AN EXPERT SYSTEM FOR WATER QUALITY MODELLING

W. G. BOOTY, D. C. L. LAM, A. G. BOBBA, I. WONG, D. KAY, J. P. KERBY, and G. S. BOWEN*

National Water Research Institute, Rivers Research Branch, Environment Canada, Burlington, Ontario, Canada.

**Water Resources Branch, Ministry of the Environment, Toronto, Canada*

(Received May 1991)

Abstract. The RAISON-micro (Regional Analysis by Intelligent System ON a micro-computer) expert system is being used to predict the effects of mine effluents on receiving waters in Ontario. The potential of this system to assist regulatory agencies and mining industries to define more acceptable effluent limits was shown in an initial study. This system has been further developed so that the expert system helps the model user choose the most appropriate model for a particular application from a hierarchy of models. The system currently contains seven models which range from steady state to time dependent models, for both conservative and nonconservative substances in rivers and lakes. The menu driven expert system prompts the model user for information such as the nature of the receiving water system, the type of effluent being considered, and the range of background data available for use as input to the models. The system can also be used to determine the nature of the environmental conditions at the site which are not available in the textual information database, such as the components of river flow. Applications of the water quality expert system are presented for representative mine sites in the Timmins area of Ontario.

1. Introduction

Models can be used to predict the characteristics of water quality conditions in aquatic systems in order to ensure that water quality objectives will be maintained under a wide variety of conditions. Models provide the ability to develop a credible and defensible water quality management program. They are continually being developed and upgraded to optimize the often competing demands of regulation, environmental protection, and efficacy and cost control measures. There are a large number of available models in the literature (Booty and Lam, 1990) which could potentially be used for water quality and wasteload allocation predictions. However, at this stage in the development of the expert system for water quality modelling, the hierarchy of models has been limited to seven models, several of which are used by the U.S. EPA (U.S. EPA, 1985). These models are used to make predictions about the concentrations of contaminants in the mixing zones as well as the areas of complete mixing of receiving waters. The system has initially been developed for application to mine effluents in Ontario as part of the MISA (Municipal Industrial Strategy for Abatement). The overall goal of the development of the system is to provide a mechanism by which a broad range of users can evaluate a wide spectrum of contaminants and aquatic systems with respect to the overall goal of sustainable development of our natural resources.

This paper presents an overview of the Mine Effluent Model Choice Advisor which can be used within the RAISON MISA Mine Effluent Models for Receiving Waters System,

Environmental Monitoring and Assessment **23**: 1–18, 1992.

the model methodology, and examples of applications to mine sites in the Timmins region of Ontario, Canada.

2. Mine Effluent Model Choice Advisor

Most environmental assessment projects share similar characteristics of uncertain or incomplete data which varies in both time and space, as well as a wide variety of possible models which could be used to simulate the behaviour of an ecological system. This is also the case for the problem of assessing the effects of mine effluents on receiving waters, where it is necessary to apply the correct model based on the characteristics of the receiver and source, often for the situation where there is imprecise or incomplete data.

The motivation for the use of expert system technology in the assessment of mine effluent effects on receiving waters is traditional in the sense that what was required was a method to encode and capture expertise. Resource managers and policy makers can only benefit from a more complete and robust information system for decision support. The expert system presented provides such inexperienced users with a systematic and intelligent model choice system to assist in the impact assessment process. Also, throughout the development process, other ostensible applications for rule based technology became apparent. For example, it was discovered that rule bases would be suitable in this particular application to infer values for missing data items and test the integrity of incoming data. Further development in this area is part of future plans for the project. The choice of rule based technology proved to be beneficial at the time of development. The knowledge engineering process is inherently iterative. Close contact by the domain expert during expert system development ensures a complete and accurate representation of domain specific knowledge. For the given system there was a clear separation of the rules from the inferencing capability, and the simulation subsystem. Such a modular design allowed the knowledge base (rules) to evolve independently of development in other areas. The result was an uncompromised translation of relevant domain expertise. The model selection process is hierarchical in nature and decisions are based on the values of input parameters, model assumptions, the availability of input data, and the nature of the assessment required. The model selection process solicits information from both the user and the information system (the databases), and based on these values, a final model selection is made.

The typical model selection yields a single model choice, yet, based on certain characteristics, the process may yield a complex result, where the simulation requires different phases. For example, consider the following dialogue from the expert system advisor:

What are the characteristics of the receiving water?

* A small creek or swampy area with a poorly defined channel
* A stream or river with a well defined channel
* A lake or river with > 20 day residence

⇒ A stream or river with a well defined channel

Calculate distance to edge of mixing zone?

* yes
* no

⇒ yes

What is the model mode?

* Steady state
* NOT steady state

⇒ Steady state

Is the chemical type conservative?

* yes
* no

⇒ no

Are the contaminants involved characterized by major decay or removal from the aqueous system via adsorption onto and settling of suspended solids?

* yes
* no

⇒ yes

The model selection is:

model 4a + model 7a

The expert system prototype was implemented using a suite of rule bases. The basic inferencing technique was data driven or forward chained. It uses left-hand side rules only for 'Yes' and 'No' to fire the right-hand side. That is, given a set of parameter values for a scenario, the inference engine would fire a rule (or a series of rules) yielding a final model choice result.

Phase II of the mine effluent system involves the communication of the model choice directives to the simulation subsystem and the subsequent application of the selected model. In the original RAISON expert system the knowledge base rules were written in the

C language directly in the RAISON program. However, this approach limited the flexibility of the system. The inferencing capability of the new system was achieved by interfacing commercial expert system shells, namely 1st-CLASS (1st-CLASS, 1988), NEXPERT (NEXPERT, 1988), and CxPERT (CxPERT, 1990) to the RAISON system. After evaluating the three expert systems, CxPERT was found to be the most compatible system for use within the RAISON system. The RAISON system is a complete processing framework incorporating a spreadsheet, DBMS, GIS and a programming language. Basic handshaking and message passing techniques facilitated the necessary communication between the model choice advisor and the RAISON system. All the components of the system (the simulation subsystem, the front-end model choice expert advisor, and the mine information system) operate on a common data context with a common protocol. This design is intended to support future incremental growth of the system as required.

3. MISA Mine Effluent Models for Receiving Waters System

At this time there are seven models which have been chosen to simulate steady state and time dependent receiving water concentrations within, and downstream from the mixing zone. These models include: Model 1: Dilution Model; Model 2: Minimum dilution near field mixing zone model; Model 3: Ambient mixing zone model; Model 4: Ambient mixing zone with decay term model; Model 5: Lognormal probability distribution model; Model 6: Lake/Reservoir model; and Model 7: Fate/Adsorption model. The models are described in detail in Booty *et al.*, 1990.

3.1. USER/DATABASE INTERFACE

The screen interface of all the implemented models uses two distinct monitors; a colour monitor for graphics, and a monchrone screen, which is used primarily for data input and manipulation. Data input to the models can be performed by direct on-screen editing using the cursor or other related keyboard keys, on-screen windows, or by retrieving data from the database given a set of indices, usually name and date. For example, for Model 1, the model user will use a combination of manual data entry and database retrieval to obtain values for the model parameters. All mono screen fields can be entered and the values can be edited by the user, an example of which is shown in Figure 1 for Model 1. This example is for Cu where Qr is the receiving water flow (m^3/s), Cr is the receiving water concentration (mg/L), Qe is the effluent flow, and Ce is the effluent concentration. The data can also be entered by using the database retrieval utility. In this case the model requires flow and concentration data for a contaminant at a specific date. This data is retrieved by selecting the appropriate titles in the screen windows.

