

# **An Information Framework for Environmental Research and Management<sup>1</sup>**

*D.D. Cowan, P.M. Koch, C.I. Mayfield, D.A. Swayne*

*D.D. Cowan, Professor, Computer Science Department,  
University of Waterloo, Waterloo, Ontario, Canada,  
tel: (519)888-4690, fax: (519)746-5422,  
e-mail: dcowan@csg.uwaterloo.ca*

*C.I. Mayfield, Professor, Biology Department,  
University of Waterloo, Waterloo, Ontario, Canada,  
tel: (519)885-1211 x4640,  
e-mail: mayfield@sciborg.uwaterloo.ca*

*P.M. Koch, 18 Royal Hunt Crescent, Ottawa, Ontario, Canada,  
tel: (613)526-0404, fax: (613)526-0929*

*D.A. Swayne, Professor, Computing and Information Science,  
University of Guelph, Guelph, Ontario, Canada,  
tel: (519)824-4120 x3411, fax: (519) 837-0323,  
e-mail: dswayne@snowwhite.cis.uoguelph.ca.*

## **Abstract**

This paper describes the current status of an ongoing applied research project which is investigating design, implementation and utility of a large scale distributed environmental information system (EIS). This EIS is currently undergoing pilot-scale implementation. A structure for an environmental framework and an accompanying high-level architecture for the EIS are presented. A number of related outstanding Computer Science issues are discussed, and the paper concludes with a brief description of current progress.

---

<sup>1</sup> The work described here has been supported by the National Water Research Institute of Environment Canada, Natural Sciences and Engineering Research Council, IBM Canada, and the Tri-Council Environmental Research Program of the Canadian Government.

**Keywords**

Environmental information system, environmental framework, environment, client-server system, sustainable management

**1 INTRODUCTION**

With half the world's population urbanized by the year 2000, sustainable management of urban bioregions is becoming one of the most important social and environmental issues of the end of the millennium. The environmental stresses on urban bioregions are unsustainable in the long term, and basic changes to their management and development are necessary.

Generally, the data describing the ecosystem and the environmental-management systems used to oversee the development of urban bioregions are fragmented, poorly integrated and even at times inconsistent, and a strategic new approach is urgently needed if sustainably managed urban development is ever to become a reality.

At the core of effective urban planning is the environmental and socio-economic information and knowledge to anticipate, and predict the impact of development on a bioregion. Currently, the necessary information, though often extensive, is fragmented, inconsistent, under-utilized and, for many key issues, simply missing. As a result, many of the important cause-effect linkages between the environment and economic development remain unclear. Action to resolve current problems must continue, of course, but what is desperately needed for the future is a substantial forward step in our ability to acquire and use data and information about a given urban bioregion. A common environmental framework is essential for assessing and interpreting regional data, for identifying critical knowledge gaps, for exploring future consequences of policies and actions, and to provide support for education about the environmental relationships in a bioregion.

Development of an environmental framework requires a multi-disciplinary, ecosystem research effort which is supported by an integrated environmental information system (EIS) containing data and knowledge bases in addition to decision-support tools for development planning in urban bioregions. An important consequence of such a system will be the availability of information which can be used as an educational medium for the general public, environmental groups, and political decision-makers.

This long-term, scientifically complex program will require creation of computer-based systems for the acquisition, storage, modeling, analysis, dissemination, presentation and creation of knowledge about environmental issues. Much of the information will be acquired from existing data as well as appropriate field work and laboratory experimentation. A number of computer technologies will be used in this research program, including: data and knowledge bases, geographic information systems (GIS), simulation and analysis systems, expert systems, multimedia and software integration methodologies. Expert systems, multimedia and software integration tools will mold these other computer technologies into a powerful decision-support tool.

Our joint research project which involves a multi-disciplinary group has been operational since 1991 and early versions of the work have been described in (Cowan, 1993a; Cowan, 1993b). The group, representing many areas of environmental research including groundwater modeling and monitoring technologies, geographic information systems, remote sensing, and transportation is determining the environmental framework, identifying interrelationships, and implementing partial prototypes of the supporting information

framework. This process has already proven the feasibility of the approach, produced some useful tools, and indicated directions for future research.

