2

The Development of Computer-based Watershed Management Systems

D.C.L. Lam[1], G.S. Bowen[2] and C.I. Mayfield[3]

[1]*National Water Research Institute, P.O. Box 5050, Burlington, Ont., Canada L7R 4A6 (phone: 905-336-4916; fax: 905-336-4582; E-mail: David.Lam@cciw.ca)*
[2]*Science and Technology Branch, Ontario Ministry of the Environment and Energy, 135 St. Clair Ave. W., Toronto, Ont., Canada M4V 1P5 (phone: 416-323-5007; fax: 416-323-5031; E-mail: boweng@gov.on.ca)*
[3]*Biology Department, University of Waterloo, Waterloo, Ont., Canada N2L 3G1 (phone: 519-888-4640; fax: 519-746-2543; E-mail: mayfield@sciborg.uwaterloo.ca)*

Abstract

A knowledge-based approach is proposed for the development of environmental decision support systems for watershed management. A generic system called RAISON was developed for the Windows system, originally for the acid rain problems and now for application in watershed management problems using expert system, neural network, genetic algorithms and other information technologies. The emphasis is on the teamwork between scientists, economists and policy advisors to meet the challenge of the paradigm shift due to the advancement of computer technologies.

Keywords

Knowledge-based system, decision support system, watershed management

INTRODUCTION

Environmental problems are complex and involve the knowledge of many scientific disciplines. Solutions to these problems require sound decisions with a good understanding not only of the science but also the socioeconomic associated with these issues. On the one hand, one must integrate the information of the various scientific disciplines such as physics, chemistry and ecology. On the other hand, the economic analysis of pollution abatement and the social impact are central to policy decisions. Mathematical modeling techniques and data analysis have been used to integrate data and process knowledge. With the advent of the computer technology particularly in the environmental software areas, these integration techniques have been made easier to implement and more user-friendly to apply. The purpose of this paper is to look at

Environmental Software Systems Vol.2 R. Denzer, D.A. Swayne & G. Schimak (Eds.)
 Published by Chapman & Hall

these computer software technologies, their present applications and possible future potential of development for environmental decision support systems. In particular, we want to emphasize the "soft" engineering technologies, such as expert systems (Lam et al., 1989), as part of a paradigm shift in computer modelling. The paradigm shift arises from bringing the science and economics closer to the decision makers, especially when computer technologies evolve in such a fast pace from mainframe, to desktops and to Internet computer systems. We will use an example of an environmental problem, namely the watershed acidification (acid rain) problem to illustrate these new concepts.

PARADIGM SHIFT IN ENVIRONMENTAL DECISION SUPPORT SOFTWARE

The development of environmental decision support software depends on new findings of scientific investigations and on economic analysis that are reactive to new social demand and political trends. It requires teamwork. The notion of having an integration team to combine data, models and knowledge to work toward solving environmental problems using a decision support system has been around for some time. The main difficulty in the past has been that of communication. This was especially true ten years ago when many worked with mainframe computers that offered excellent computational facilities but poor graphical capabilities. In recent years, however, the advancement in microcomputer technologies for environmental application (Lam et al., 1994) including those for workstation computers has changed many aspects of graphical presentation. Not only can data and model results be presented in colour and various formats, they can be shown in motion with textual and audio-visual explanation. The manipulation of documents, maps, sound, photos, video and other forms of information or objects is much easier with microcomputers. New programming languages (e.g., C, C++ and Visual Basic) emerge and can handle various types of objects (Lam et al., 1994). This object-oriented programming ability enhances the communication in the solution team. Recently, the Internet has provided another impetus to new forms of communication. The idea of shared information and data within a large community with virtually no geographical boundaries is now available. Over a short time span of about ten years, the difficulties with mainframe computing are overcome with new technologies that are easy to use and accessible by many including the public. Time is due for a paradigm shift for the design and implementation of environmental decision support systems.

