. 1991

Rizzoli, A.E., and Young, W.Y. Delivering Environmental Decision Support Systems: software tools and techniques. Environmental Modelling and Software. 12(2-3),237-249, 1997.

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

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

Sànchez-Marrè, U. Cortés, J. Lafuente, I. R.-Roda, and M. Poch., DAI-DEPUR: a Distributed Architecture for Wastewater Treatment Plants Supervision. *Artificial Intelligence in Engineering*, 10(3), 275-285, 1996.

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