International Congress on Environmental Modelling and Software

2nd International Congress on Environmental Modelling and Software - Osnabrück, Germany - June 2004

Jul 1st, 12:00 AM

Implementation of the STREAMES environmental decision-support system

David Cabanillas

Esther Llorens

Joaquim Comas

Manel Poch

Luigi Ceccaroni

See next page for additional authors

Presenter/Author Information
David Cabanillas, Esther Llorens, Joaquim Comas, Manel Poch, Luigi Ceccaroni, and Steven Willmott

Implementation of the STREAMES environmental decision-support system

David Cabanillas[1], **Esther Llorens**[2], **Joaquim Comas**[2], **Manel Poch**[2], **Luigi Ceccaroni**[1] **and Steven Willmott**[1]

[1]Universitat Politècnica de Catalunya,
LSI department, Campus Nord, Jordi Girona 1-3, E-08034 Barcelona
Catalonia, Spain
e-mail: {david / luigi / steve}@lsi.upc.es

[2]Universitat de Girona,
Campus Montivili, E-17071 Girona
Catalonia, Spain
e-mail: {quim / esther / manel}@lequia.udg.es

Abstract: As part of the European-Commission--funded STREAMES project, a system is being developed with the objective of capturing knowledge from water managers and environmental-science experts, regarding nutrients-excess effects in streams and of combining this knowledge into a user-friendly tool to assist water managers in evaluating streams' nutrient-retention capabilities. In this paper, we summarize the decision-support knowledge components which have been identified in previous work and, based on these, present an implementation of a prototype of an environmental decision-support system. The decision support provided by the system to water managers consists of: (1) diagnosis: inferring possible stream problems, assessing the alteration degree of the stream, and evaluating the source and magnitude of nutrient loads; (2) actions: offering alternative, ranked courses of action to solve possible problems; (3) forecast: providing several scenarios to simulate the effect of the different actions proposed as solutions.

***Keywords**:* Environmental decision-support system, implementation, river, rule-based expert system, water management

1. INTRODUCTION

In the context of European Commission's Fifth Framework Programme (1998-2002) and Water Framework Directive (WFD) (2000/60/EC), the STREAMES project[1] aims to analyze nutrient cycles in a particular, human-altered environment: the river ecosystem, with special emphasis on the Mediterranean region. The decision-making process involved in altered-rivers management requires extensive human expertise from people directly implicated in day-to-day stream problems (water managers and environmental-science experts), knowledge from different science fields and complex calculations over large amounts of numerical and symbolic data. Therefore stream optimal management requires an integrated and interdisciplinary approach. To face this complexity, the STREAMES project aims to develop and implement a knowledge-based system, which will contribute achieving a good ecological state in rivers with bad water quality. This system manages general knowledge extracted from different literature sources as well as specific knowledge acquired by processing empirical data collected from the project's study-sites and from interviews and meetings with human experts.

[1] STream REAch Management, an Expert System. Human effects on nutrient cycling in fluvial ecosystems: The development of an ES to assess stream water quality management at reach scale. [http://www.streames.org, EVK1-CT-2000-00081].

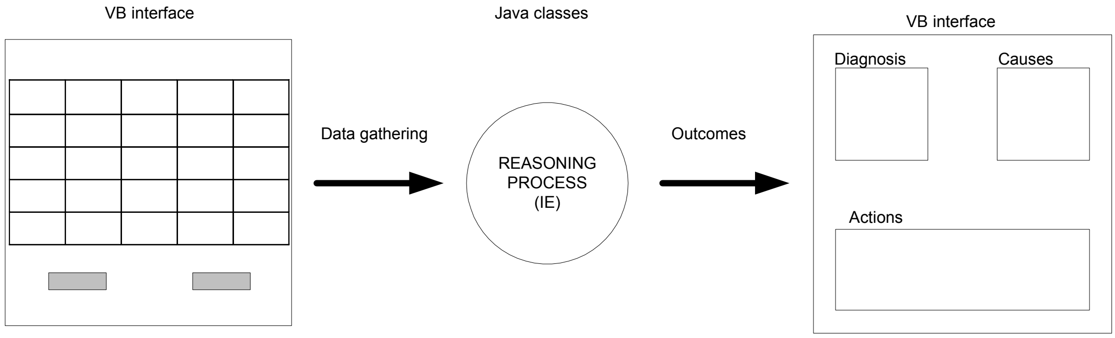

Figure 1. From data to outcomes (simplified)

In the system, artificial intelligence techniques are applied to the water-management field in the form of an environmental decision-support system (EDSS).