However, we need a new conceptual framework and methodologies to facilitate the communication between scientists and policy advisors. The data, knowledge and models from the scientists should be made more understandable to policy advisors, particularly the non-numeric and descriptive knowledge. For example, there are already many existing water quality models, but only few critical reviews on their strengths and weaknesses are available. The knowledge on model limitations and usage resides with those scientists who developed the models or others who used them frequently. Similarly, scientists who did the research know best the meta-data or descriptions about their data. We need a new conceptual framework to formalize and incorporate this important but largely descriptive body of knowledge, in addition to the numeric data and information, for designing environmental decision support systems.

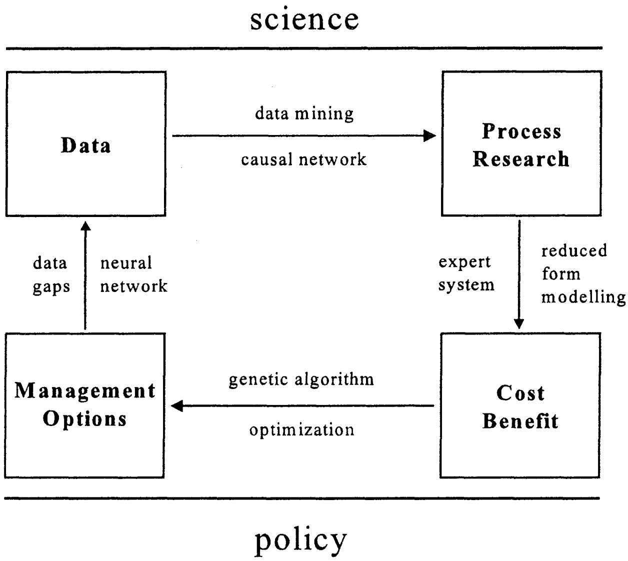

Figure 1 A knowledge-based approach to decision support for science and policy

For example, a simple approach is shown in Figure 1. We let data and process research representing science, and cost benefit and management options indicating policy. The linkages among these components can be investigated by using knowledge-based tools. For instance, from observational data, we can improve our knowledge on processes by data mining methods (Lam et al., 1994) such as neural network to find out relationships, if any, among the variables. From process knowledge, we can develop models to simulate various processes, design expert system (Lam et al. 1989) to choose appropriate models or reduce the model complexity by using input and output relationships (Olson et al., 1983). The economic models including cost benefit analysis can be tied to management options through optimization procedures such as genetic algorithms (Goldberg, 1992). Finally, to complete the iterative path, scenario testing of management options can lead to identification of data gaps for further investigation. This simple conceptual framework can be expanded and adapted to different environmental problems. The following example serves to explain further this knowledge-based system approach.

INTEGRATED ASSESSMENT FOR WATERSHED ACIDIFICATION PROBLEMS

The so-called acid rain phenomenon pertains to the industrial release of sulphur and nitrogen oxides into the atmosphere which travel long distance and return in the form of acidified deposition to the ground and into streams and lakes. Emission control strategies have been in place in North America and Europe and are now considered seriously in Asia. The cost and social impact of these control strategies are estimated to be in billions of dollars, while environment damages must be reduced to an acceptable level. This is the essence of the decision process. A decision support system for the acid rain problem therefore must consider the balance between economy and environment. For example, we can start with the economic production model, e.g., for energy consumption. The model calculates sulphur dioxide emissions, based on production activities and pollution control costs. Long range transport models for the prediction of the deposition of sulphur dioxide at given receptors can be computed by using the sulphur emission data as input. From the output of the atmospheric transport model, the acidified deposition is in turn used as input to the terrestrial and aquatic chemistry models. The output of the soil and water chemistry model is then fed into the ecological impact model to determine the damages made to the fish, wildlife, trees and other ecosystems. This linked modelling approach forms the basis for developing the integrated assessment model for acid rain problems (Lam et al., 1989; Lam et al. 1997).

The integrated assessment modelling approach grows out from the necessity of combining knowledge of the air, water and soil sciences with social and economical aspects. It has been adopted by the UN-ECE for use in many countries (Amann, 1990) for environmental assessment. The knowledge-based approach (Lam et al., 1994; Lam et al., 1997) adds to the integrated assessment modelling by providing an open architecture framework that accepts any component model and links them with knowledge-based tools as explained in Figure 1. Thus, it seeks optimal solutions for emission reduction for policy makers, given an acceptable ecological damage cost (upper loop, Figure 2). It tests economic production scenarios and analyzes the ecological damages and associated uncertainties (lower loop, Figure 2). The following are some examples of the application of knowledge-based linking tools.

