Pathways to Enhance Environmental Assessment Information Systems*

David J. LePoire, John J. Arnish, Timothy R. Klett, Robert L. Johnson, & Shih-Yew Chen
Environmental Assessment Division at Argonne National Laboratory, Argonne, IL, USA

dlepoire@anl.gov, jarnish@anl.gov, tklett@anl.gov, rljohnson@anl.gov, sychen@anl.gov

Abstract

In the environmental field, modeling plays a critical role in connecting current data and knowledge with predictions of future events and environmental states. Environmental problems are quite challenging to solve because of the complex relationships among many contributing factors, both natural and man-made. Moreover, these problems need to be addressed not only by environmental engineers and regulators but also by a larger community consisting of concerned members of the public and nongovernmental organizations. Their demands on environmental modeling often conflict because predictions need to be accurate yet easily understood, communicated, and explored. The requirements for environmental modeling are generally identified as accuracy, defensibility, transparency, efficiency, and effectiveness. An approach to meet these requirements based on layered components, role assignments, and flexible integration strategies is identified, developed, and tested with prototypes. The results from the prototypes suggest possible enhancements for further advancing the use and communication of environmental modeling. Qualitative comparisons with other approaches are framed with identified criteria.

Keywords: environmental modeling, community informatics

* Work supported by the U.S. Department of Energy under Contract No. W-31-109-ENG

Introduction

In the environmental field, modeling plays a critical role in connecting current data and knowledge with predictions of future events and environmental states (Figure 1). Environmental problems are quite challenging to solve because of the complex relationships among many contributing factors, both natural and man-made (Constanza, Wainger, Folke & Mäler, 1993). Moreover, these problems need to be addressed not only by environmental engineers and regulators but also by concerned members of the public and nongovernmental organizations. Their demands on environmental modeling often conflict because predictions need to be accu-

Editor: John Paynter

rate yet easily understood, communicated, and explored. The increasing complexity of environmental codes also places a demand on the end user, who must translate the real environmental problem into the conceptualization allowed by the model and its options. Information on assumptions and options must be conveyed to the user to ensure that the model is applied and interpreted correctly. Open communications about the model, interface, and data components would enable software applications to be more easily developed and applied.

