Lawrence Berkeley National Laboratory
Recent Work

Title
Integrated environmental modeling: A vision and roadmap for the future

Permalink
https://escholarship.org/uc/item/27n679nb

Journal
Environmental Modelling and Software, 39

ISSN
1364-8152

Authors
Laniak, GF
Olchin, G
Goodall, J
et al.

Publication Date
2013

DOI
10.1016/j.envsoft.2012.09.006

Peer reviewed
Integrated environmental modeling: A vision and roadmap for the future

Gerard F. Laniak\textsuperscript{a,∗}, Gabriel Olchin\textsuperscript{b}, Jonathan Goodall\textsuperscript{c}, Alexey Voinov\textsuperscript{d}, Mary Hill\textsuperscript{e}, Pierre Glynne\textsuperscript{f}, Gene Whelan\textsuperscript{a}, Gary Geller\textsuperscript{f}, Nigel Quinn\textsuperscript{g}, Michiel Blind\textsuperscript{h}, Scott Peckham\textsuperscript{i}, Sim Reaney\textsuperscript{j}, Noha Gaber\textsuperscript{k}, Robert Kennedy\textsuperscript{l}, Andrew Hughes\textsuperscript{m}

\textsuperscript{a}US Environmental Protection Agency, Office of Research and Development, USA
\textsuperscript{b}US Environmental Protection Agency, Office of the Science Advisor, USA
\textsuperscript{c}University of South Carolina, Department of Civil and Environmental Engineering, USA
\textsuperscript{d}University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), Netherlands
\textsuperscript{e}US Geological Survey, National Research Program, USA
\textsuperscript{f}Jet Propulsion Laboratory, California Institute of Technology, USA
\textsuperscript{g}Berkeley National Laboratory, USA
\textsuperscript{h}Deltares, Netherlands
\textsuperscript{i}INSTAAR, University of Colorado, Boulder, USA
\textsuperscript{j}Durham University, Department of Geography, UK
\textsuperscript{k}US Environmental Protection Agency, Office of the Administrator, USA
\textsuperscript{l}US Army Corps of Engineers, Engineer Research and Development Center, USA
\textsuperscript{m}British Geological Survey, Keyworth, UK

\textbf{Abstract}

Integrated environmental modeling (IEM) is inspired by modern environmental problems, decisions, and policies and enabled by transdisciplinary science and computer capabilities that allow the environment to be considered in a holistic way. The problems are characterized by the extent of the environmental system involved, dynamic and interdependent nature of stressors and their impacts, diversity of stakeholders, and integration of social, economic, and environmental considerations. IEM provides a science-based structure to develop and organize relevant knowledge and information and apply it to explain, explore, and predict the behavior of environmental systems in response to human and natural sources of stress. During the past several years a number of workshops were held that brought IEM practitioners together to share experiences and discuss future needs and directions. In this paper we organize and present the results of these discussions. IEM is presented as a landscape containing four interdependent elements: applications, science, technology, and community. The elements are described from the perspective of their role in the landscape, current practices, and challenges that must be addressed. Workshop participants envision a global scale IEM community that leverages modern technologies to streamline the movement of science-based knowledge from its sources in research, through its organization into databases and models, to its integration and application for problem solving purposes. Achieving this vision will require that the global community of IEM stakeholders transcend social, and organizational boundaries and pursue greater levels of collaboration. Among the highest priorities for community action are the development of standards for publishing IEM data and models in forms suitable for automated discovery, access, and integration; education of the next generation of environmental stakeholders, with a focus on transdisciplinary research, development, and decision making; and providing a web-based platform for community interactions (e.g., continuous virtual workshops).

Published by Elsevier Ltd.

\textbf{1. Introduction}

Integrated environmental modeling (IEM) is a discipline inspired by the need to solve increasingly complex real-world problems involving the environment and its relationship to human systems and activities (social and economic). The complex
and interrelated nature of real-world problems has led to a need for higher-order systems thinking and holistic solutions (EPA, 2008b; Jakeman and Letcher, 2003; MEA, 2005; Parker et al., 2002). IEM provides a science-based structure to develop and organize multidisciplinary knowledge. It provides a means to apply this knowledge to explain, explore, and predict environmental-system response to natural and human-induced stressors. By its very nature, it breaks down research silos and addresses the perceptions of 'black-box' modeling, and increases awareness and detection of unintended consequences of decisions and policies.