3.1.1. *Background Data*

A wide range of background data may be available for a particular mine and the receiving waters for its effluents. This data is stored within the RAISON system according to specific data types. For the MISA Mine Effluent Models for Receiving Water System, these data include:

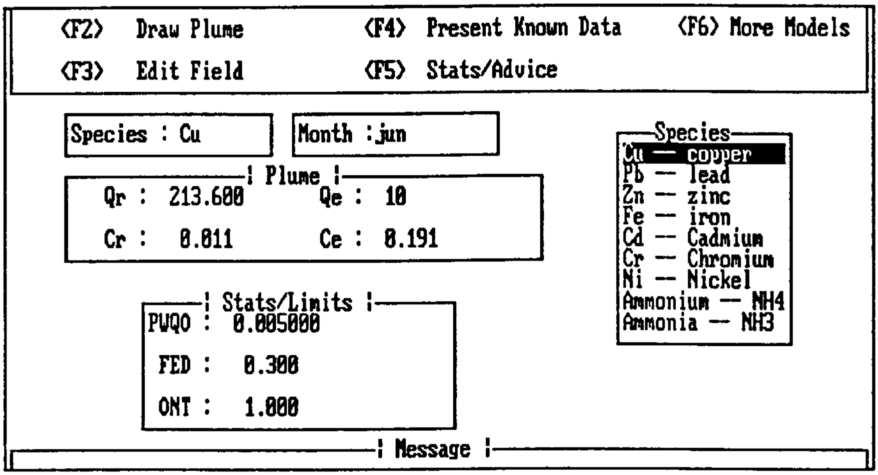


Fig. 1. Model 1 (Dilution Model) input screen.

(1) flow data (background flows, effluent flows, critical flow rates) for specific locations and dates;
(2) concentration data (background, effluent) for specific chemicals, locations and dates;
(3) temperature data (receiving water and effluent) for specific locations and dates);
(4) textual information (data qualifiers, mine operation information, physical characteristics of the discharge system and the receiving water system, etc.).

If flow data is not available for a particular river system, a hydrological model is available within the RAISON system (Bobba and Lam, 1984) which can be used to generate river flows using local meteorological data. In Figure 2 the model predicted versus measured flow data are presented for the Porcupine River for 1983, near Timmins, Ontario. The model can also be used to predict the significance of the various flow components which contribute to the total river flow, as shown in Figure 3.

3.1.2. *Guideline Data*

This set of data include the federal and provincial water quality guideline limits for specific chemicals. At this stage, the guideline limits and the observed data are stored for a particular mine. In the future, the guideline limits will become a separate entity and will be stored separately. The observed data will be extracted from the mine effluent database and compared with the guideline values.

3.1.3. *Parameter/Coefficient Data*

A number of the models require values for first order decay coefficients as well as adsorption coefficients. Values for these terms as a function of temperature, pH, and type of suspended solid, etc. are also available for specific chemicals within the database.

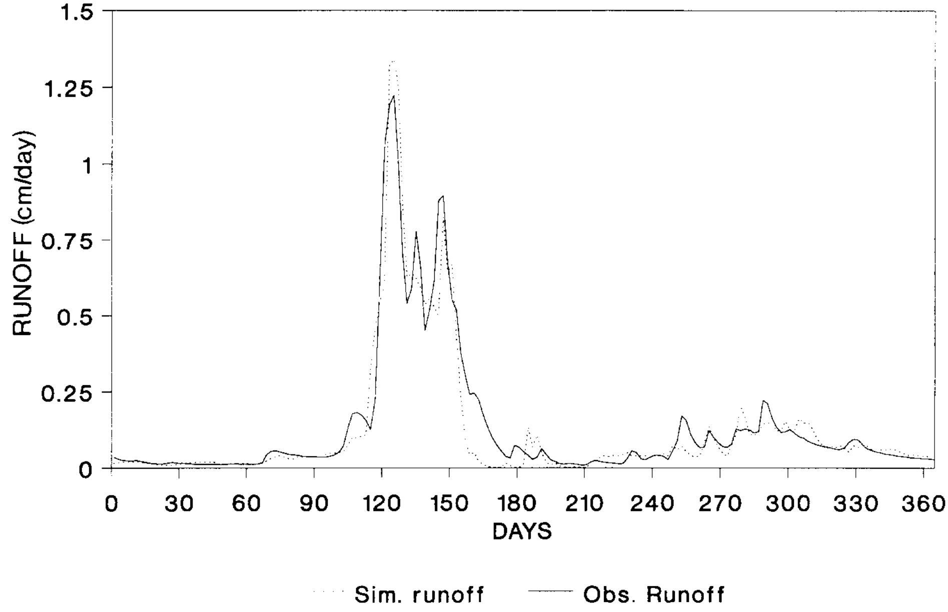


Fig. 2. Hydrological submodel simulated versus observed daily flows for the Porcupine River, 1983.

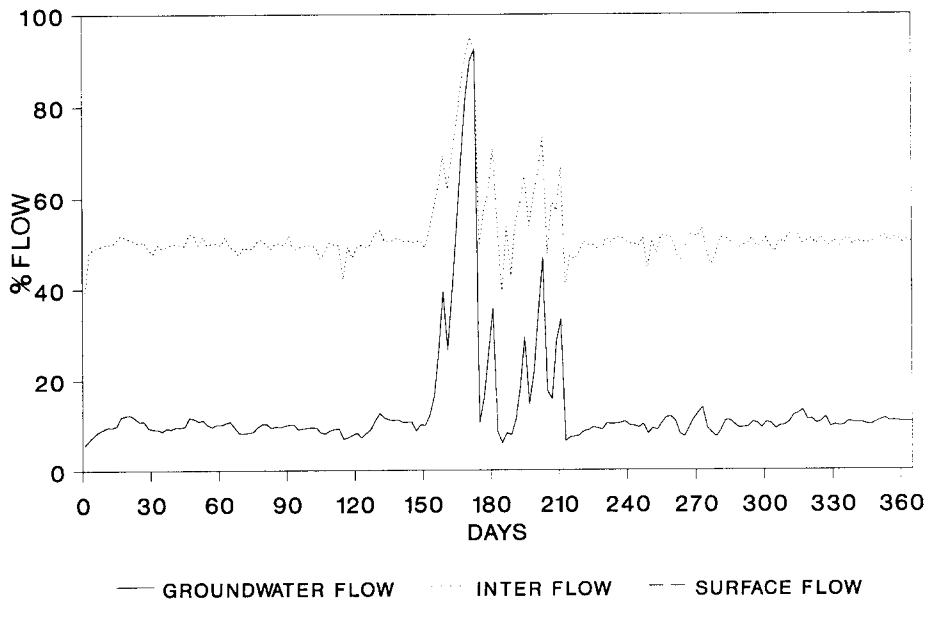


Fig. 3. Porcupine River contributing flow components for 1983.