EDSSs are a subset of decision support systems (DSSs), which in turn are a subset of computer-based information systems (CBIS). Some examples of EDSSs developed recently and applied to the water domain are described by, among others, Chang et al. [1997], Davis et al. [1998], De Marchi et al. [1999], Rousseau et al. [2000], Rodríguez-Roda et al. [2002], the Great Lakes Commission for the Great Lakes States and Provinces [2003], Matthies et al. [2003] and Ceccaroni et al. [2004].

1.1. EDSS development

The development of the STREAMES EDSS has been carried out following a methodology composed of a series of phases, each with its own inputs, activities and outputs (modified from Poch et al. [2002]):

1. environmental problem analysis
2. data collection and knowledge acquisition
3. system analysis and design
4. *problem-solving method* (PSM) selection[2]
5. PSMs integration
6. system implementation
7. validation
8. maintenance

In the case of the STREAMES EDSS, most phases from *problem analysis* to *PSM selection* are described in a previous work by Comas et al. [2002]. With respect to PSM selection, the following ones were chosen:

PSM1. A *rule-based system*, to resolve stream management problems, whose diagnosis and solution involves qualitative data and knowledge processing.

PSM2. *Numerical and statistical models*, to estimate point and non-point nutrient inputs and self-purification capacity.

In the STREAMES project, these PSMs are complemented with *geographical information systems* and the rules of the expert system are grouped into four modules (see 2.3).

2. IMPLEMENTATION

In this paper, we analyze the implementation of the *rule-based expert system* (RBES) and of the graphical user interface. RBESs are mainly composed of a knowledge base (KB) and an inferential engine.

2.1. Inferential engine

The inferential engine (IE) works with rules (see PSM1 in section 1.1) and provides the reasoning mechanism. In our case the inference is backward chaining. With the objective of being able to reuse the RBES, we developed an IE shell that can be adapted and customized. In case of reuse in different domains, the KB would need to be redeveloped with new knowledge-components (as described in section 2.2). The current implementation of the IE is Java-based (platform independent) and is integrated with a friendly user-interface in Visual Basic (VB) (see Figure 1).

Once the system is started the user has to fill in different forms of data-input that the system presents via the user-interface. In the IE, the rules correspond to the decision trees described in section 2.2.1 and the facts correspond to the data introduced by the user.

[2] The term *PSM* corresponds to the term *model* used by Poch et al [2002].

Decision tree ID	Decision tree name	Represented problems
DT1	Nitrogen	- Excess of ammonium - Excess of nitrate - Excess of nitrite
DT2	Phosphorous	- Eutrophication
DT3	Organic matter	- Excess of organic matter - Anoxia
DT4	Suspended solids	- Excess of suspended solids - Clogging
DT5	Salts	- Anthropogenic alteration of salinity
DT6	Stream characteristics	- Low river--self-purification

Table 1. Decision trees and related diagnosed problems

Then the inference process starts, trying to find out if the facts match some of the antecedents of each one of the active rules. If a rule is triggered, new facts can be introduced into the facts base, as a result of the inference. This process finishes when the IE has tested all the facts with all the active rules. Afterwards, the IE delivers the results to the interface component that parses and shows the results to the user in an appropriate format.

2.2. Knowledge components

For building and validating the KB of a decision-support system for our given practical domain, four knowledge components (KC) are needed:

KC1. a domain ontology for nutrient cycles in river ecosystems (to formally describe terminology and processes);
KC2. a decision-support ontology to formalize the output of the system (see section 2.4);
KC3. a library of decision trees (see section 2.2.1) or an equivalent rule representation scheme;
KC4. a set of domain requirements that are used to select a suitable set of elements of KC3.

In Comas et al. [2002, 2003], KC3 and part of KC4 were made explicit; we developed the two ontologies and the remaining part of KC4.

2.2.1. Decision trees

STREAMES' KB is codified by means of rules, which are sets of *conditions and conclusions*. As a prior step to build the KB, knowledge is structured and represented in decision trees (DTs) [Comas et al., 2003]. Every DT refers to a set of specific problems (shown in Table 1) and is composed of two modules: one for problem diagnosis and one for cause detection.