IEM concepts and early models are now more than thirty years old (Bailey et al., 1985; Cohen, 1986; Mackay, 1991; Meadows et al., 1972; Walters, 1986). With the emergence of problems related to regional-scale land-use management, impacts of global climate change, valuation of ecosystem services, fate and transport of nanomaterials, and life-cycle analysis, the application of IEM is growing. National and international organizations have commissioned studies to determine research directions and priorities for integrated modeling (Blind et al., 2005a, 2005b; EC, 2000; ICSU, 2010; NSF; Schellekens et al., 2011). Senior managers in government, academia, and commercial organizations are restructuring operations to facilitate problem solving for which the expertise needed for integrated modeling are actively pursuing inter-organizational collaborations (e.g., Delsman et al., 2009; ISCMEM; OpenMI, 2009). Environmental assessors are utilizing IEM science and technologies to build integrated modeling systems that will address specific problems at varying scales (Akbar et al., 2013; Bergez et al., 2013; Linker et al., 1999; Mohr et al., 2013; Quinn and Jacobs, 2006). Finally, policy developers and decision makers are asking for and processing information synthesized from holistic systems-based modeling approaches (EPA, 2008b; ABARE—BRS, 2010).

The primary motivation and input for this paper are drawn from a series of workshops held during the past several years (Table 1). The workshops were open forums convened to share knowledge, experience, and future visions related to IEM. The workshops were attended by a cross-section of IEM practitioners including environmental modelers, software technologists, decision analysts, and managers. Participants represented government, academia, and the private sector.

The principal message from the workshops is a call to elevate solutions to key IEM issues and challenges to a level of community above individual groups and organizations. In effect, to establish an open international community environment for pursuing the ability to share and utilize the broad science of IEM by communicating ideas, approaches, and utilizing modern technologies and software standards. The purpose of this paper is to synthesize the knowledge and perspectives shared during the workshops and present a holistic view of the IEM landscape and a roadmap, consisting of goals and activities, to guide its navigation. The remainder of this introduction is intended to

Table 1

<table>
<thead>
<tr>
<th>Workshop title</th>
<th>Sponsor</th>
<th>Date</th>
<th>Organizations represented</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>iEMSs 2010 Conference</td>
<td>iEMSs</td>
<td>July 2010</td>
<td>- 75 attendees&lt;sup&gt;e,k,j,m,n,p,v&lt;/sup&gt;</td>
<td>This roadmap paper</td>
</tr>
<tr>
<td>The International Summit on Integrated Environmental Modeling</td>
<td>USGS</td>
<td>US EPA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> US NRC: US Nuclear Regulatory Commission (http://www.nrc.gov/).
<sup>b</sup> US EPA: US Environmental Protection Agency (http://www.epa.gov/).
<sup>c</sup> DOE: Department of Energy (US) (http://energy.gov/).
<sup>d</sup> US ACoE: US Army Corps of Engineers (http://www.usace.army.mil).
<sup>e</sup> NGO: Non-Governmental Organizations.
<sup>f</sup> EC: Environment Canada (http://www.ec.gc.ca/).
<sup>g</sup> EU: European Union (http://europa.eu/).
<sup>h</sup> ISCMEM: Intergency Steering Committee for Multi-media Environmental Modeling (US Federal Agencies) (http://iemhub.org/topics/ISCMEM).
<sup>i</sup> CEH UK: Center for Ecology and Hydrology, UK (http://www.ceh.ac.uk/).
<sup>j</sup> iEMSs: International Environmental Modeling and Software Society (http://www.iemss.org/society/).
<sup>k</sup> OGCI: Open Geospatial Consortium (http://www.opengeospatial.org/).
<sup>l</sup> CUAHSI: Consortium of Universities for the Advancement of Hydrologic Science, Inc. (http://www.cuahsi.org/).
<sup>m</sup> OpenMI: Open Modeling Interface (Association) (http://www.openmi.org/).
<sup>n</sup> USDA: US Department of Agriculture (http://www.usda.gov/).
<sup>o</sup> CSDMS: Community Surface Dynamics Modeling System (http://csdms.colorado.edu/wiki/Main_Page).
<sup>p</sup> NRC (Italy): National Research Council (Italy) (http://www.cnr.it/sitocnr/Englishversion/Englishversion.html).
<sup>q</sup> NSF: National Science Foundation (http://www.nsf.gov/).
<sup>r</sup> ONR: Office of Naval Research (US) (http://www.onr.navy.mil/).
<sup>s</sup> NASA: National Aeronautics and Space Administration (US) (http://www.nasa.gov/).
<sup>u</sup> NOAA: National Oceanic and Atmospheric Administration (US) (http://www.noaa.gov/).
<sup>v</sup> BGS: British Geological Survey (http://www.bgs.ac.uk/).
provide a definition of IEM relative to several similar terms, describe the role of IEM in the decision making and policy development\(^1\) process, and establish a conceptual view of IEM as a landscape with interdependent elements. Sections 2–5 then present each element of the IEM landscape, including an integrated roadmap of activities that addresses the associated collection of issues and challenges. Conclusions and a summary are presented in Section 6.