4. Model Applications and Results Presentation

Model results can be displayed as numerical data on a spreadsheet, a graph, as colour coded sections of a plume in the mixing zone, or as colour coded river segments on a G.I.S. (Geopgraphical Information System) map. Examples of applications of the seven models within the MISA Mine Effluent Models system are presented for sites within the Timmins mining district of Ontario, Canada.

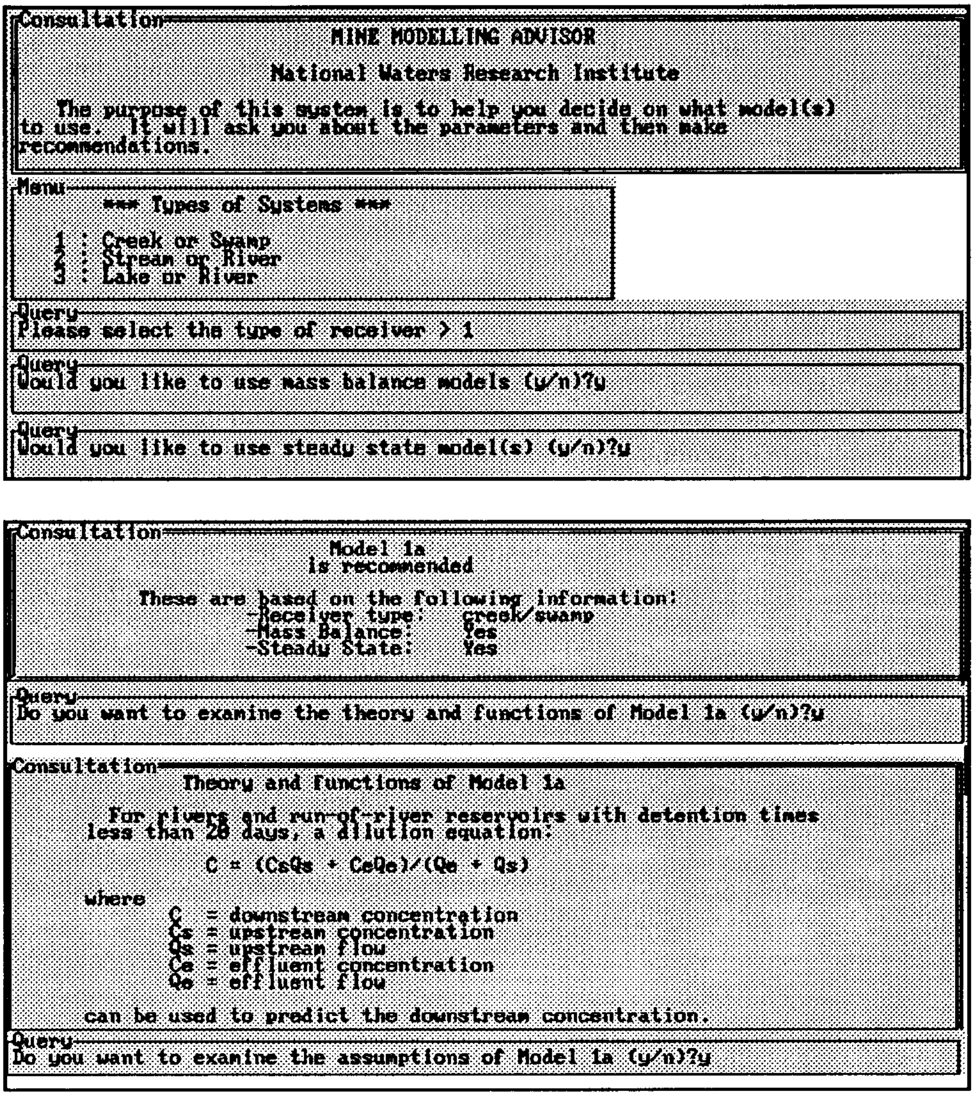
4.1. APPLICATION #1

In this example, a small 400 ton/day gold mine in the Timmins area is the subject of an environmental assessment of the impact of its wastes to a receiving water. The model choice advisor is initially used to select an appropriate model based on the physical and chemical data available for the mine and based upon the information required for the assessment. In this case the regulatory agency is interested in determining the number of times per year that the mine will exceed the $7Q_{20}$ limits for copper. The dialogue shown in Screens 1 and 2 is generated with the expert system advisor.

The results for Model 1 include the chemical name, month, effluent, downstream and upstream flows and concentrations and the percentage of the river flow required to bring concentration levels down to PWQO (Provincial Water Quality Objective) standards. Model results are presented in spreadsheet form and also in graphical form, as shown in Figure 4. In this example, the model predicted monthly Cu concentrations are presented for mean, 50 year and 100 year critical low flows. Using these model results, the mine operators may choose to only discharge effluent to the receiving water during the months of September–November. Alternatively, the model may be used to determine the monthly effluent discharges rates that would be required to meet the receiving water guidelines.