Reduced-form modelling

Mathematical models form an integral part of many decision support systems. Some environmental problems require complex models to explain the complicated processes. However, these complex models are not suitable for application in a decision support system as they may require lengthy computational time and large input databases. One alternative is to run these models many times to cover sufficient temporal ranges and spatial domains and to relate the input and output with simple statistical relationships identified by neural network or other statistical methods. For example, trajectory model results of sulphur dioxide can be reduced to a source-receptor relationship for 40 source regions and 15 receptors in North America (Olson et al., 1983).

Expert system

In some cases, there are often too many models to choose from. Some models are better than others for a certain input or output range, because model assumptions may satisfy a given data better. The use of this modelling knowledge in expert system is proposed in Lam et al. (1989) to select and combine model results for acid rain impact analysis. The rule set first used in this application was a crisp set, i.e., each rule is based on precise values of the attributes. For example, one of the rules says if the TD

(Trickle Down) Model result is greater or equal to 0.1 meq/l and the CDR (Cation Denudation Rate) Model result is less than 0.1 meq/l, then the expert system result is the average of the two model results. However, one can also use a smooth continuous membership function to describe condition before and after the critical value of 0.1 meq/l (as compared to a sharp discontinuity in the crisp rule set). Different membership functions can be used to produce a value that satisfies the rules but allows for a smooth transition of outcome in the neighbourhood of the critical threshold (Wong et al., 1995).

Neural network

Neural network techniques are useful for those environmental problems where a priori knowledge is absent and must require a search for relationships among variables. For example, the relationship between dissolved organic carbon (DOC) and other water quality variables such as colour, aluminum, iron and ammonia could be derived by training a neural network (Lam et al., 1994), using existing data, based on a structure with four input layers, two hidden layers and one output layer. After the training was completed, the network was used to predict DOC. Since DOC was not often measured, while the four water variables were, the neural network was used to generate estimated values of DOC, thereby expanding the database. When compared to observed alkalinity data in four aggregates totalling 568 sites in Eastern Canada, the original crisp rule-base prediction produced an aggregate median of relative errors from 15% to 18.7%; the fuzzy expert system without the neural network input yielded 14.7% to 18.2%; and the expert system with fuzzy logic and neural network input gave 14.3% to 15.2%. Thus, the hybrid neural network and expert system approach led to smaller errors and better fit with observed data.

Opimization and genetic algorithm

The integrated assessment model can be used to derive optimal emission control strategies by using the concept of "critical load". A critical load is the level of wet sulphate deposition at a given location that will cause certain degree of ecological damage. For example, at the receptor site at Kejimkujik Park near Halifax, Nova Scotia, Canada, the critical load for a 24.2% lake damage (i.e. below a pH threshold of 6 for clearwater lakes) is about 13.5 kg/ha/y of sulphate deposition. This is slightly below the current deposition level of about 14.5 kg/ha/y, and can only be reached by reducing the emission from somewhere. To achieve this critical load, we can minimize the emission reduction (see upper loop, Figure 2), by searching through all source regions, every one of which is equally probable but each is allowed to cut by up to, say, half of its emission, as an example for economic considerations. The optimal solution to this question can be obtained by either linear programming or genetic algorithm (Goldberg, 1992) procedures. The computational time required for genetic algorithms was longer, but they appeared to have the advantage of avoiding problems with local minima and to have many desirable features of global search methods. The optimal solution calls for a reduction of about 26% in the Ohio source region, 22% for Pennsylvania, with smaller reduction in other source regions.