1.1. Related terms

Within the literature, multiple terms related to “integrated environmental modeling” create a distracting confusion for practitioners and users alike. From our view, the terms are more similar than dissimilar; and are collectively defined within the greater context of environmental decision making and policy development (see Box 1). In any of the definitions, we perceive the term integrated to convey a message of holistic or systems thinking (sensu Tress et al., 2005) and assessment as a message of decision or policy relevance (Tol and Vellinga, 1998), while modeling indicates the development and/or application of computer based models.

1.2. The environmental decision and policy development process and the role of IEM

The environmental decision and policy development process are illustrated in Fig. 1; similar descriptions are presented by others (CMP, 2007; Jakeman et al., 2006; Liu et al., 2008; Van Delden et al., 2011; Zagone, 2002). The process can be described as a loop containing two principal stages: decision/ policy and modeling/monitoring. Stakeholders in this process can be grouped by stage. Decision stakeholders are primarily concerned with the problem, its impacts, representation of interests/ concerns, management scenario development, and decisions related to solving the problem. Science stakeholders are primarily concerned with the organization and application of science-based knowledge in the form of data, models, and methods for the purpose of informing decisions.

The process begins in the decision/policy stage with the formulation of a problem statement that defines the causes of concern, policy or decision context, boundaries and objectives, management scenarios and options, solution criteria (including tolerance for uncertainty), and resource constraints. The decision stage is coupled to the modeling/monitoring stage by the system conceptualization, which represents a high level view of the social—economic—environmental system within which the problem occurs. It represents the common view of the system constructed jointly by all stakeholders. Formulating the system conceptualization requires the merging of often different world views held by the stakeholders. The conceptualization forms the basis for developing a detailed modeling- and/or monitoring-based solution. The relevant science, in the form of models, data, and assessment strategies, is organized and executed in the modeling stage. The modeling stage is coupled back to the decision stage by a process of information synthesis which consolidates and interprets modeling results for use by decision makers. Feedback and iteration within and across these stages are essential aspects of the modeling (Jakeman et al., 2006) and stakeholder-driven IEM and decision processes.

The role of the IEM modeler in the decision/policy process is often not limited to the modeling stage or the technical details of the modeling effort. Kragt et al. (2013) present an in depth discussion of the various roles modelers may play in structuring and executing integrative research projects. Because of the modelers natural systems orientation he/she may perform roles that include facilitator, knowledge broker, technical specialist, and leader. These roles are equally applicable in the larger decision/policy context.

1.3. IEM landscape

To organize the myriad of topics discussed during the workshops and provide an intuitive structure for presenting the roadmap, we characterize IEM as a landscape and recognize four interdependent elements: applications, science, technology, and community. These elements are presented in Box 2 with descriptions that reflect the current and envisioned future of IEM. In the following sections, we briefly discuss these elements from the perspective of current practices and the issues and challenges that must be addressed to advance IEM. To address the issues and challenges, a set of activities in the form of a roadmap are organized and presented.

---

\(^1\) For efficiency, in this paper when we refer to decision making alone we intend to include policy development as well.

---

Box 1. Terms related to integrated environmental modeling.