4.2. APPLICATION #2. MODEL 2 (MINIMUM MIXING ZONE MODEL)

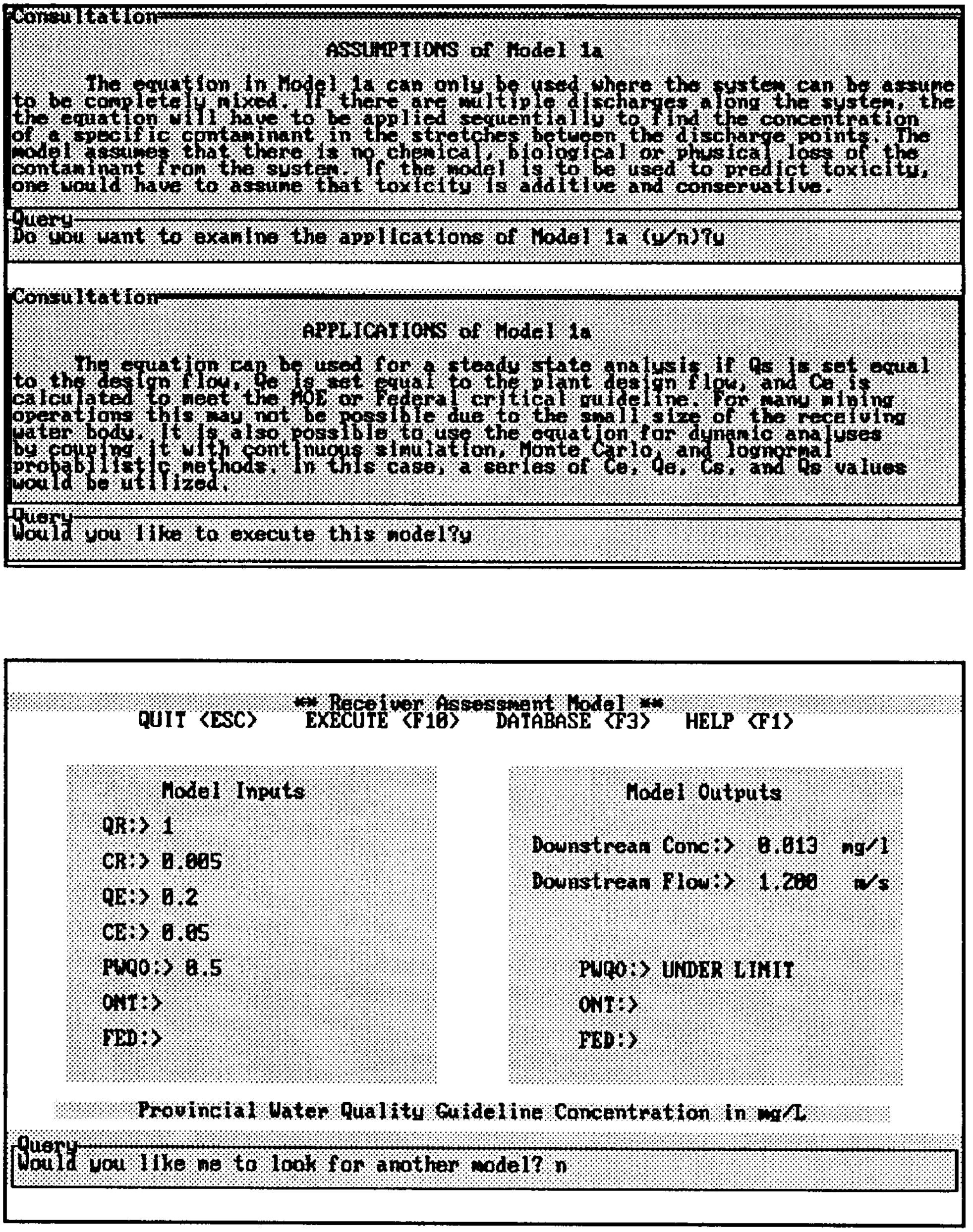
Model 2 is normally run in series with Model 3 due to the fact that the model requires the distance to the edge of the mixing zone, which is one of the outputs of Model 3. Other inputs required for Model 2 are the diameter of the effluent discharge pipe and the width of the river. The minimum mixing zone model is used to determine the minimum amount of mixing that can be achieved within the mixing zone due to discharge induced mixing. The position of the mixing zone plume is shown in Figure 5 along with the input data and calculated flow (QD) and concentration (CD) of Cu at the edge of the mixing zone. For this case, the distance to the edge of the mixing zone is 5.6 metres. The minimum flux induced mixing or dilution factor (S) is equal to 0.86. Results can also be generated in tabular form or in graphical form, as shown in Figure 6. In this example, the expected flux-averaged dilution factors for a fixed distance to the edge of the mixing zone of 35.0 metres and a river width of 5.0 metres is presented for varying discharge pipe diameters.



Screen 1.

4.3. APPLICATION # 3. MODEL 3 (AMBIENT MIXING ZONE MODEL)

The ambient mixing zone model is initially run to determine the distance required to achieve complete mixing. In this model it is assumed that the width of the zone that is mixed by discharge induced mixing is much smaller than the width of the river. It is also assumed that the effluent chemical is conservative and consequently the distance to achieve complete mixing is not a function of the contaminant but rather the physical conditions of the river system downstream from the effluent discharge point. An example output is



Screen 2.

shown in Figure 7 for various monthly critical flow scenarios. The concentrations at the edge of the mixing zone for these flow scenarios are compared with the PWQO guideline of 0.005 mg/L in Figure 8. The model is also capable of calculating the concentration at any point within the mixing zone plume. In Figure 9 the concentration of Cu at various

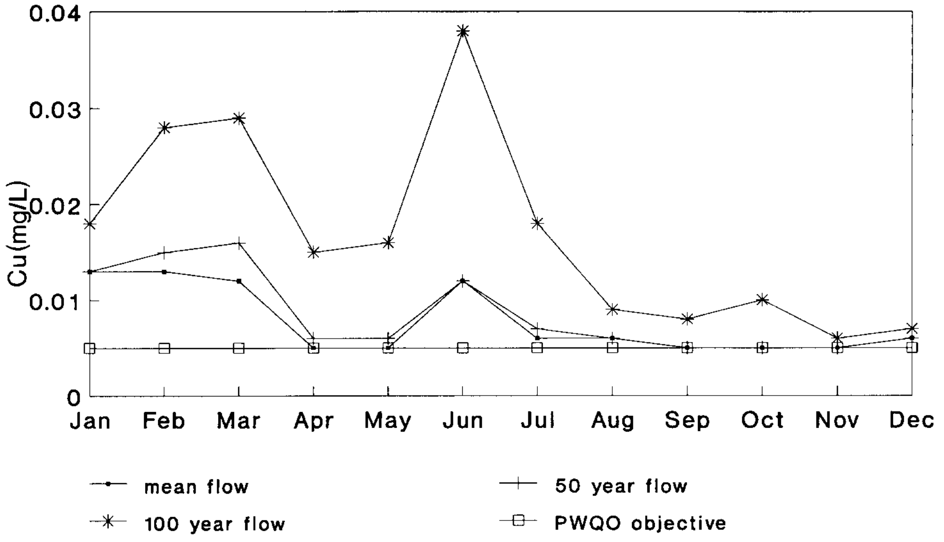


Fig. 4. Dilution model predictions of downstream monthly Cu concentration for various flow scenarios.

distances from the discharge point within the mixing zone are shown. In this example, it is assumed that Cu behaves as a conservative chemical, at least within the mixing zone. However, if this were not considered to be a valid assumption, then Model 4 would be indicated as the correct model to use by the model choice advisor.

4.4. Application #4. Model 4 (Ambient Mixing Zone Model with Decay Term)

In this model the effluent chemical is assumed to be non-conservative. Figure 10 shows the pop-up window that is used for input data entry. Besides the physical data for the river, the model also requires the pH of the receiving water since the decay term is a function of pH and temperature. The temperature values for the different months of the year are already entered in the database as are values for the decay coefficients for 10 common mine effluent chemicals. In Figure 11 the results of running the ambient mixing model with the decay term are shown for the same system shown in Figure 8.

4.5. Application #5. Model 5 (Lognormal Probability Distribution Model)

This model is useful for predicting the frequency and duration of contaminant concentrations in a river for which time series data is lacking. The model can only be used to predict the concentration of a substance for a zone of complete mixing and for the case where it can be assumed that no significant decay or transformation has occurred. The input data are the receiver flow and concentration and the effluent flow and concentration. The user is prompted for the contaminant and its database filename, and the effluent and upstream receiver flowrate database filenames. An example is shown in Figure 12 for Bell Creek total heavy metals data (Canamax, 1990).