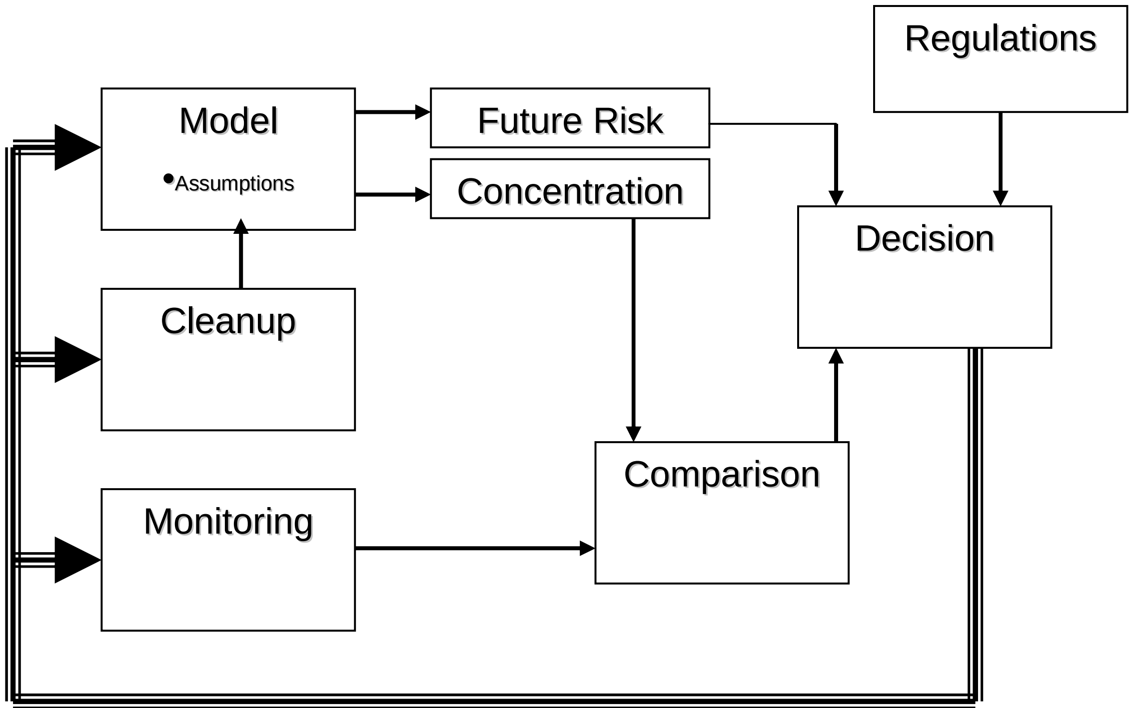

Figure 1. Environmental modeling plays an important role in the environmental decision-making process by 1) facilitating an understanding of the design of the cleanup and monitoring activities and 2) predicting future consequences of various actions.

The problem of developing and applying software modeling tools and communication technologies can be viewed in the context of "Community Informatics," which is defined as "the science and application of information and communication technologies to support human communities and their processes" (Rathswohl, 2003). In this case, the human community is the wide distributed group that is interested in the environmental issues, and the processes are the decisions, monitoring, clean-up, and mitigation efforts taken to solve or control the issues.

The complex and conflicting goals placed on environmental modeling are illustrated in recently stated goals of an environmental modeling system (Interagency Steering Committee, 2001):

- Maintain safety and protection of the environment (Accurate and Defensible),
- Increase public confidence (Transparent),
- Increase efficiency and effectiveness of decisions (Efficient and Effective results), and
- Reduce unnecessary burdens on stakeholders (Efficient and Effective input).

One modeling approach might not be sufficient to satisfy these goals. For some purposes, the most detailed models and data are appropriate to predict a situation. However, the results from these models might not facilitate a good understanding of the situation, or they might place an undue burden on the stakeholders as they try to learn how the models work and try to gather sufficiently verified data. Moreover, the results might not be directly integrated into regulatory processes, so an interpretation of the results, although accurate, would be difficult to accomplish. Another approach is to use simplified models with conservative values for data to facilitate an understanding of the results and to place bounds on them. This approach enables the end user or interpreter to focus on the important issues. Many regulatory processes include this type of analysis to help users explore the issues and decide whether a more detailed analysis is justified.

There is a wide gap between these two approaches and between their contributions to understanding and the decision-making process. Both can be enhanced with tools that allow sensitivity analysis, uncertainty analysis, and visualization and manipulation of the data.

To realize these goals in the future, the models must be flexible enough to incorporate changes in both the demands for environmental modeling and also the means to develop this modeling. These goals might be accomplished through modularization and integration based on the supply of new information technologies (ITs) originally developed for businesses. This modularization leads to separation of tasks that are handled more efficiently by developers assigned to more specific roles. ITs such as those for object-oriented code development, network-based distributed processes, database storage and manipulation, graphical user interfaces (GUIs), work-flow processing and communication, geographic information systems (GISs), and graphical visualization, have become available for integration into environmental modeling in recent years. These tools and techniques, which required large investments by commercial vendors, are available relatively inexpensively because of the demand from the business community. Software packages for environmental modeling should take advantage of these tools and techniques in responding to the field's unique needs and demands. The emphasis should not be on redeveloping industry standards but on meeting environmental modeling's unique needs by leveraging these tools and techniques.

Approach

If a next-generation risk-modeling environment is to be successful, it must address these needs for conflicting goals in a quickly changing environment. The approach taken was to identify 1) areas for shared components, 2) roles for developing components and integrated packages, and 3) multiple ways to integrate the components. The flexible integration of the components could then be used for both detailed modeling with emphasis on manipulating large data sets and for scoping models where understanding and regulatory requirements are more important. The detailed models would place additional demand on the end user for properly connecting submodels by ensuring the assumptions in each are consistent. Simpler models would be used to explore the relationship among the parameters and assumptions leading to a focused plan for further modeling action such as collecting more data or using more complex models.