- **Conventional Modeling**: a process of creating a simplified representation of reality to understand it and potentially predict and control its future development. Models are generally single purpose (i.e., represent a single modeling discipline) and can come in a variety of forms and implementations, including mental, verbal, graphical, mathematical, logical, physical, etc. (Voinov, 2008).
- **Integrated Modeling**: includes a set of interdependent science-based components (models, data, and assessment methods) that together form the basis for constructing an appropriate modeling system (EPA, 2008b, 2009).
- **Integrated Assessment**: seeks to provide relevant information within a decision making context that brings together a broader set of areas, methods, styles of study, or degrees of certainty, than would typically characterize a study of the same issue within the bounds of a single research discipline (Parson, 1995; Weyant et al., 1996; Jakeman and Letcher, 2003).
- **Integrated Assessment Modeling**: an analytical approach that brings together knowledge from a variety of disciplinary sources to describe the cause–effect relationships by studying the relevant interactions and cross-linkages (Rotmans and van Asselt, 2001; Rosenberg and Edmonds, 2005).
- **Integrated Environmental Decision Making**: an approach for evaluating complex environmental problems holistically by integrating resources and analyses to address the problems as they occur in the real-world; including input from appropriate stakeholders (EPA, 2000).
- **Participatory Modeling**: a generic term used for modeling strategies that rely upon stakeholder involvement and participation in various forms. In various applications also known as group model building, mediated modeling, companion modeling, shared vision planning, participatory simulation, etc. (Voinov and Bousquet, 2010).
2. IEM applications

IEM applications are the stakeholder community’s methods for defining, selecting, integrating, and processing the combination of environmental, social, and economic information needed to inform decisions and policies related to the environment (i.e., implementation of the process shown in Fig. 1). During the past two decades, there has been a steady evolution toward IEM in the range and complexity of environmental issues and problems, related decisions and policies, and the modeling performed to inform the decisions. While decision makers will continue to address traditional problem sets involving environmental quality standards and compliance, management challenges are now framed in ecological, social, and economic terms (MEA, 2005). The literature contains a growing number of studies involving the application of integrated environmental modeling concepts and approaches. Table 2 lists a number of such examples organized by dominant IEM characteristics.

Dominant themes throughout workshop discussions concerning IEM applications included stakeholder involvement, adaptive management strategies, education, peer review, and reusability. In the following sections we discuss these topics from the perspective of current practices, issues, and challenges that lead to the IEM applications roadmap presented first, in Fig. 2.

### 2.1. Stakeholder involvement

Stakeholders have become an intrinsic part of the systems analytical approach, which is essential for environmental management. The idea of transdisciplinarity is based on stakeholder involvement in solving multidisciplinary problems, where stakeholders are used to estimate the understanding across formal and informal knowledge bases and to glue together the data and theories originating from different disciplines. The definition of stakeholders in this case is quite broad (Krueger et al., 2012) and in the case of IEM applications includes experts (scientists, engineers, educators, and decision makers) as well as non-experts (in the traditional sense). An often overlooked but potentially valuable group of stakeholders is represented by citizen—science networks. Citizen—science networks can contribute in many different ways, including direct monitoring of natural resources and environmental conditions, facilitate knowledge transfer between scientists and lay public, test IEM and monitoring technologies and processes (e.g., Smartphone apps), and provide historical knowledge and local stakeholder continuity to ensure persistence and improvement of IEM application efforts.

While the importance and value of involving the full stakeholder community in the decision and application process is recognized there remains a significant need for guidelines for managing, facilitating, and reporting the dynamic interactions among stakeholders (Arciniegas et al., 2013). These interactions are critically important in establishing a common understanding and appreciation of the problem, the relevant social–economic–environmental system, the role of modeling, and the information provided by the modeling. The process of merging differing world views, priorities, and value systems into a unifying and objective approach to problem elucidation and resolution will require social science expertise (Kalougher et al., 2013). Additional issues related to the science content of these interactions are discussed in Section 3.1.

### 2.2. Adaptive decision process

Adaptive management (AM) refers to the realization that with respect to the complexities of the human–environmental system we are “learning as we go”. IEM–based decisions and policies are based on existing knowledge, understanding, and observations and often prove to fall short in terms of intended outcomes. There is, therefore, an intrinsic need for iteration and adaptation in the decision process. Combined with integrated modeling, adaptive management provides the stakeholder community with a means to jointly build an understanding of the system, conduct experiments related to the exercise of management options, and refine and update management strategies when coupled with ongoing monitoring.