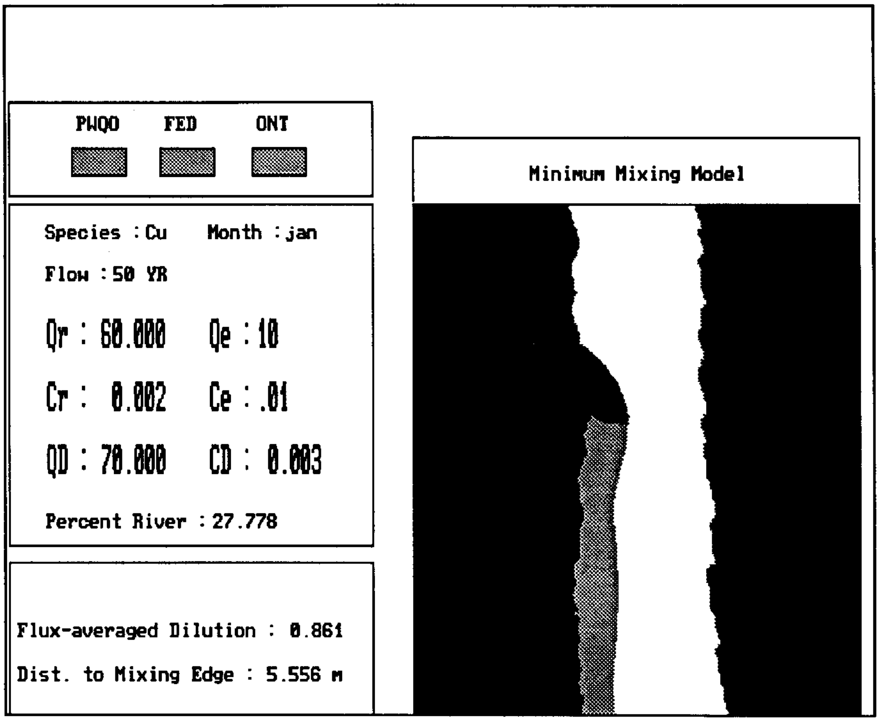


Fig. 5. Model 2 (minimum mixing zone model) graphics output screen.

4.6. Application #6. Model 6 (Lake/Reservoir Model)

This model is designed to determine the steady state concentration of a contaminant in a lake or reservoir with a residence time greater than 20 days. It is also assumed that the decay rate of the contaminant is constant. As shown in Figure 13, the model output includes a listing of all the model parameter values as well as a lake whose water is colour coded to show the concentration of the contaminant relative to the PWQO guideline. In this case, the concentration of Cu is less than 50% of the PWQO.

4.7. Application #7. Model 7 (Fate/Adsorption Model)

In this model it is assumed that adsorption onto and settling of suspended solids are the key processes controlling the fate of a metal in a stream or river. There are a significant number of parameters which must have values assigned to them. This can be done by manual entry or through database acquisition. The user must also select a river segment to analyze. This is performed by moving the cursor cross-hair to the segment and pressing the enter key (keyboard or mouse). Access to the Model 7 database is accomplished through a series of pop-up windows. Currently, the model parameter values are stored in their own database file with a range of values for each. Retrieval of the model parameter values requires the user to select either a high, low, or average value for the entire set. Once

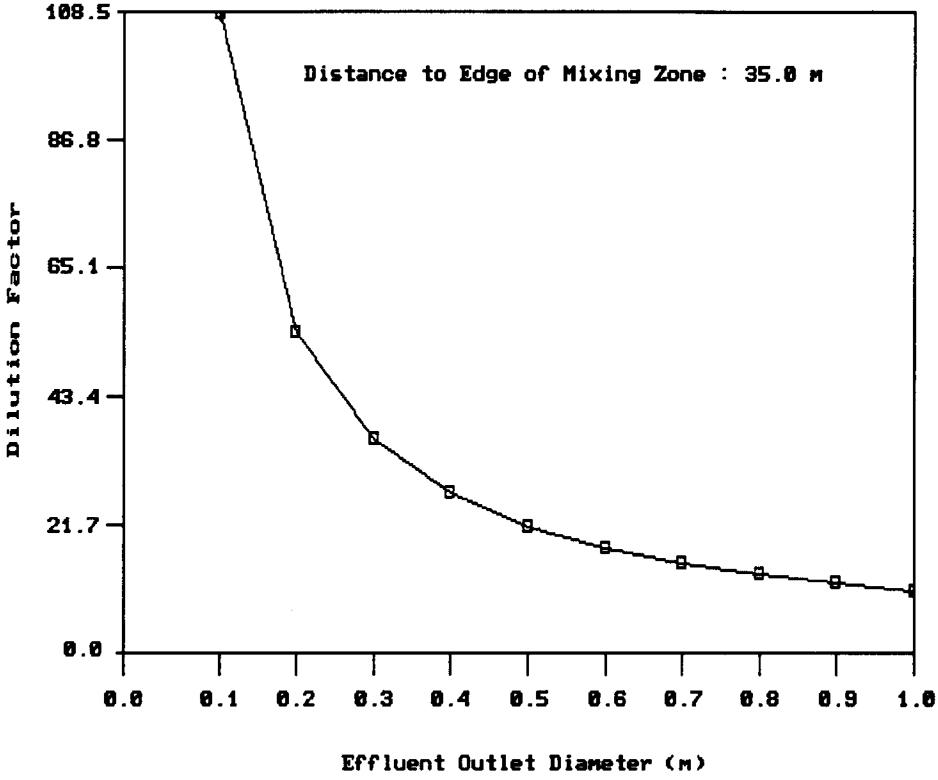


Fig. 6. Minimum mixing zone model predictions of the flux-averaged dilution factor for various discharge pipe diameters.

retrieved, the user can change any of the parameter values using the editing facilities. The graphics screen allows the user to view the concentration levels of the river segment chosen. This is accomplished by colour coding the segments for high, medium and low levels. An example of the graphics screen is shown in Figure 14. Spatial and time series graphs can also be produced as output as shown in Figures 15 and 16, respectively.

5. Discussion

The above examples are based on the limited data that were available during the startup of the MISA program. Due to the restricted amount and range of data available, it was not possible to use the results of the model tests as feedback to modify the model rules. This feedback process is a necessary part of the development of the system, but it must be based on results generated from a wide range of actual sites rather than artificial scenarios.

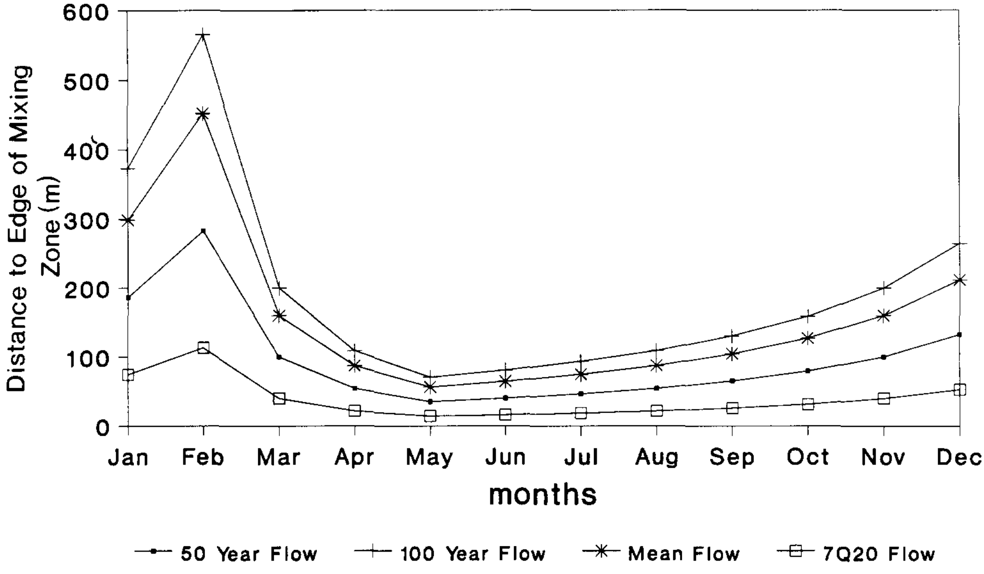


Fig. 7. Model 3 monthly predictions of distance to the edge of the mixing zone for various flow scenarios.