We, the authors, are in a unique position in that we are involved in all identified roles (LePoire et al., 2001). For example, in the modeler role we have developed new groundwater models for radionuclide transport, new external groundshine models, and new air transport models. In the integrator role, we have developed a suite of software application codes to address radiological contamination issues for soil, buildings, and recycled products. As end-users we have applied these codes to specific sites or for specific regulations. In general however these roles are separated. Our unique position has allowed us to conduct training sessions for a variety of users, participate in code development for new environmental issues, and to participate in model comparisons for validation such as the Chernobyl accident. Our position will allow us to qualitatively assess the approach from each of the role's perspectives.

Components

The areas for shared components have been identified in the general IT field and refined to the environmental modeling field over a period of six years by a group of agencies (U.S. Department of Energy [DOE], NRC, EPA, U.S. Department of Defense [DOD], and others) (Whelan & Nicholson, 2001) (Figure 2).

- Model connectivity (model layer): Development of models at different levels of detail and complexity that are important to environmental analysis such as contaminants, transport pathways, and receptors. Analysis tools can also be modularized such as sensitivity and parameter uncertainty analysis.
- Information architecture (data layer): Input parameters, intermediate results, and final results must be stored, modified, passed between models, and displayed. A common tool to facilitate these processes is a relational database which can store not only the data but also the metadata concerning data source, reliability, defaults, and assumptions. Communication components were developed for connecting the shared database with model interfaces, user interfaces, and display tools such as reports, graphs, and GISs.
- Framework connectivity (application layer): Allow for integration of components into a self-consistent model with sufficient documentation for quality assurance purposes on the assumptions.

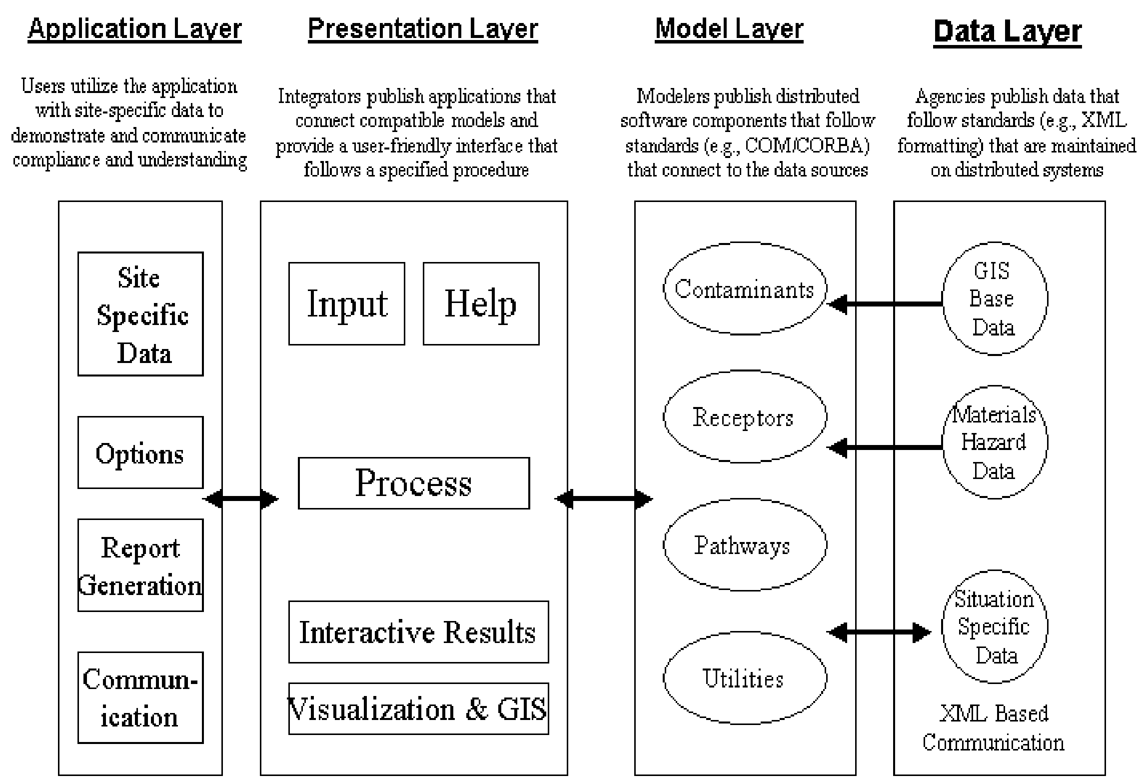

Figure 2. Areas for shared components identified in the general IT field and refined to the environmental modeling field.