According to Stankey et al. (2005), the specific idea of AM, as a strategy for natural resource management, can be traced to the seminal work of Holling (1978), Walters (1986), and Lee (1993). It is a framework that promotes iterative learning–based decision making (Holling, 1978) from management outcomes and making adjustments as understanding improves (Williams, 2011) and will probably never converge to a state of equilibrium involving full knowledge and optimum productivity (Walters, 1986). Walters (1986) defined AM as consisting of three essential tasks: structured synthesis and analysis, use of formal techniques that consider uncertainties and result in optimal decisions and policies, and design and implementation of monitoring programs to collect data needed to measure the effectiveness of decisions and advance

---

2. Box 2. Envisioned IEM landscape.

1. **IEM Applications** reflect problem formulations and solution approaches that are transparent and holistic, establishing a view of the human–environmental system that recognizes the interdependent relationships among responsible organizations, management scenarios (decision options), co–occurring stressors, the environment, and socio–economic structures. This integrative systems approach engages all stakeholders throughout the decision process.

2. **IEM Science** is transdisciplinary, involving the integration of social, economic, and environmental science into modeling systems that describe and forecast the behavior of the human–environmental system in response to natural and human induced stress. The science provides methods for model evaluation, including the characterization and communication of model sensitivities and uncertainty to inform both decision and science stakeholders according to their role and related needs for information. IEM is recognized as a science discipline and is taught in schools and universities.

3. **IEM Technology** provides the means to express, integrate, and share the science of IEM. It provides standards and tools to facilitate the discovery, access, and integration of science components (data, models, and assessment strategies) by stakeholders world wide. Integrated modeling systems are constructed and executed on a variety of platforms serving research, applications, and education at a variety of spatial, temporal and complexity scales.

4. A **Community** of IEM stakeholders and associated organizations engages in, invests in, and contributes to, the shared and open development of integrated environmental science and related computer–based technologies. The community is open and grows to further engage scientists, engineers and educators, as well as interested or concerned citizens, decision makers, and their associations and gains their support for the development and application of IEM.

---

Bormann et al. (1994) defined four phases (plan, act, monitor, and evaluate), whereas CMP (2007) identified five steps (conceptualize, plan and monitor, implement and monitor, analyze/use/adapt, capture/share learning), where monitoring is an important component in each approach.

Implementing adaptive management strategies is complicated by the involvement of a diverse stakeholder community and the idea that complex problems have many potential solutions, each perhaps appealing to a subset of the stakeholder community. The challenge for IEM is to reflect this adaptive management process in the design and execution of applications and, in particular, building modeling systems that can elucidate these solutions and the implications of choosing one versus another (Van Delden et al., 2011; Voinov and Bousquet, 2010).

2.3. Education, peer review, and reuse

A review of current applications shows a wide range of approaches for designing, executing, and documenting applications, making it difficult to understand, review, and reuse applications.

Table 2
Examples of IEM applications.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Context</th>
<th>Examples and references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interdependent relationships among multiple stressors, multiple environmental compartments, and multiple endpoints</td>
<td>San Joaquin River Deep Water Ship Channel, Stockton, CA USA</td>
<td>Jassby and Cloern (2000), Lehman et al. (2001), and Quinn and Jacobs (2006)</td>
</tr>
<tr>
<td>Venice Lagoon, Italy</td>
<td>Life-cycle analyses of chemical-based stressors at a global scale</td>
<td>Sommerfreund et al. (2010), Sleeswijk and Heijungs (2010)</td>
</tr>
<tr>
<td>Pinios River, Greece</td>
<td>Groundwater-surface water flooding in United Kingdom basins</td>
<td>Makropoulos et al. (2010), Hughes et al. (2011)</td>
</tr>
<tr>
<td>Ecological applications focused on decision/policy objectives with alternative management strategies</td>
<td>Quantifying the trade-offs among ecosystem services in complex, dynamic systems, Agent-based modeling of land use and land cover, Willamette River Basin, Oregon, USA Unified metamodel of the biosphere</td>
<td>Farber et al. (2006), Bohte et al. (2006) and Guzy et al. (2008), Hulse et al. (2008), Boumans et al. (2002), See also: Maxwell and Costanza (1995), Daniels (1999), Noth et al. (2000), Costanza et al. (2002), Sengupta and Bennett (2003), and Schaldach et al. (2011)</td>
</tr>
<tr>
<td>Applications involving a diverse set of stakeholders</td>
<td>Multi-criteria integrated resource assessment with science and decision stakeholders</td>
<td>Stahl et al. (2011, 2002)</td>
</tr>
</tbody>
</table>
Rouwette et al., 2002 review the system dynamics literature in an attempt to characterize the effectiveness of such stakeholder driven model building exercises. They found a similarly wide variation in approaches and present guidelines for reporting the process and assessing effectiveness. To improve this situation for IEM, there is a need to move toward conceptual standardization of the application process (i.e., defining and documenting the elements of IEM applications according to a community recognized process and set of practices).