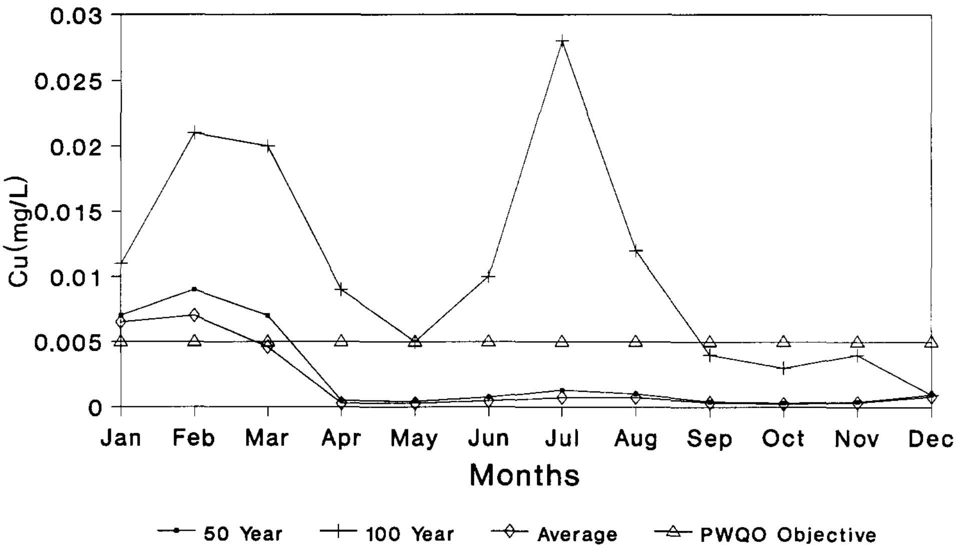


Fig. 8. Monthly concentrations of Cu at the edge of the mixing zone for various flow scenarios.

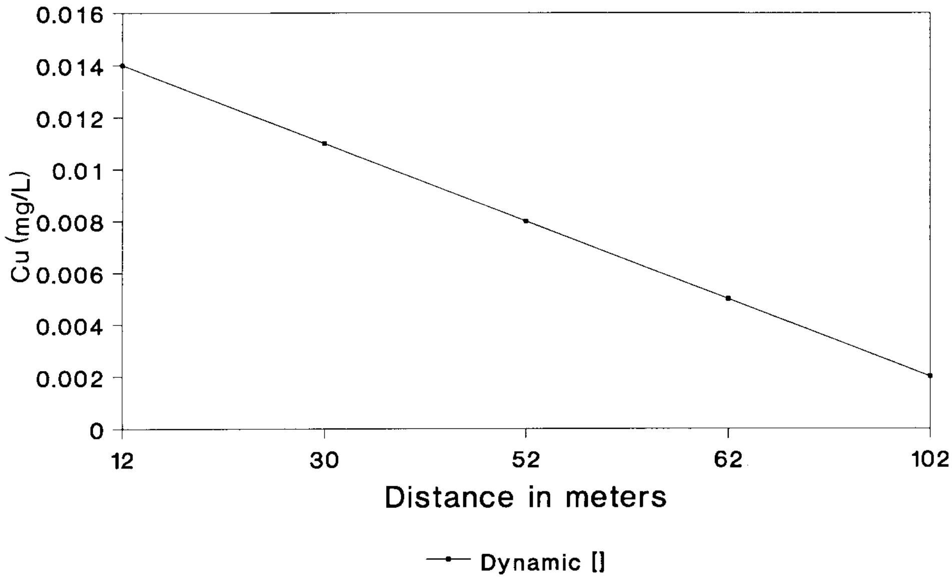


Fig. 9. Cu concentration as a function of distance from the discharge point.

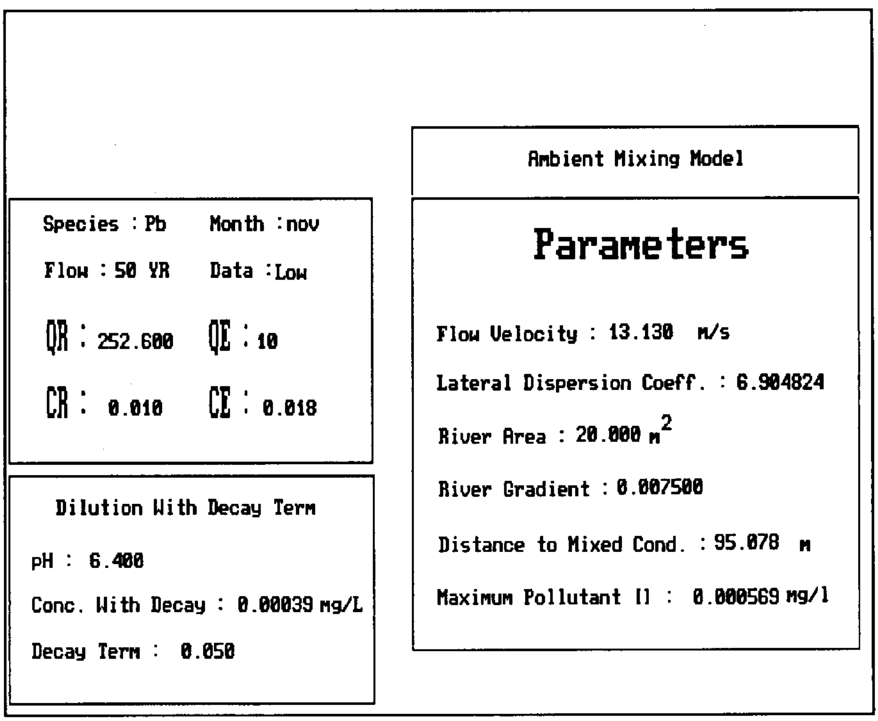


Fig. 10. Model 4 (ambient mixing zone model with decay) input screen.