The developed DTs correspond to those problems for which water managers and environment experts expressed a greater interest and preoccupation. Six DTs have been developed: one for nitrogen-related problems, one for eutrophication[3], one for organic-matter problems (which include part of the anoxia problems[4]), one for suspended solids and clogging, one for salinity problems and one for alterations of the stream ecosystem. While the first five ones are related to physico-chemical elements in the water, the last one is focused on the physical, biological and morphological characteristics of the river ecosystem (riparian zone and streambed), which can affect the river's functionality and self-purification capacity. The self-purification capacity is in turn an important aspect to be taken into account in water pollution problems.

The module of cause detection of the DTs includes a set of pre-defined causes. For example, if a low river--self-purification due to a physical alteration of the system is detected, causes such as the following ones are evaluated: riparian banks destruction, dredging, morphological alteration of the riverbed by human activities, modification of flow regime.

2.3. Rule modules

In the STREAMES KB, rules are grouped into four modules, or *steps*: the first one, *symptom discovery*, is derived from KC4, while the following three ones codify the DTs (KC3). The sequence of the steps is:

1. *Symptom discovery*. If certain symptoms are detected, this meta-rules module activates one or more DTs.
2. *Problem diagnosis*. This module represents the knowledge necessary to diagnose the problem corresponding to

---

3 Eutrophication problem is evaluated by means of the N:P molar-ratio calculation.

4 Oxygen depletion may be due also to ammonium oxidation and eutrophication problems; these situations are not considered by the current version of the EDSS.

the symptoms. A specific problem and its possible side-effects are confirmed and communicated to the user.
3. *Cause detection*. Different, possible causes of the problem under analysis are deduced and evaluated.
4. *Actuation*. A set of actions, corresponding to the causes, is proposed to solve the problem.

For a full understanding of the KB implementation and functioning, as well as the interaction with the user, we present a complete use case of the EDSS.

The process starts with the selection by the user of one of the following two options:
1. evaluation of possible stream problems;
2. assessment of the alteration degree of the stream.

In the following, we consider the first option because it is the one related to the implementation of the RBES.

2.3.1. Symptom discovery

The system begins to gather data, asking questions to the user about groups of significant descriptors (DS), or quality elements for the classification of ecological status. Some of these DS are in accordance with the WFD; other ones have been defined by the authors according to their experience and the knowledge acquired from diverse sources (e.g., EPA manuals by Barbour et al. [1999]):
1. River basin DS. These elements are related to the location of the river in its river catchment, to the characterization of the basin and to the identification of diffuse pollution sources (e.g., geology, predominant land use).
2. Streambed characterization DS. These elements are to estimate the quality of the river in relation to the riverbed. We distinguish two classes:
    a. Biological and habitat DS. These are related to the micro-scale aspects, e.g.: color of sediments, presence of bio-film, fishes, algae, macro-invertebrates.
    b. Streambed DS. These are related to larger-scale aspects, e.g.: types of streambed, channel sinuosity.
3. Hydromorphological DS supporting the biological DS. Examples of these elements are: stream width, water velocity and, in general, the hydrological regime and the river continuity.
4. Water quality DS. Examples of these elements are: nitrogen and phosphorous data, water odor, conductivity, water color, water temperature, pH.
5. Point nutrient-source DS. Identification, location and characterization of the existing point sources of nutrients in the river catchment, e.g.: input of wastewater, ammonium.
6. Riparian DS. These elements characterize the riparian zone and help to estimate the quality of the river in relation to it. Examples are: types of riparian vegetation, soil permeability.

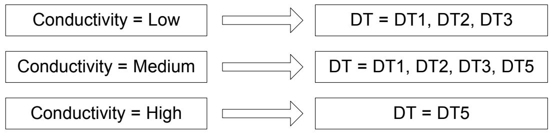

Figure 2. Symptom-discovery meta-rules

These data and a set of meta-rules representing domain requirements are used to select the DTs to be activated (see Figure 2 for an example of these rules).

2.3.2. Problem diagnosis

When, for instance, DT2 (*phosphorous*) is selected, its problem-diagnosis module is activated. Part of the problem-diagnosis rule-inference is shown in Figure 3.