IEM applications are resource intensive, and thus, the ability to reuse an application, in part or in whole, can result in significant resource savings and more problems and decisions served. To achieve reusability of applications will require formal documentation and archiving strategies that preserve not only the software technology utilized (data, models, etc.) but also the expertise that sets up, executes, and interprets the results of the modeling system. Assumptions and model parameterization schemes should be documented in machine readable formats. Janssen et al., 2009 discuss these issues and propose the use of assessment project ontologies for describing and documenting scenarios and assessments. This level of documentation and transparency is also necessary to facilitate quality assurance and peer review.

3. IEM science

The science of IEM provides the knowledge and integrative strategies that support the decision process. In conducting the IEM process, scientists do not directly pursue new knowledge within individual disciplines, but rather concern themselves with issues that arise when domain-specific knowledge bases must be integrated. The goal is to construct and apply systems-based approaches to explore, explain, and predict system response to changes in natural or managed environmental systems. Workshop discussions related to IEM science focused on several areas including holistic systems thinking and integrated modeling, data, model evaluation, and peer review. In the following sections we discuss these topics from the perspective of current practices, issues, and challenges that lead to the IEM science roadmap presented in Fig. 3.

3.1. Holistic systems thinking and integrated modeling

At the core of IEM science is the concept of holistic thinking (i.e., assessing a problem in the context of the larger system — of systems — within which it occurs). This systems approach is necessary to serve the decision makers' needs to understand the working system, compare impacts among decision scenarios, analyze trade-offs among options, ask “What if?” questions, avoid the creation or transfer of problems in pursuing solutions to the problem at hand, adapt strategies based on future monitoring of the system, and respond to unintended consequences.

A primary challenge for IEM is the merging of knowledge domains into coherent and appropriately complex representations of the relevant system (Kragt et al., 2011; Lancaster, 2007; Otto-Banaszak et al., 2011; Voinov and Bousquet, 2010; Voinov and Gaddis, 2008; Zagonel, 2002). Coherence exists when modeling components are scientifically consistent across the system with respect to complexity, data requirements, and uncertainty (EPA, 2008b). Complexity is a direct function of the problem statement, decision objectives, system understanding, and data availability. It is not possible (and unnecessary) to include all known science related to the social, economic, and environmental disciplines (Liu et al., 2008; Oreskes, 2003). The challenge is to determine which of
the detailed processes are important in simulating the system behavior at an appropriate scale of application (Sidle, 2006). These challenges must be addressed in a consistent manner across each step of the modeling process, beginning with the formulation of a problem statement and followed by the system conceptualization, an integrated modeling methodology, and the synthesis of the modeling results (Hinkel, 2009; Liu et al., 2008). We discuss each of these modeling steps to emphasize the manner in which IEM issues and challenges manifest and also to point out that to efficiently share IEM science products across the global community we will need to be more explicit and compartmentalized in our implementations.

A problem statement is a question that requires the application of a structured approach to a solution. A problem statement contains a targeted interest or concern, context, objectives, questions to be answered, scenarios, solution criteria (i.e., tolerance for uncertainty), and available resources (Arnold, 2013, presents an interesting discussion of this topic from the perspective of the resource manager). The purpose of the problem statement is to provide the information needed to guide the subsequent steps of an IEM application. Decision stakeholders are primarily responsible for the content of the problem statement and the science stakeholders must ensure that the content is sufficiently focused and detailed to achieve its purpose. Currently, there are no widely accepted protocols for developing and documenting problem statements. A challenge for IEM is to establish appropriate guidelines for defining and documenting the full expression of a problem statement.

System conceptualization captures the essence of the real-world problem including the processes, cycles, and flows that characterize the relevant social—economic—environmental components of the system (Fischenich, 2008). System conceptualization serves as a basis for communication between decision and science stakeholders. Several methods for creating and documenting a conceptual model have been developed. Luna-Reyes (2003) presents examples of mapping tools used to graphically represent dynamic systems. Other methods, based upon specific quantitative modeling approaches, include systems dynamics (Boumans et al., 2002; Fenner et al., 2005; Muetzelfeldt and Masheder, 2003), fuzzy cognitive mapping (Özesmi and Özesmi, 2004; Samarasinghe and Strickert, 2013), and Bayesian inference (Reckhow, 2003).