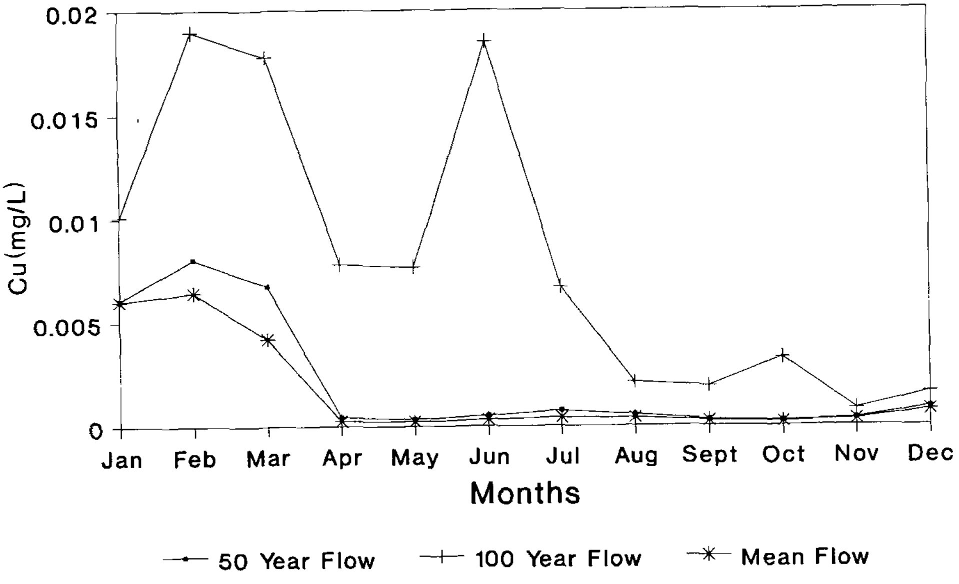


Fig. 11. Model 4 predictions of Cu concentrations at the edge of the mixing zone for various flow scenarios.

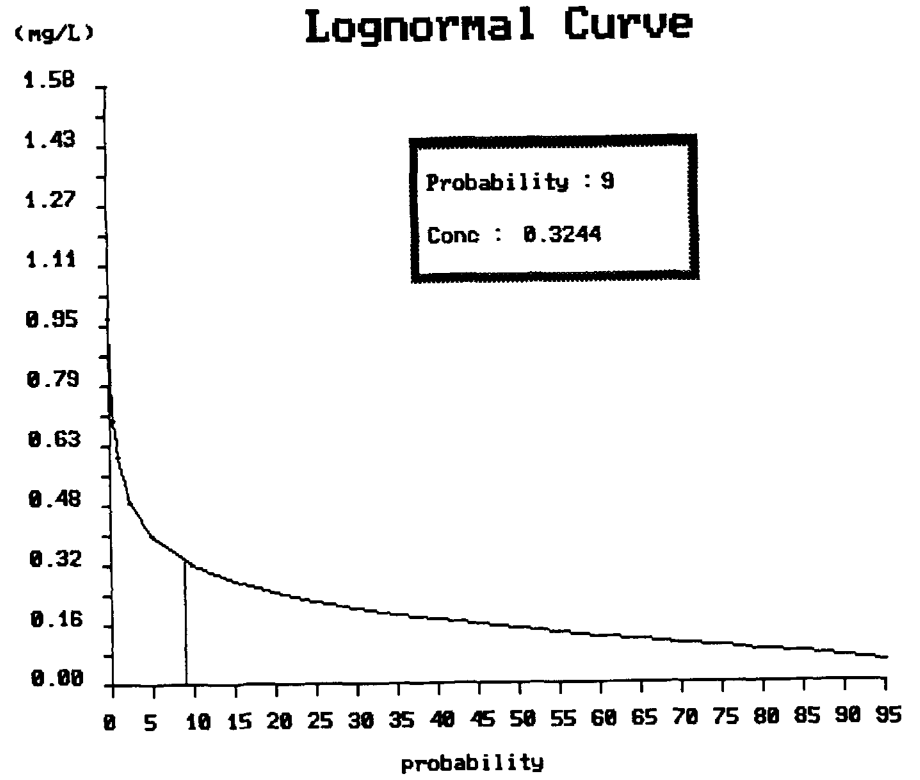


Fig. 12. Lognormal probability distribution of total heavy metals downstream of the Bell Creek Mine.

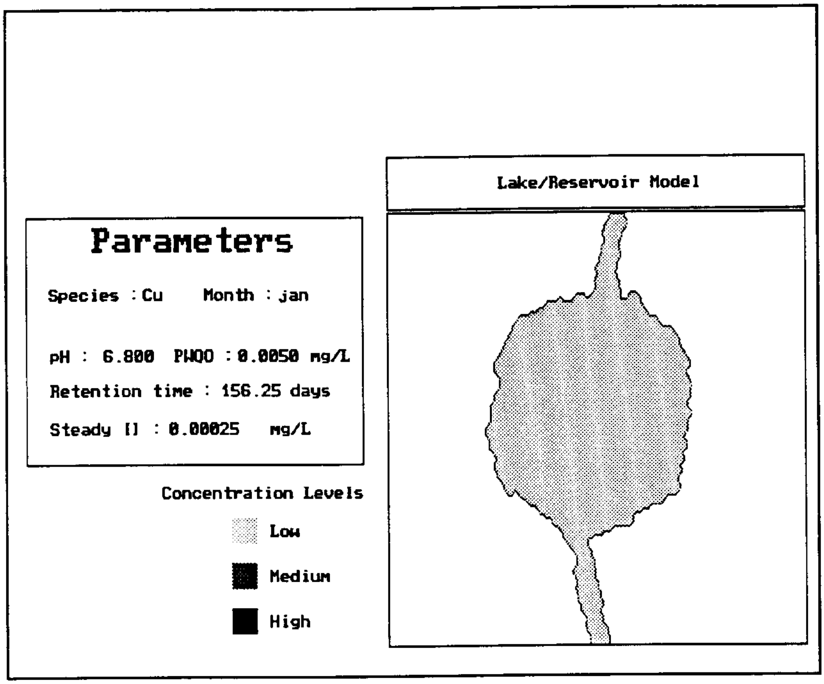


Fig. 13. Lake/Reservoir model predictions of Cu concentration with respect to the PWQO guideline.

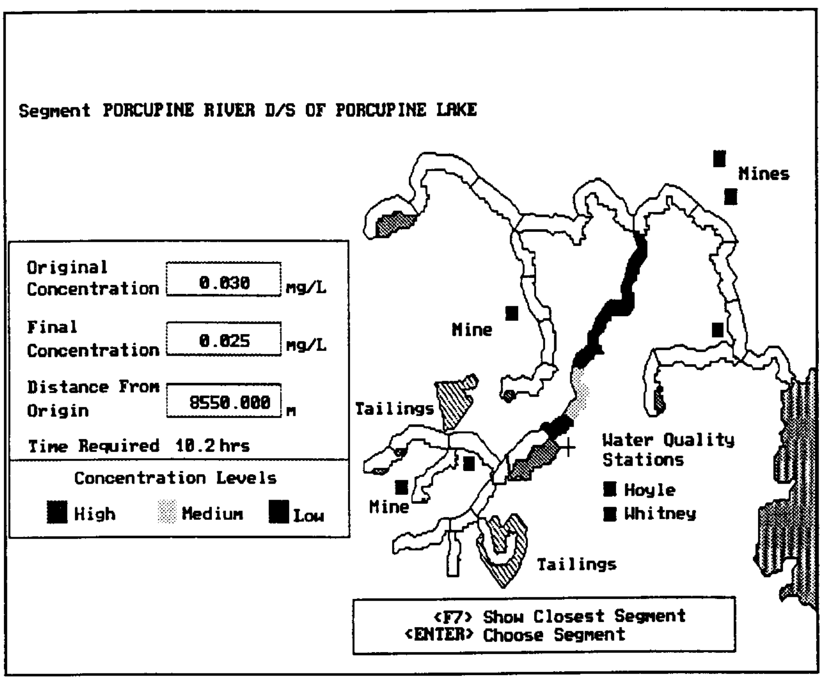


Fig. 14. Fate/adsorption model output graphics screen for the Porcupine River.