- Web-based access (network layer): Allow for easily accessible software, training, and shared data sets to create a sustainable critical set of users.
- System functionality (presentation layer): The application should appear to the user as a consistent application and not a mixture of interface designs and strategies. This requires flexibility in connecting both the input and output to interface components.

Roles

Three roles are proposed for developing and using model applications, as seen in Figure 3. First, modelers develop domain-specific models and document their assumptions. Second, integrators create an application from the available models and data. Third, end users specify data and options through the integrated user interface and communicate the results to the regulators and public.

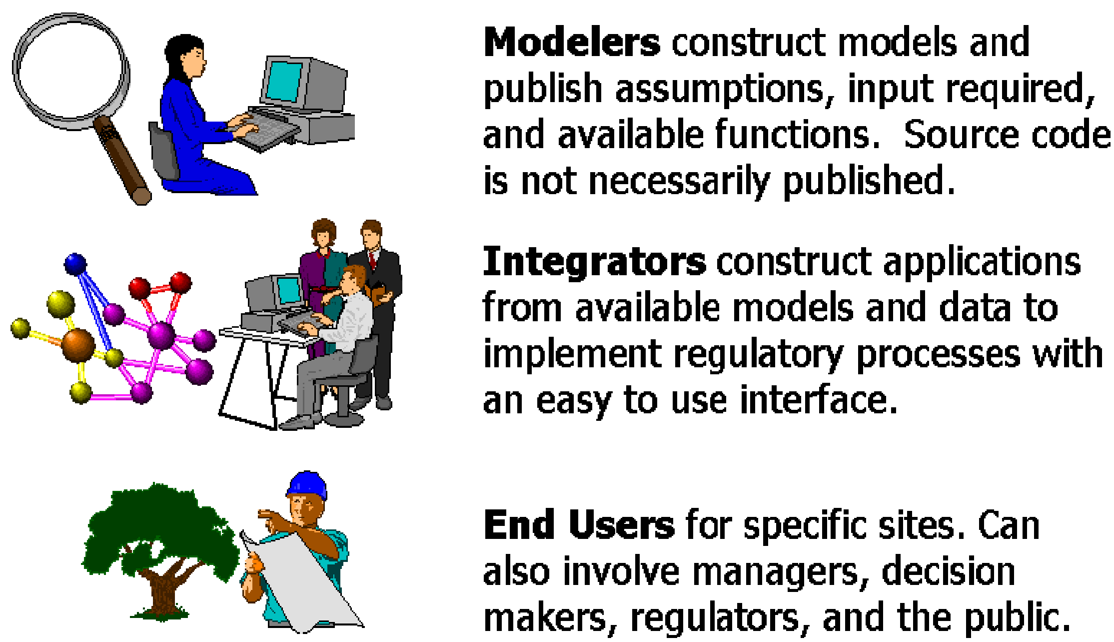

Figure 3. Three roles are proposed for developing and using model applications. First, modelers develop domain-specific models and document their assumptions. Second, integrators create an application from the available models and data. Third, end users specify data and options through the integrated user interface and communicate the results to the regulators and public. The participants usually do not work together synchronously. Products from each are generally accompanied with reports of activities that last from months to years.

This distinction of roles follows the historical trend of environmental modeling. In its earliest phase, models were created, integrated, and applied by an individual or a small team. Later more generic software applications were constructed for end users but the modeling and integration were still entwined. Some component framework systems allow independent model component development but then place the burden of integration of components and application on the end user.

This separation of responsibilities might alleviate some of the uncertainty in environmental modeling. In one study environmental risk estimates varied by six orders of magnitude for a particular case (Travis, Obenshain, Regens & Whipple, 2001). The sources of the variation were attributed to 1) model technical differences, 2) model integration differences (i.e., the various applications were not able to address the specific scenario exactly), and 3) variation in model application by the user. The first issue is addressed by modelers focusing on developing standard models. The second issue is addressed by integrators having more control to develop model integration to match complex sites. The third issue is addressed with end users better trained with tools that better fit the situations, having access to more pertinent data, and having the ability to share information.