Despite the availability of methods and evidence that engaging stakeholders in the system conceptualization process is a growing priority (Voinov and Bousquet, 2010) their widespread use across the community of IEM remains an issue. There is a need to promote best practices with respect to the social process of eliciting and merging the array of stakeholder mental models.

The integrated modeling methodology is a combination of a modeling system and an implementation strategy. The modeling system represents an integration of data and knowledge from across relevant science domains and represents the quantitative and computational form of the conceptual model. The mathematical form of the modeling system may be empirical, statistical, process-based, or a combination of the above (Linker et al., 1999; Hart et al., 2009; Schwarz et al., 2006). The implementation strategy
specifies how the modeling system will be deployed in the context of an IEM application. Deployment may include such strategies as applying the modeling system to representative locations across a regional or national landscape (Marin et al., 2003), executing the modeling system within a Monte Carlo simulation protocol to address uncertainty (Johnston et al., 2011), or applying the modeling system repeatedly within an adaptive management strategy (Akbar et al., 2013). Multiple implementation strategies may be applied with the same modeling system.

Key science aspects of these integrated systems are ensuring the conceptual compatibility among the components (ontology) and specification of the information to be exchanged between components (semantics). Achieving semantic and ontological consistency is particularly challenging for IEM system design due to the trans-disciplinary nature of components and the common practice of linking existing modeling components (e.g., legacy models) not originally designed for such integration. Voinov and Cerco (2010) and Voinov and Shugart (2013) discuss these issues and point out several challenges and potential pitfalls regarding the construction of IEM systems.

No general guidelines, best practices, or standards exist for defining and harmonizing the semantic and ontological information. In practice, these issues are resolved either implicitly or explicitly by the development team responsible for the integration of modeling components. For example, Akbar et al. (2013), in building an emergency response modeling system, select a fixed set of models on the basis of ontological consistency and specify a mapping of variables from one model to another. FRAMES (Johnston et al., 2011) defines data dictionaries that contain semantics that reflect controlled vocabularies and relationships among the variables. Each dictionary represents a standard set of information either produced or consumed by a type of model (e.g., watershed model) and thus allows for application specific configuration of models. Ontological consistency is enhanced with the data dictionaries but not ensured. Finally, SEAMLESS (van Luttersum et al., 2008) defines formal expressions of the combined semantics and ontologies associated with the modeling components needed to construct workflows for agricultural based modeling assessments. Any modeling component that conforms to these definitions can be used in the workflows with assurance of both semantic and ontological consistency.

The challenge for the IEM community is to expose and standardize the model integration process and explicitly express a model’s semantics and ontology, thus facilitating the ability to interoperate with a wider array of available models. To guide this movement toward higher degrees of interoperability, we look to Wang et al. (2009), who describe six levels of interoperability (Table 3) in order of increasing capacity for interoperation. Differences between these levels reflect the type and content of information to be exchanged, not the technology that implements the exchange. While some standards are available for expressing this information in software technology, until this science-based information is defined and standardized, it will not be possible to address the technology issues of software reuse and interoperability, which we will discuss further in Section 4.

The final step of the modeling process, synthesis of modeling results, represents a key interface between science and decision stakeholders. The objective is to interpret, consolidate, and present the results of complex integrated modeling to stakeholders and decision makers. Synthesis must produce information that is not only of high scientific quality but also useful to the decision makers. McNie (2006) discusses the challenges of “reconciling the supply and demand” of scientific information between scientists and decision makers, defining “useful” as a combination of salient, credible, and legimate information. Salience implies contextual relevance, credibility refers to scientific veracity, and legitimacy refers to a lack of bias. While scientists routinely synthesize results via scientific journals, the synthesis of scientific information for decision making is not well understood and executed (NRC, 2005). The challenge for IEM is to serve diverse users and stakeholders who may require multiple synthesis streams designed hierarchically to move seamlessly from very general displays of overall results to highly detailed, component-based visualizations (Ellarby and Kite, 2006; Liu et al., 2008).