## 2 AN ENVIRONMENTAL FRAMEWORK

An environmental framework is a collection of data, information, and knowledge about the environment in an urban bioregion. This framework is designed to assist environmental planners, engineers, and scientists to create and assess plans which are consistent with sustainable development in a given urban bioregion, to isolate and solve specific environmental problems, and to identify gaps in knowledge relevant to the urban ecosystem. The collection should include geographic and socio-economic GIS data, groundwater data, runoff water data, and environmental impact knowledge to name just a few examples. The environmental framework must be presented in a coherent model so the various relationships among the data, information, and knowledge can be identified, maintained, and exploited.

Much of the environmental data collected today is or can be identified by reference to location or map grid coordinates. Therefore, this is a natural way to relate environmental data sets to each other. Forms of knowledge such as toxicity of various chemical compounds, water quality, and environmental impact assessment rules are not directly related to map coordinates, rather they are used as standards against which the geo-referenced environmental data can be measured.

Since this environmental framework focuses on urban regions and ultimately even neighbourhoods, the environmental framework model must scale effectively as the size of the locale of interest decreases. Thus, the system supporting the framework must be capable of storing and manipulating extremely large volumes of data, knowledge, and relationships between them, and easily expand by allowing the incorporation of new types of data and knowledge as they are identified.

At present environmental information is held by many different public and private organizations that have specific interests in its collection and maintenance. This situation will continue as long as organizations collect information and thus current environmental information will always be distributed both geographically and politically. Clearly the environmental framework is a very large collection of data and knowledge which must be accessed and maintained on a distributed network of computer systems.

## 3 AN ENVIRONMENTAL INFORMATION SYSTEM (EIS)

The Environmental Information System (EIS) shown in Figure 1 is a distributed information system specifically designed to support the environmental framework. The EIS can be viewed as a client/server system where the client is accessing distributed databases controlled by many different jurisdictions in both the public and private sector. The client supports four major components: modeling, simulation and analysis tools; visualization systems; user interfaces and software integration tools. The client system should also provide support for data access, and the location of new sources of data, software, and expertise.

The entire system is intended to be accessible to environmental planners, scientists, and engineers for their specific research, planning, and problem-solving activities. As well, the information framework must provide educational facilities for environmental groups,

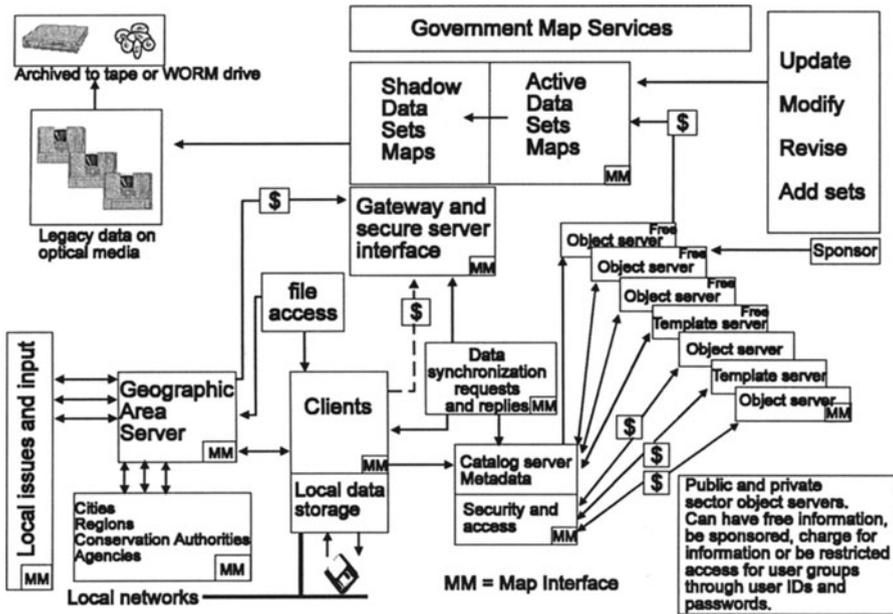


Figure 1 The Structure of an EIS.

politicians, and the public. Since the uses of an EIS are quite broadly based, most applications will be constructed from domain-specific software components using a set of software integration tools. Because the creation of an environmental framework is a research project as well as an environmental planning tool, the components of the system will remain loosely coupled to allow easy replacement of current constituents and insertion of new components.