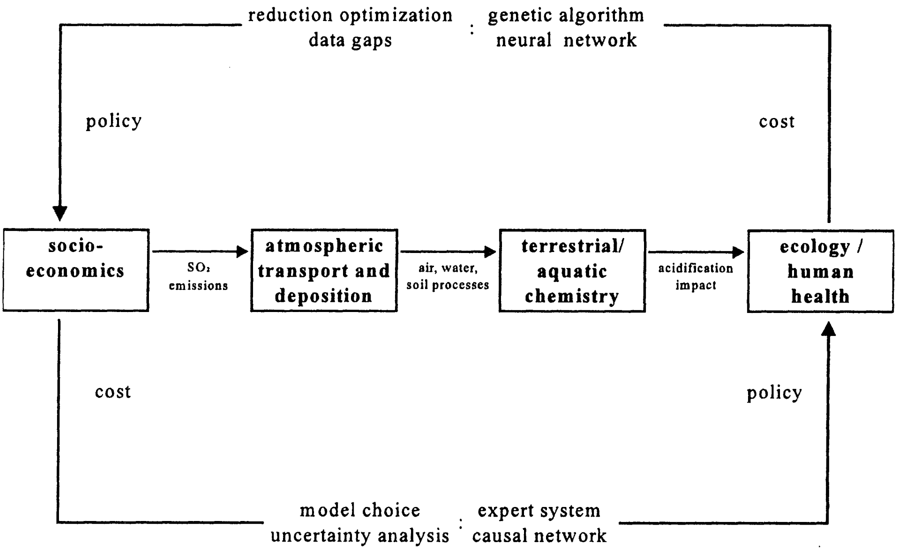

Figure 2 A schematic showing the knowledge-based software design for the acid rain problem

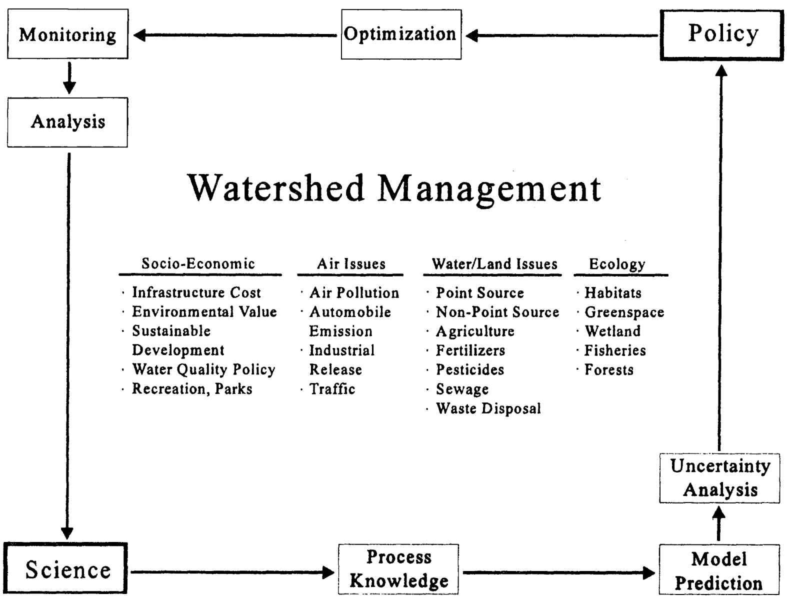

Figure 3 An iterative approach between science and policy for watershed management

Uncertainties and Belief Causal Network

It is difficult to quantify the uncertainties with linked model results as discussed here. A qualitative check on model uncertainty, however, can be obtained through the causal or belief network technique, e.g., to track the propagation of uncertainties in the source-receptor matrix model. Statistical errors and standard deviations derived during the computation of trajectories were obtained and translated into a simple high and low ranking of uncertainty, for such factors as source-receptor input (I), interpolation methods (K) and internal variability (V). For the receptor site at Kejimkujik, Nova Scotia, as an example, with five main source regions considered in Canada, five in the U.S. and one for other areas. The probability for high uncertainty was in the range between 10% to 30% for the variables I, V and K. We combined probabilities by using the relation P(A|B).P(B) = P(A,B), where P(A,B) is the joint probability of events A and B. Thus, the uncertainties for I, V and K were combined into an aggregate probability of about 26 to 30% for high uncertainty to be found in the model results from these source regions to the receptor site at Kejimkujik (Lam et al., 1997).