Integration Strategies

Integration strategies must address both the preparation and integration of existing applications and the incorporation of new commercial technologies. The process of integrating previously existing applications involves: 1) separation of each package into model, data, and interface components;
2) integration of the multiple applications models into one model by connecting model inputs into outputs, integration of the multiple data structures into a common data structure through mapping, and integration of the user interface by determining required inputs and navigation, and 3) integrating these three levels (model, data, and interface) into a new package.

The integration technology might be a specific framework, an Internet or Windows environment, or as a hybrid using web services. Some specific framework model integration tools for environmental modeling include Argonne National Laboratory's (ANL) DIAS system (Sydelko, Majerus, Dolph & Taxon, 1999), the U.S. Environmental Protection Agency's (EPA) MIMS system, Pacific Northwest National Laboratory's (PNNL) FRAMES (Whelan et al., 1997), and GoldSim (GoldSim, 2004). These tools offer a system of utilities for model integration and data communication. There are also many other ways to accomplish object construction, wrapping, and object integration with commercial tools such as J2EE (Perrone, Venkata & Schwenk, 2003), ColdFusion (Forta, 1998), Microsoft [MS] .NET (Hollis & Lhotka, 2001) that offer a flexible integration with commercial components.

These last options of using generic commercial development environments (.NET and ColdFusion) were chosen to be explored for the integration technology. They offer a path to continued technological progress through commercial tool development without the need for developers' exclusive dedication.

Methods

Demonstration projects were chosen both to address a current need among radiological analysts and to be potentially useful in later applications. The projects demonstrate the wide variety of integration techniques and ways to use components based on existing software packages, new models, and commercial components (LePoire et al., 2001).

- *Connecting models and user-interfaces using a standard database structure*: Often there is a need for a new environmental analysis which requires models of two separate applications to be combined. The applications might be well validated and verified but the underlying source code might not be available or written in legacy style. In these cases it is useful to have the ability to place wrappers around this legacy code to allow for more flexible use of the models.

  In this example case, the DUST application calculated the release of radionuclides over a period of time from an engineered landfill. The model however does not consider subsequent movement of the radionuclides through the groundwater to a potential individual and the harm that this may cause. This second set of models is incorporated in the RESRAD-OFFSITE package. Using the NRC's DUST (Sullivan, 1993) package and a modified RESRAD-OFFSITE package (Yu et al., 2001), the models were integrated into a desktop application, with DUST providing a leaching source term to the groundwater and RESRAD providing the multipathway dose assessment from that point. The user-interface and model assumptions were maintained.

This project demonstrates a model linkage between two independent tools. To accomplish the integration, each model was separated into separate layers: user interface (Presentation), model interface (Model), and data components (Data). These layers were integrated together and then packaged in a new application (Application). By maintaining the look and feel of both RESRAD and DUST, this application can be easily utilized by users of either program (Effective & Efficient). Because each model's assumptions were maintained, this application's results can be easily verified (Defensible). Both DUST and RESRAD are well-documented and have clear assumptions. By making the transition between these two programs seamless for the user, this application maintains those same easily understood assumptions (Transparency).

In order to wrap the code, the input data was extracted from the database and passed to the model. The conversion from the database to the DUST input file was tedious because of the formatting, data array structures, and exceptions.

- *Wrapping calculational components with database access for integration with web based GIS tools:* Many of the models are incorporated in desktop applications. Some applications have a well separated interface between the graphical user interface and the model calculations. In these cases it would be nice to be able to web-enable the models and incorporate them with an interface that includes current commercial visualization and interaction tools. In this example an application with models for environmental transport and individual exposure estimation for radioactive contamination was prepared for web access and connected to a web-based commercial GIS system. Specifically, the RESRAD model was wrapped with a preprocessor and postprocessor for web execution. The pre- and postprocessors allowed simple connections to a customized, simplified, web-based user interface and commercial visualization graphing and GIS packages.

  This project integrated new technologies into an existing code. The interface (Presentation) to the RESRAD input parameters was accomplished through the web-based interface. The RESRAD model (Model) was accessed through a simple FORTRAN DLL that called the RESRAD system batch file. The preprocessor was developed to read the input data from a database and place them in the RESRAD input file. The postprocessor read the RESRAD results and inserted them into a table in a database that was accessed by the graphing and GIS packages (Data). This application makes the RESRAD code more flexible in that it makes the results of the run accessible to other programs/modules. The web-based interface for the application is available to any user with an internet connection, and is compatible with PDAs (Effective & Efficient). The RESRAD calculations are not changed; therefore the same results would be obtained with an input file generated through the actual RESRAD interface (Defensible).