The servers may supply data, information, knowledge and domain-specific software applications. The “main” database servers will contain base maps showing the primary geographic and physical features of a political region such as a state or province, and will normally be accessible through a government agency. Significant geo-referenced features such as buildings, well data, historical information, or hazardous materials, as well as domain-specific application templates are usually obtainable from local government authorities or organizations such as private consultants, environmental groups or historical societies. Government departments at all levels are cooperating with this project and providing support, since they anticipate that many of the results will help in solving some of their own accessibility and information management problems.

The initial thrust of the project has focused on potential applications of an EIS, since even a partial implementation could have value for a large user community. For this reason, some members of the research team have been focusing on the construction of a prototype client. The client supports local databases and basic tools that have been developed and combined to produce templates for several of the processes associated with environmental planning, monitoring, and visualization including: community-based monitoring, wetlands monitoring, watershed planning, environmental education, and cultural landscape presentation. A number

of demonstration systems have been created using these templates. Although the client is not currently connected to the servers, various ad-hoc techniques are used to load and support local databases on the client.

Although the work on developing prototype clients is continuing, we are now focusing on research issues related to accessing the servers over the Internet and prototypes of the Information Highway. The next section describes some of the issues related to this next stage.

## 4 COMPUTER SCIENCE RESEARCH ISSUES

The information framework which supports the environmental framework must eventually consist of an extensive distributed computer system. Many of the components of the information framework can be implemented now using current computer technology, but a working large-scale EIS will require significant computer science research.

### 4.1 Databases

The environmental framework will contain data, text, knowledge, graphics, and relationships among these various types. The term database is used in this document as a generic term to describe this entire collection.

Much of the data is structured and describes names of such entities as objects and compounds, and numeric data such as map coordinates, water flow rates, and chemistry. Text will encompass documents which describe specific areas within the map defined by the framework, or various standards which could be used in assessing the environmental health aspects of an urban bioregion. The knowledge will include rules which might describe environmental impact assessment procedures, or water quality tests. Graphic data could be single images, complete video presentations or more complex structures to support visualization of ecosystem behavior. The relationships provide a connection among the various components of the framework, connecting both data to text and knowledge. For example, data on ground water might be connected to water quality assessment rules.

The data in the environmental framework is likely to be collected, maintained and administered by various authorities depending on its relevance to a specific political area. Thus, the aggregate databases will be distributed over a number of computer systems and will be accessible over a computer network. Locating this information and related resources on a distributed database and making such a distributed database system work are challenging research problems. For these reasons, resource discovery (Bowman, 1994), resource browsing (Berners-Lee, 1994) and directory construction (Bumbulis, 1993) are research topics of significant interest.

Once data is located, providing access will still present problems. For example, determining if the correct data set has been chosen for a data analysis or visualization program may require preliminary viewing or computations on a large scientific database of images occupying tens or even hundreds of megabytes per image. Moving this data across the network to the application program can be time-consuming enough to be impractical. One approach under consideration is to provide multi-resolution access to remote data. Preliminary examination or browsing of large amounts of data can be performed using low-resolution data. Tentative analyses can be performed using intermediate-resolutions, and the final analysis can be performed using the data's full resolution. A complementary approach is

to reduce data transfer delays, by using the fact that data analysis and visualization programs can often provide substantial information about their future data needs.

## **4.2 Modeling and Analysis Tools, and Expert Systems**

The databases are to be used to support the environmental planning process. In this context it becomes necessary to interpolate and extrapolate the numeric data in the database over space and time to determine likely values for missing data, and to predict the consequences of present actions for the future health of the urban ecosystem. In addition, present data needs to be compared with historical information to observe changes which reflect possible detrimental modifications to the environment. Relationships need to be created between the numeric data in the database and standards so that, eventually, deterioration in environmental data will automatically be brought to the attention of the environmental planner and engineer.

Expert systems and mathematical models exist to assess water quality or monitor land use, and to predict behavior such as the dispersal of toxic wastes in a large body of water, or the effects of acid rain. These are already valuable tools in the environmental planning process. Such tools are still in their infancy and must continue to develop as part of the framework.