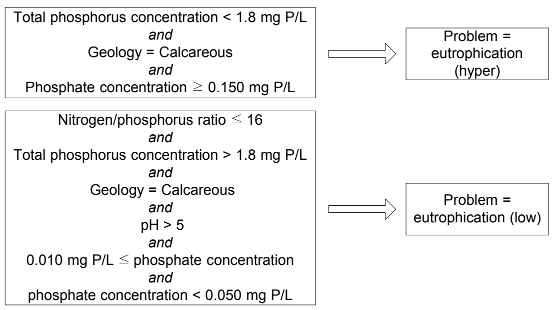

Figure 3. Problem-diagnosis rules for the *phosphorous* decision tree.

In the same way, inference is carried out in the rest of DTs activated by the meta-rules.

2.3.3. Cause detection

For each problem diagnosed, the cause-detection module of the corresponding DT is activated. Part of the cause-detection rule-inference is shown in Figure 4.

2.3.4. Actuation

Once the system executed all triggered rules in activated DTs, it shows the user a set of

<diagnosis, cause> pairs (DCPs), for him to analyze.

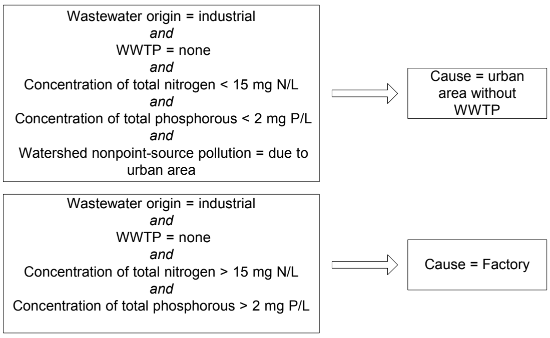

Figure 4. Cause-detection rules for the *phosphorous* decision tree.

The user chooses the DCPs he is interested in and, for each one, the actuation category (hydromorphology, chemistry, biota, best practices, hydrology) and the actuation geographical-scope (river basin, riparian zone, river body). With these data, the system is able to offer an ordered list of recommended courses of action to carry out (see an example in Figure 5), as well as, when possible, a series of complementary parameters, such as: chances of success, feasibility, response time, effort vs. environmental benefit, references.

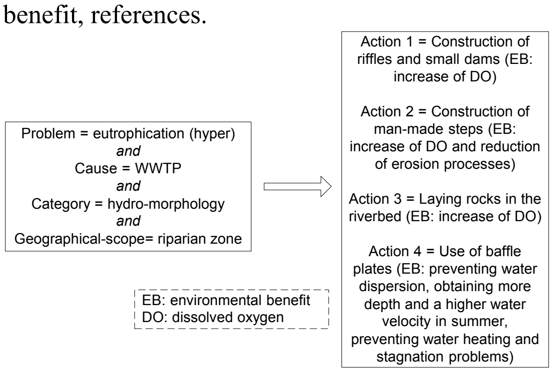

Figure 5. Recommended actions in the *actuation* step (simplified).

2.3.5. Forecast

The system forecasts what improvements would take place in the river if one of the actions suggested were carried out. As outcome, the system shows the user a comparison of the current problematic state versus the state after the application of the action, as well as a measure of the improvement in the quality of water.

2.4. Decision support

In summary, the decision support supplied by the system consists of providing:

1. Diagnosis: inferring possible stream problems, assessing the alteration degree of the stream, and evaluating the source and magnitude of nutrient loads.
2. Actions: offering alternative, ranked courses of action to solve possible problems.
3. Forecast: providing several scenarios to simulate the effect of the different actions proposed as solutions.

An example of the outcome of the system is as follows. The EDSS detects that the stream undergoes a *hyper-eutrophication* problem. Also, the EDSS has been able to infer that the cause related to this diagnosis is a point source (a WWTP without nitrogen removal). According to this diagnosis and cause, the EDSS proposes several actuations: restoration of riparian vegetation, optimization of the nitrification/de-nitrification process, nitrogen removal. Furthermore, the EDSS allows estimating the effect of the actuations proposed for stream improvement. If, for example, *nitrogen removal* were implemented, the nutrient loads into the river would decrease, the problem would be partially solved and the prediction would be *low eutrophication*.

3. CONCLUSIONS AND FUTURE WORK

Recently, attention has been focused on providing decision support for evaluating streams' nutrient-retention capabilities. Such support is needed to guide a water manager in planning actions regarding relief from nutrients-excess effects in streams. This paper contributes to the efforts for building and validating decision-support knowledge-bases (KBs) for the streams domain. We identified four knowledge components explicitly required to develop the KB of an environmental decision-support system (EDSS). These include, in the case of the river domain: (1) a domain ontology for nutrient cycles in river ecosystems (to formally describe terminology and processes); (2) a decision-support ontology to formalize the output of the system; (3) a library of decision trees or an equivalent rule-representation scheme; (4) a set of domain requirements that are used to select a suitable set of decision trees. We summarized the knowledge components which have been identified in previous work and, based on these, presented an implementation that exploits rule-based expert systems to aid water managers in planning practical and effective courses of action in response to early symptom discovery.