- *Distributed computing using web services:* Certain data structures and corresponding objects with defined methods are usually needed in applications. While local reuse of objects and data can be accomplished for a small team, a larger distributed set of developers require a distributed solution. In this example, a nuclide data and object component was developed to allow common access to applications of data structures through a web service. The nuclide data were obtained from a distributed server and used by a local application that could then manipulate the nuclide structure in a common technique.

  This project demonstrated the use of web services to pass data and methods with the use of HTTP and XML. A method that would take input on a radionuclide and deliver decay chain information was developed. The data were recursively extracted from two database tables (Data). Web services make the development of the integrated package transparent for the application integrators. Calls are written to the remote object, and the developer no longer has to be too concerned about the computer “handshaking” and passing of data via XML to the server.

- *Online Training:* Web-based training systems have been explored for guiding users through a standard analysis. The workshop material has been captured in dynamic multimedia (slides and audio).
- *Animation:* Scaled schematic diagrams give feedback of the input parameters but results also could be made more meaningful by generating conceptual animations of the forward movement of the contaminants and the reverse analysis of exposure. Some simple techniques for displaying one of the more complicated models—movement into the groundwater and subsequent extraction—was demonstrated but not refined enough for incorporation into the software. One of the major problems is determining the scale of distance and time, since the models depend on multiple parameters that vary over several orders of magnitude.

Results

Table 1 lists the methods that are implemented and being considered to support the four criteria (defensibility, transparency, effectiveness, and efficiency) and four layers (model, data, presentation, and application). While many previous approaches emphasized defensibility and efficiency, new technologies allow for computing resources to be used to enhance effectiveness and transparency. Actions that could improve transparency include the interactive exploration of intermediate results, sharing of information by distributed users with the ability to maintain metadata, and inclusion of options to explore model assumptions.

Another aspect of transparency is the ability to focus on the important factors and explain the fundamental cause and effects simply. A conceptual diagram of how this conflicts with a traditional approach of defensibility is depicted in Figure 4. In modeling the exposure of a receptor to some contamination there may be many potential pathways and factors. For defensibility, the process is detailed and all possible pathways are considered in high accuracy. For transparency, the focus is shifted to the important aspects of the model and simplified to provide a higher-level understanding.

Table 1. Grid of criteria and software layers for both the present and potential enhancements for the future.

		Defensibility and Accuracy	Transparency	Effectiveness	Efficiency
Model	Present	Pathway validation; Integrated Validity, Verification	Manual with assumptions	Options for model assumptions	Integrated model optimization
	Future	Ability to select a range of models based on different scales	Intermediate results displayed in diagrams or animations	Partitioning with documented connections	Partitioning at appropriate level
Data	Present	Uncertainty analysis	Data Collection Handbook stating need for parameter, measurement technique and further data sources.	Sensitivity Analysis	Defaults; Scenario Templates
	Future	Semi automated process to determine source and need for uncertainty analysis	Allow users to add metadata and notes to input parameters.	Web-based data sharing	Web-based parameter data-base; Web service for data
Presentation	Present	Complete textual report with input, intermediate results, and final results	Scaled diagram of physical layout.	Interactive ta-bles and plots of final results	Organized graphic user interface with batch command option.
	Future	Results database for use in independent visualization tools.	Scaled diagram of contamination movement	Interactive in-termediate results.	Distributed partitioning of monte carlo probabilisitic runs
Application	Present	Construct data-base of scenarios considered by others	Manual	Tie to Sampling tool	Ability to integrate at a very granular level
	Future	Maintain search-able user's comments and ratings	Generate template of generic assumptions	Tie to GIS Tools	Enable applica-tions in web sites for real time distributed analysis