For example, the proposed EIS could be used to examine the issues surrounding the exploitation and protection of groundwater resources, since over half the population of North America obtains drinking water from this source. It is important to anticipate future threats to groundwater resources by conducting numerical simulations of the flow of water in the subsurface. The results of such simulations can be used in visualizations to determine optimum placement of landfill sites away from aquifers, and to develop remediation strategies for toxic waste spills. These models require, as input data, a three dimensional description of the geological formation of the subsurface, as well as topographic data, location of streams, rivers, wells and boreholes, well pumping rates, and infiltration data. Currently, development and analysis of the model input is a time-consuming and error-prone task. Much of the data must be obtained from different sources, and re-entered by hand. Addition of new information to these models such as recent borehole test data becomes a tedious process, and hence the models become rapidly out of date. The EIS could eliminate much of this difficulty.

## **4.3 Computer Graphics**

“A picture is worth a thousand words” is an aphorism which has a bearing on environmental planning. Visualization of results in the form of pictures, charts, and graphs, particularly those with some associated animation provides vital support to this process. Graphical systems can also provide a user interface which is more user tolerant and friendly.

There are now many techniques in computer graphics which can produce quite realistic views from computer-generated data and should be available as part of the information framework. New results in graphics and user interface technology should be incorporated into the information framework as they are created. Research in these areas directed toward support of the environment should be encouraged. For example, once simulations are completed, visualization methods could be used to determine the possible extent of subsurface contaminant plumes through examination of fluid flow patterns, or to demonstrate the effectiveness of various groundwater remediation schemes. The Environmental Information System (EIS) could be used to provide widespread access to the resulting high-quality visualizations for both analysis, and to assist in the scientific and political decision-

making process. Different bodies such as citizen's action groups and municipal organizations, will want to view visualizations of various simulations in order to verify, understand, and discuss the issues. This will require easy access to large video data sets, and efficient query access to large amounts of simulation data, a problem mentioned in Section 4.1.

#### **4.4 Software Integration Tools**

The tools, graphics systems, and databases just described will be used in a variety of ways to support the environmental planning process, and must be configurable to support tasks that are not yet well understood, or even contemplated. This need to use software in an unpredictable and timely fashion means that the various components will have to be loosely coupled, and capable of being "glued" together in different configurations. There is a requirement for a set of software integration tools which will bind together existing components into tools for a new application. Furthermore, the set of software integration tools must allow this binding to occur when the tool is required. For example, tools to monitor and assess a toxic spill cannot wait for the advance planning of the normal software development cycle. Preferably the software integration tools can be used by the application specialist such as the environmental planner with minimal expert help from the programmer. Such tools and the methodologies to use them are starting to become available (Wiederhold, 1992; Cowan, 1992; Cowan, 1993c; NIST, 1994), but significant progress is still required. One approach to this problem is to develop software application templates for restricted environmental processes such as watershed planning, wetlands management, or historical properties' inventory. We are building examples of these templates, which can be configured by the environmental expert to handle specific problems within an application domain. This approach will minimize the need for a programmer except when a new application template is required. Since these templates will use many common presentation, visualization, modeling and analysis tools, we expect to assemble them from various reusable software components created within the context of an application framework for environmental software. Building software applications from reusable parts presents many challenging problems and is a concern of the component-based software initiative in the United States (NIST, 1994). By restricting our work to a single application domain we expect to make significant progress in producing useful tools.

#### **4.5 Networks**

Network transmission and switching speeds are increasing dramatically (Aiken, 1992; CANARIE, 1993), and application interfaces are being defined to allow the real-time manipulation and presentation of large volumes of data such as video and other graphic images. Although fundamental network research is not part of this project, certainly network developments will be followed closely, and the EIS will be configured to take advantage of new developments.

### **5 ENVIRONMENTAL RESEARCH ISSUES**

There are two main areas where environmental scientists will have input into the process: data and knowledge evaluation and generation for various environmental models, and research into processes in the environment.

## **5.1 Data and Knowledge Evaluation and Generation**

Data and knowledge already exist for many parts of the environmental framework, but the quantity and quality of this information varies considerably. Approaches need to be developed so that this information can be integrated into the model. The process of integration will assist in identifying gaps in information, and should provide the impetus to create a base standard against which the quality of information can be assessed. One aspect of this base standard will be a set of metadata standards. Metadata will describe the format of the data and its provenance, that is, its origin and quality. A second function of the framework will be to identify and quantify the information needed to continue and enhance the modeling process. Ecosystem level studies are essential in this context since they will lead to superior integration of any information generated.