GENERIC SOFTWARE FOR OTHER WATERSHED MANAGEMENT PROBLEMS

Over the past decade, we have developed an environmental information system for regional analysis by intelligent systems on microcomputers (acronymed RAISON), for both DOS and Windows platforms (Lam et al., 1994). It offers a generic framework to integrate data, text, maps, photos, video and other input and to provide the user with a library of software tools, including models, optimization procedures, expert systems and neural network, to produce customized interfaces that generate output on data interpretation, scenario tests, strategic analysis and policy recommendation.

The RAISON system is developed primarily based on the watershed acidification study as explained above. However, instead of developing it as a specific application, we have modularized it into generic components so that it can be applied to other problems, such as watershed management applications. Figure 3 shows some current studies on watershed management problems, with more complex requirements for socio-economic issues such as infrastructure costs, recreation and parks, as well as more facets of air, water, land and ecological issues than acid rain problems. Again, one can design a watershed management decision support system with the knowledge-based approach as explained for Figure 1. It is iterative and starts with monitoring and research data (Figure 3) to improve scientific knowledge, which is used to build models and to quantify uncertainties for recommending policy options. The data and knowledge gaps can be fed back to the monitoring program to complete the iterative cycle (Figure 3). However, these paths can happen in parallel steps. Some examples will be presented at this conference and can be found in this proceedings (Booty et al., 1997; Young et al., 1997; Leon et al., 1997).

CONCLUSIONS

The knowledge-based approach to decision support system development is a viable method to respond to the paradigm shift in environmental software technologies. The approach can be implemented as a specific application or as a generic system for many environmental applications. Further studies are required to prove its effectiveness.

Acknowledgements. Thanks are due to the help from D.A. Swayne, I. Wong, W.G. Booty, D. Kay, P. Fong, K. Brown, A. Storey and other partners in this study.

REFERENCES

Amann, M. (1990). Energy use, emissions and abatement costs in Europe. In: J. Acamo et al.(e.ds.), *The RAINS Model of Acidification, Science and Strategies in Europe*. Kluwer Academic Publ., Dordrecht, the Netherlands.

Booty, W.G., Lam, D.C.L., Tseng, T. and Siconolfi, P. (1997) Great Lakes toxic chemicals decision support system, this proceedings.

Goldberg, D.E. 1989. Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA.

Lam, D.C.L., Swayne, D.A., Storey, J. and Fraser, A.S. (1989) Watershed acidification using the knowledge-based approach, *Ecol. Modelling* **47**, 131-152.

Lam, D.C.L., C.I. Mayfield, D.A. Swayne and K. Hopkins, 1994a. A prototype information system for watershed management and planning. *J. Biological Sys.*, 2(4):499-517.

Lam, D.C.L., Puckett, K.J., Wong, I., Moran, M.D., Fenech, G., Jeffries, D.S., Olson, M.P., Whelpdale, D.M., McNicol, D., Mariam, Y.K.G. and Minns, K. (1997) An integrated acid rain assessment model for Canada using environmental monitoring data, to appear.

Leon, F.L., Lam, D.C.L., Swayne, D.A. and Farquhar, G.J. and Soulis, E.D. (1997) Integration of a nonpoint source pollution model with a decision support system, this proceedings.

Olson, M.P., E.C. Voldner and K.K. Oikawa, 1983: Transfer matrices from the AES-LRT model. *Atmosphere-Ocean*, 21, 344-361.

Wong, I., Lam, D.C.L., Fong, P., Storey, A. and Swayne, D.A. (1995) A neuro-fuzzy system for acidic deposition modelling. *Proc. Int. ICSC Symp. Fuzzy Logic*, (Ed. by N.C. Steele), ICSC Academic Press, Zurich, Switzerland, **C**: 136-141.

Young, W.J., Lam, D.C.L., Ressel, V. and Wong, I. (1997) Development of an environmental flows decision support system, this proceedings.

Biography

Dr. David Lam is the Senior Scientist and Project Chief, Environmental Information Technology at the National Water Research Institute, Environment Canada, Burlington.

Gary Bowen is a Scientist with the Watershed Management Section, Science and Technology Branch, Ontario Ministry of the Environment, in Toronto, Canada.

Dr. Colin Mayfield is a Professor in the Biology Department, University of Waterloo.