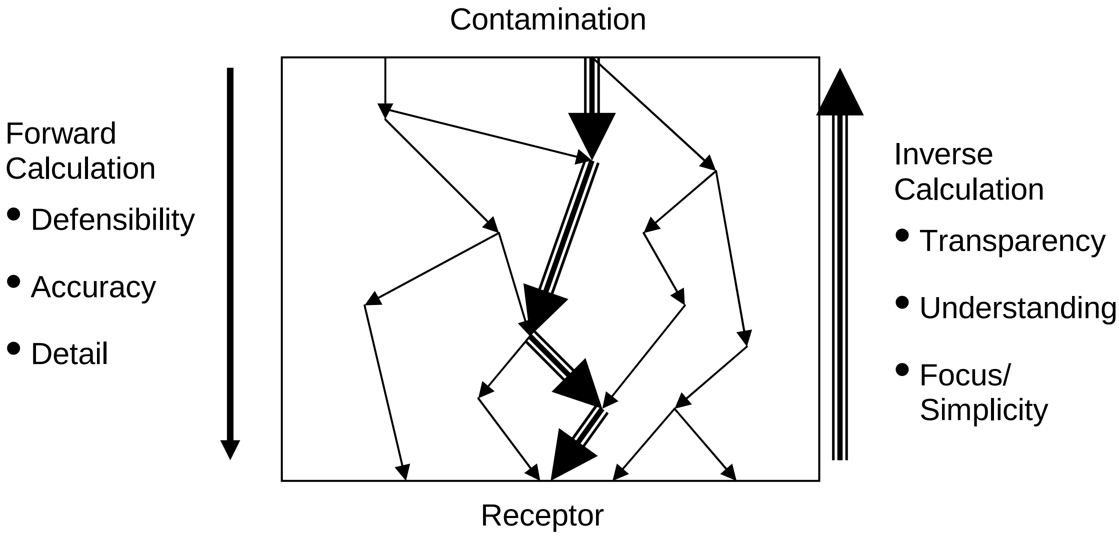

Figure 4. Difference in requirements for meeting the criteria of defensibility and transparency. In modeling the exposure of a receptor to some contamination there may be many potential pathways and factors. For defensibility the process is detailed, all possible pathways are considered in high accuracy. For transparency the focus is shifted to the important aspects of the model and simplified to provide a higher-level understanding.

Discussion

This approach organizes the necessary disparate groups (modelers, integrators, and end-users) efficiently and effectively. Note that the approach differs from traditional Group Decision Support Systems in that the groups work asynchronously to develop different class of solutions and are only weakly coupled through long-term communication about new models and environmental issues. That is the end-user usually does not work directly with an integrator or modeler but depends on their work through their published applications.

While although the example prototypes demonstrate the advantages of this flexible approach towards the goals articulated by the Interagency Steering Committee, there are further issues to resolve before the approach is accepted and implemented. Surveys and response to demonstrations could be conducted with a wider set of participants encompassing the 3 roles. A general implementation of this flexible approach would first require some agreement within a subset of the environmental modeling and software development community. The path to the implementation would possibly require extensive preparation of legacy but well verified and validated software. This investment in preparation would only be recovered if the components are widely reused.

References

Constanza, R., Wainger, L., Folke, C. & Mäler, K. G., (1993, September). Modeling complex ecological economic systems. *BioScience, 43*, 545-555.

Forta, B. (1998). *The coldfusion web application construction kit*. Indianapolis, IN: Que.

GoldSim (2004, January), Radioactive and hazardous waste management applications using GoldSim. GoldSim White Paper. Retrieved September 2004 from http://www.goldsim.com/Solutions/WhitePapers.asp

Hollis, B., & Lhotka, R. (2001). *VB.NET programming with the public beta*. Birmingham, UK: Wrox Press.

Interagency Steering Committee. (2001). Summary Report on Meeting of the Interagency Steering Committee on Research and Development of Multimedia Environmental Models -- Rockville, Maryland, June 18-19. Retrieved September 2004 from http://iscmem.org/Documents/Steering_2001_06_18.pdf

LePoire, D. J., Arnish, J., Gnanapragasam, E., Klett, T., Johnson, R., Chen, S.Y., et al. (2001, October). *Openlink: A flexible integration system for environmental risk analysis and management.* Argonne National Laboratory. Retrieved February 2004 from http://web.ead.anl.gov/openlink/ANL-EAD-TM-114.pdf

Perrone, P., Venkata, S. R., & Schwenk, T. (2003). *J2EE developer's handbook*. Indianapolis IN: SAMS Publishing.