## **5.2 Research on Processes in the Environment**

Some basic environmental processes are not well understood at present, and yet knowledge of these processes is essential to the maintenance of a healthy urban environment. Examples of these environmental processes include: groundwater effects on surface water, diffuse sources of pollutants and their importance, effects of complex mixtures of pollutants on ecosystems, sequestration and release of materials in sediments and consequences of atmospheric emissions. Conducting experiments, gathering data, integrating data into the framework, and modeling the results in an urban context will provide many opportunities for cross-disciplinary studies in environmental processes. Such research always leads to new questions and new approaches, providing new insights into the ecosystem. In general, accumulation of information into the model will result in modifications to the model through an interactive process.

As the model is developed and new ecosystem research occurs, many new correlations should appear; some will deserve further investigation. Creating an environmental framework should produce a coherent and consistent baseline data set. Such baseline data is extremely important in assessing any future effects observed in the bioregion. By using a standard framework the data will be much easier to access and use, because it will be in a consistent form with standard access procedures. Access standards imply accessibility; data will be available to a wider range of users. Improvements to the environmental review and environmental assessment processes should occur because the model could provide a common database agreed to by major participants in such reviews.

## **6 STATUS OF THE ENVIRONMENTAL INFORMATION SYSTEM**

This paper contains a justification for an Environmental Framework, an outline for an Environmental Information System (EIS) to support the Framework, and presents some of the Computer Science research problems which need to be addressed. We are currently implementing the EIS in an incremental fashion and the remainder of this section contains a description of the progress to date, and some of the problems encountered.

The initial implementation concentrates on the client portion of the system because previous experiences and interviews with many potential users indicated that a properly structured component-based system that could be used to support a wide variety of environmental applications was of fundamental importance. We have constructed a prototype of such a

system, and used it to build applications to support environmental assessment and cultural landscape activities. In general we have found that the prototype can be easily configured to support new disparate applications within a few hours, thus giving us confidence in our design and in our choice of basic components as our building blocks. We are currently implementing a more robust version of this prototype, while still using the prototype for continuing experiments.

The prototype version of the EIS now operates over the Internet, in that environmental data such as maps, and map objects and their corresponding attributes can be downloaded from servers to the client system. We have also developed a prototype user interface to allow users to select the data required for their specific applications. Obviously many problems need to be solved as data is acquired from multiple sources. For example, how do we connect map objects which may come from one source with their attributes from possible different sources? Currently such a connection is performed manually. However, we will need to develop some standard approaches to encoding information so that such a process can, as a minimum, be performed semi-automatically. Locating and charging for data are other issues that must be addressed. Various directory structures such as those present in the X.500 standard (Bumbulis, 1993) need to be investigated as well as metadata standards. These and other issues related to the development of an EIS are discussed further in (Cowan, 1995).

## 7 POTENTIAL CONSEQUENCES OF THE FRAMEWORK

If successful, the process of constructing the environmental framework model should be adaptable to other regions in Canada and to other geographic areas. Third world countries would not have the data available to the same extent, but the model should provide better management decision capability in any given situation. If the framework identifies critical gaps in knowledge in certain areas it will serve an important function. As experience with other areas accumulates, the modeling process will improve. The modeling process can be applied on many scales from very small (university campus), through the municipal or regional level (small towns, cities, metropolitan areas, and watersheds) and could conceivably become a part of a Green Plan for an entire province or country. It is anticipated that it could become a major tool for Planning/Development departments, consultants, public interest groups, governments, research organizations, groups concerned with environmental assessment, and the general public. If access to the information is relatively easy, it would be a significant educational tool for both students and the public.

## 8 CONCLUSIONS

This paper has examined the structure of an Environmental Information System (EIS) for managing the interactions between human activity and the urban ecosystem. Specific reference is made to current prototype implementations and likely future directions. A complete EIS will require many years and a substantial amount of computer science and environmental research, but the current state of computer technology will allow the creation of an EIS that can have significant impact on our interaction with the environment.