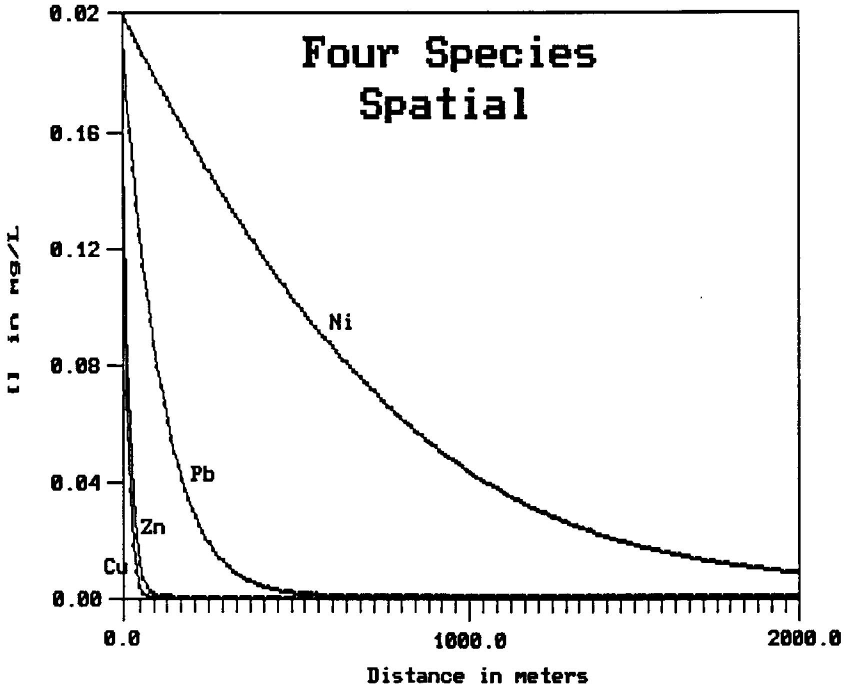


Fig. 15. Comparison of Cu, Pb, Zn, and Ni concentrations over a 2 km stretch of river.

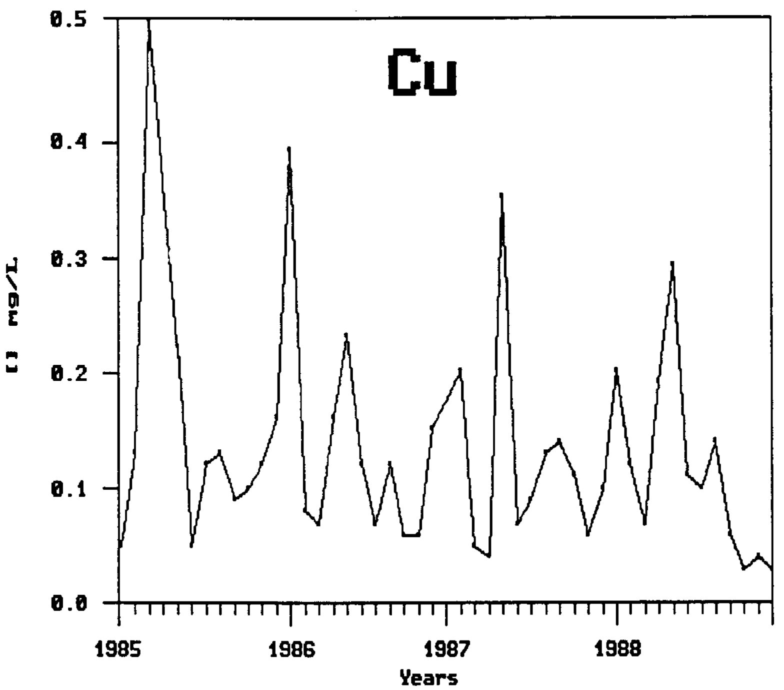


Fig. 16. Time series concentrations for Cu at a point 4500 meters downstream from the discharge point.

However, the examples presented demonstrate the range of applications that the system is currently designed to simulate.

6. Summary and Conclusions

It is important that computer models are user friendly in order that a wide range of users can easily utilize them. It also is equally important that the model users choose the best model for a particular application. To do this the model user must be familiar with the basic assumptions upon which the model is based. The preliminary model choice advisor which has been described in this paper is a convenient way of ensuring that a model user is aware of the basic assumptions and application limitations of available models. At this stage in the development of the RAISON Mine Model Expert System, insufficient data were available for calibrating the relatively simple models currently included. However, as the full spectrum of MISA data become available, more sophisticated models will be incorporated and tested along with the more advanced model choice advisor.

Future directions of the project include the development of a generalized RAISON expert system shell with the capabilities of both forward and backward chaining and knowledge representation through rule bases. Such a design will give the system developer the flexibility to use a combination of both conventional and expert system techniques to solve a problem without the limitations imposed by disjoint systems.

References

Bobba A G., and Lam, D. C. L.: 1984, 'Application of Linearly Distributed Surface Runoff Model for Watershed Acidification Problems', *Proc. Can Hydrol. Symp.* **1**(15), 381–398.

Booty, W. G., Lam, D. C. L., Wong, I., Kay, D., Kerby, J. P. and Bowen, G.: 1990, 'A Preliminary Study of MISA Mine Effluent Models for Receiving Waters', Report submitted to Water Resources Branch, Ontario Ministry of the Environment, Toronto, Ontario.

Booty, W. G., and Lam, D. C. L.: 1990, 'Freshwater Ecosystem Water Quality Modelling', In *Modelling Marine Systems, Vol. II.* A. Davies (Ed.), CRC Press, Inc. Boca Raton, p. 387–432.

Canamax Resources, Inc.: 1990, 'Operating Review of the Bell Creek Mill Tailings Waste Facilitiy for 1988 and 1989', A report submitted to Ontario MOE, March 6.

CxPERT: 1990, Version 3.0, Software Plus, Ltd., Crofton, MD, U.S.A.

1st-CLASS Expert System.: 1988, Version 3.70, 1st-CLASS Expert System Inc. Wayland, MA. U.S.A.

Lam, D. C. L., Swayne, D. A., Storey, J., Fraser, A. S. and Wong, I.: 1988, 'Regional Analysis of Watershed Acidification using the Expert Systems Approach', *Environ. Software* **3**(3), 127–134.

NEXPERT.: 1988, Version 1.1, Neuron Data Inc. Palo Alto CA. U.S.A.

Swayne, D. A. and Fraser, A. S.: 1986, 'Development of an Expert System/Intelligent Interface for Acid Rain Analysis', *Microcomputers in Civil Engineering* **1**, 19–22.

Wong, I., Lam, D. C. L., Bowen, G., and Novak, Z.: 1989, 'A Preliminary Application of the RAISON-micro System for the Development of Mine Effluent Assessment Methodologies', National Water Research Inst. Contribution 89–48.

U.S. Environmental Protection Agency.: 1985, 'Technical Support Document for Water Quality-Based Toxics Control', Office of Water Regulations and Standards, Washington, D.C.