Rathswohl, E. (2003). Introduction to the special series on community informatics. *Informing Science Journal, 6*, 101–102. Retrieved from http://inform.nu/Articles/Vol6/v6p101-102.pdf

Sullivan, T.M. (1993, May). *Disposal unit source term (DUST) data input guide*. NUREG/CR-6041, BNL-NUREG-52375. Upton, NY: Brookhaven National Laboratory.

Sydelko, P. J., Majerus, K. A., Dolph, J. E., & Taxon, T. N. (1999). A dynamic object-oriented architecture approach to ecosystem modeling and simulation. *Proceedings of the 1999 American Society of Photogammetry and Remote Sensing (ASPRS) Annual Conferenc*e. Portland, OR, May 19–21. Retrieved September 2004 from http://www.dis.anl.gov/DIAS/papers/ASPRS.doc

Travis, C. C., Obenshain, K. R., Regens, J. L., Whipple, C.G. (2001). Limitations of multimedia models for use in environmental decision making. *Environmental Monitoring and Assessment, 71*, 51–60.

Whelan, G., Castelton, K. J., Buck, J. W., Hoopes, B. L., Pelton, M. A., Strenge, D. L., Gelston, G. M., & Kickert, R. N., (1997, October). *Concepts of a framework for risk analysis in multimedia environmental system*s. Richland, WA: Pacific Northwest National Laboratory. Retrieved September 2004 from http://mepas.pnl.gov/FRAMESV1/frames_doc.pdf

Whelan, G., & Nicholson, T. (Eds.). (2002, May). *Proceedings of the Environmental Software Systems Compatibility and Linkage Workshop*. Richland, Wash.: Pacific Northwest National Laboratory. Retrieved September 2004 from http://www.nrc.gov/reading-rm/doc-collections/nuregs/conference/cp0177/0177.pdf

Yu, C., Zielen, A. J., Cheng, J. J., LePoire, D. J., Gnanapragasam, E., Kamboj, S., Arnish, J., et al. (2001, July). *User's Manual for RESRAD Version* 6. ANL/EAD-4. Argonne, IL: Argonne National Laboratory.

Biographies

David J. LePoire is an Environmental Information System Analyst in the Environmental Assessment Division at Argonne National Laboratory. He holds a Ph.D. in Computer Science from DePaul University in Chicago and a bachelor degree in physics from the California Institute of Technology in Pasadena. He has over 14 years experience in the design, development, and implementation of software and models for environmental risk analysis. His research interests include technological innovation, environmental modeling, and risk analysis.

John J. Arnish is an Engineering Specialist in the Environmental Assessment Division at Argonne National Laboratory. He holds a masters degree in nuclear engineering from the University of Tennessee and a bachelors degree in physics from Southern Illinois University. His research interests include modeling environmental transport of radionuclides and applications of Bayesian analysis.

Timothy R. Klett is a Scientific Programmer in the Environmental Assessment Division at Argonne National Laboratory. He holds a B.S. in Technology Management from the University of St. Francis and expects a M.S. in Computer Science from the Illinois Institute of Technology in December of 2004. He has over 4 years of experience designing, implementing, testing, and documenting computerized systems and models for environmental risk analysis. His research interests include Software Engineering, Database and Information Systems, and Scientific Computing.

Robert L. Johnson is the Environmental Information Management & Communication Program manager at the Environmental Assessment Division at Argonne National Laboratory. He holds a Ph.D. in Hydrological Systems from Cornell, a masters degree in Environmental Engineering Systems from Johns Hopkins, and a Mathematics B.S. from Calvin College.

Shih-Yew Chen is currently Senior Environmental Systems Engineer at Argonne. He has over 25 years experience in the development and application of methods pertaining to radiological health, safety, environmental risk assessment, and waste management. Since 1997 he has served as the Strategic Area Manager in Risk and Waste Management of Argonne National Laboratory's Environmental Assessment Division. At Argonne, he developed the Integrated Risk Assessment Program that addresses the following areas: environmental cleanup; transportation of hazardous and radioactive materials; environmental impact; environmental policy analysis, and the emerging areas such as homeland security. His research interests focus on the development of risk-informed approaches to environmental management.