## 10 REFERENCES

- Aiken, Robert. (1992) NSF Implementation Plan for Interagency Interim NREN. [gopher://es.net/NREN/NREN documents/impl.txt](mailto:gopher://es.net/NREN/NREN%20documents/impl.txt), May.
- Bumbulis, P.J., Cowan, D.D., Durance, C.M. and Stepien, T.M. (1993) An Introduction to the OSI Directory Services. *Computer Networks and ISDN Systems*, 239-249.
- Bowman, C.M., Danzig, P.B., Manber, U. and Schwartz, M.F. (1994) Scalable Internet Resource Discovery: Research Problems and Approaches. *Communications of the ACM*, 37(8):98-107, August.
- Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H.F. and Secret, A. (1994) The World-Wide Web. *Communications of the ACM*, 37(8):76-82, August.
- CANARIE. (1993) CANARIE: The Canadian Network for the Advancement of Research, Industry and Education. CANARIE Inc., 401 Laurier Avenue West, Suite 1120, Ottawa, Ontario, K1P 6H5.
- Cowan, D.D., Ierusalimschy, R. and Stepien, T.M. (1992) Programming Environments for End-Users. In *Proceedings of IFIP 92*, Volume III, 54-60.
- Cowan, D.D., Grove, T.R., Mayfield, C.I., Newkirk, R.T. and Swayne, D.A.B. (1993a) An Integrative Framework for Environmental Management and Research. In *Proceedings of the Second International Conference/Workshop on Integrating Geographic Information Systems and Environmental Modeling*, 6 pages.
- Cowan, D.D., Grove, T.R., Mayfield, C.I., Newkirk, R.T. and Swayne, D.A.B. (1993b) Managing Environmental Data An Extension to the GIS Architecture. In *Proceedings of the Workshop on Advances in Geographic Information Systems*, 10 pages.
- Cowan, D.D., Ierusalimschy, R., Lucena, C.J.P. and Stepien, T.M. (1993c) Application Integration: Constructing Composite Applications from Interactive Components. *Software Practice and Experience*, 23(3):255-276, March.
- Cowan, D.D., Koch, P.M., Mayfield, C.I. and Swayne, D.A.B. (1995) Implementing an Information Framework for Environmental Research and Management. In *International Proceedings of the Conference on Computing and Information*, 10 pages, IEEE Press.
- NIST. (1994) ATP Component-Based Software Program. [gopher://zserve.nist.gov:79/1/.menus/.atp94/9406](mailto:gopher://zserve.nist.gov:79/1/.menus/.atp94/9406).
- Wiederhold, G., Wegner, P. and Ceri, S. (1992) Towards Megaprogramming. *CACM*, 35(11), November.

## 11 BIOGRAPHIES

Donald D. Cowan received a B.A.Sc. degree in Engineering Physics from the University of Toronto in 1960, and the M.Sc. and Ph.D. degrees from the University of Waterloo in Applied Mathematics in 1961 and 1965, respectively. He has been on the Faculty of the University of Waterloo since 1962 and is currently a Professor in the Computer Science Department and Director of the Computer Systems Group. His research interests include software engineering, and the design and implementation of computer systems to support environmental management and planning. Dr. Cowan is a member of the IEEE Computer Society and the ACM.

Paul M. Koch received the B.A.Sc. and M.A.Sc. degrees in Chemical Engineering from the University of Waterloo in 1963 and 1964, respectively. He has been a member and Vice-Chairman of the University's Board of Governors. His interest in environmental issues developed from attendance at the UN Conference on the Human Environment in Stockholm, in 1972. He spent the majority of his business career in marketing assignments with IBM Canada Ltd. Since 1992, he has been self-employed and specializes in developing and facilitating business strategies that use new approaches to education and the use of information technology in handling environmental issues.

Colin I. Mayfield received a B.Sc. and a Ph.D. in Microbiology from the University of Liverpool in England in 1966 and 1969, respectively. He has been on the Faculty of the University of Waterloo since 1970 and is currently a Professor in the Biology Department. His research interests include environmental toxicology, the bioremediation of contaminated groundwater, the design and implementation of computer systems to support environmental management and planning.

David A. Swayne received a B.Sc. and Ph.D. from the University of Waterloo and an M.A. from York University. He is currently a Professor in the Computing and Information Science Department and the School of Engineering at the University of Guelph, and an adjunct professor in the Computer Science Department at the University of Waterloo. His research has focused, since 1983, on applications of computing science and information technology to the environment. Since 1991, he has collaborated with the University of Waterloo's Environmental Information Systems Project. He is a member of the Society of Industrial and Applied Mathematics.