Future work includes: (1) integration with other technologies and models, such as the nutrient emission model MONERIS and geographical information systems, to improve the EDSS; (2) introduction of more powerful rules, using fuzzy sets and new operators; (3) automatic rule generation and validation; (4) comparison with other knowledge-based systems, such as case-based and model-based reasoning systems.

4. ACKNOWLEDGMENTS

The research described in this paper is partly supported by the EC projects STREAMES (EVK1-CT-2000-00081) and @LIS TechNET (ALA/2002/ 049-055). The opinions expressed in this paper are those of the authors and are not necessarily those of the STREAMES project's partners. The authors would like to thank the anonymous reviewers of this paper in its draft form, who have provided many constructive comments and suggestions.

5. REFERENCES

Barbour, M.T., Gerritsen, J., Snyder, B.D. and Stribling, J.B., Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C., 1999.

Ceccaroni, L., Cortés U., and Sànchez-Marrè, M., OntoWEDSS: augmenting environmental decision-support systems with ontologies. *Environmental Modelling & Software, in press*, 2004.

Chang, N. B., Chen, Y. L. and Chen, H.W., A fuzzy regression analysis for the construction cost estimation of wastewater treatment plants. *Journal of Environmental Science and Health. Part A - Environmental Science and Engineering and Toxic and Hazardous Substance Control*, *32*(4), 885-899, 1997.

Comas, J., Llorens E., Martí, E., Puig, M. A., Riera, J. L., Sabater, F. and Poch, M., Knowledge acquisition in the STREAMES project: the key process in the Environmental Decision Support System development*, in AI Communications 16*, 2003.

Comas, J., Llorens, E., Poch, M., Markakis, G., Battin, T., Gafny, S., Maneux, E., Martí, E., Morais, M., Puig, M. A., Pusch, M., Riera, J. L., Sabater, F., Solimini, A. G. and Vervier, P., The STREAMES Project: Linking Heuristic And Empirical Knowledge Into An Expert System To Assess Stream Managers*. 1st International Conference on Environmental Modelling and Software (iEMSs2002)*, *3*, 444-449, 2002.

Davis, J. R., Farley, T. F. N., Young, W. J., and Cuddy, S. M., The experience of using a decision support system for nutrient management in Australia, *Water Science and Technology*, *37*(3), 209-216, 1998.

Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy, *Official Journal L327*, *2000-10-23*, *P0001*.

Great Lakes Commission for the Great Lakes States and Provinces. *Toward a Water Resources Management Decision Support System for the Great Lakes-St. Lawrence River Basin*. Great Lakes Commission, Eisenhower Corporate Park, 2805 S. Industrial Hwy, Suite 100, Ann Arbor, MI 48104-6791, 2003. The full report is available online at [http://www.glc.org/wateruse/wrmdss].

De Marchi, C. , Ivanov, P., Jolma, A., Masliev, I., Smith, M. G. and Somlyody, L., Innovative tools for water quality management and policy analysis: DESERT and STREAMPLAN. *Water Science and Technology*, *40*(10), 103-110, 1999.

Matthies, M., Berlekamp, J., Lautenbach, S., Graf, N. and Reimer, S., Decision Support System for the Elbe River Water Quality Management. In Proceedings of the *International Congress on Modelling and Simulation (MODSIM 2003)*, Townsville, Australia. 2003.

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

Rousseau, A. N., Mailhot, A., Turcotte, R., Duchemin, M., Blanchette, C., Roux, M., Etong, N., Dupont, J., and Villeneuve, J. -P., GIBSI – An integrated modelling system prototype for river basin management. *Hydrobiologia* 422/423: 465–475, 2000.

Poch, M., Comas, J., Rodríguez-Roda, I., Sànchez-Marrè, M. and U. Cortés, *Ten Years of Experience in Designing and Building Real Environmental Decision Support Systems. What Have We Learnt?, 1st International Conference on Environmental Modelling and Software (iEMSs2002)*, *3*, 22-33, 2002.