Marin et al. (2003) and Babendreier and Castleton (2005) provide an example of a successful synthesis involving the application of 17 science-based models to predict national-scale human and ecological exposure and risk due to chemical releases from waste disposal facilities. The effort included hundreds of thousands of individual simulations and resulted in output too voluminous to store, much less hand to the decision makers. A database of modeling results was constructed along with a graphical user interface to allow decision makers to ask and receive answers to very specific policy questions related to risks, protection levels, human versus ecological impacts, etc. Identifying and promoting exemplars will help focus attention on this important science issue.

### 3.2. Data for IEM

Environmental, social, and economic data drive model development, application, and evaluation. Discovering, accessing, processing, and preparing data for IEM tasks is particularly challenging due to a combination of cross-disciplinarity, volume, and disparate sources presenting data with varying formats and semantics. Fig. 4 describes an integration framework that illustrates the typical data and processing needs for an agency like the US Food and Drug Administration. In this example, data reflecting biological levels ranging from molecular to population are contained in a series of databases owned and maintained by a variety of institutions. Viewed generically, this example illustrates the IEM data challenge described above.

Recognizing these issues and the importance of data to decision making, many government agencies and offices are consolidating access to environmental data. Table 4 summarizes several such efforts from the United States, Europe, and Australia.

As seamless access to data becomes available, the next challenge is to process and transform the data for use in IEM systems. The gateways, like IEM systems, are designed with internal semantic and ontological consistency but not for seamless integration across systems. This data integration task requires reconciliation of varying semantics and establishing a system’s level operational ontology that honors the relationship among physical, chemical, and biological entities across the components. This may include such procedures as statistical processing (e.g., averaging, interpolation, etc.), geo-processing (e.g., re-projection, clipping overlays, merging, etc.), and processing specific to physical interfaces between elements of the modeling domain (e.g., catchment and

### Table 3

<table>
<thead>
<tr>
<th>Level of interoperability</th>
<th>Information defined</th>
<th>Content defined</th>
</tr>
</thead>
<tbody>
<tr>
<td>L6: Conceptual</td>
<td>Assumptions, constraints, etc.</td>
<td>Documented conceptual model</td>
</tr>
<tr>
<td>L5: Dynamic</td>
<td>Effect of data</td>
<td>Effect of information exchanged</td>
</tr>
<tr>
<td>L4: Pragmatic</td>
<td>Use of data</td>
<td>Context of information exchanged</td>
</tr>
<tr>
<td>L3: Semantic</td>
<td>Meaning of data</td>
<td>Content of information exchanged</td>
</tr>
<tr>
<td>L2: Syntactic</td>
<td>Structured data</td>
<td>Format of information exchanged</td>
</tr>
<tr>
<td>L1: Technical</td>
<td>Bits and bytes</td>
<td>Symbols of information exchanged</td>
</tr>
<tr>
<td>L0: None</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
stream segment connectivity). Currently, the execution of this process for individual applications can be characterized as a semi-automated task of discovering data sources and subsequently cutting and pasting file fragments to form a single coherent dataset.

Enhanced solutions to this issue are beginning to emerge. The GEON3 (Ludascher et al., 2003) is an open collaborative project funded by the US National Science Foundation to develop cyberinfrastructure with the ultimate goal of linking heterogeneous scientific data and information for the purposes of knowledge discovery, sharing, and integration into scientific workflows. Data for Environmental Modeling (D4EM – Johnston et al., 2011) is an open source software system developed expressly to access, retrieve, and process data for IEM. D4EM is currently linked to several US Government databases and performs all data processing required to serve data directly to an integrated modeling system designed to simulate interactions among watersheds and aquatic ecosystems. CUAHSI (Maidment et al., 2009) has designed and implemented standards for exchanging hydrologic data over the web including the WaterOneFlow web services and the Water Markup Language (WaterML). These were used to construct a Hydrologic Information System with a service-oriented architecture and able to integrate hydrologic observational data from international, federal, state, local, and academic data providers (Tarboton et al., 2011). The design principle is to keep databases separate and autonomous while providing standards that allow software programs to query and extract data from them using standardized approaches (Goodall et al., 2008; Horsburgh et al., 2009). Other important efforts include the development of controlled vocabularies that promote common referencing and facilitate variable matching across sources. Examples include the NetCDF Climate and Forecast (CF) Metadata Convention4 and the CUAHSI – HIS Controlled Vocabulary5 for hydrology data (Beran and Piasecki, 2009). Applying and integrating these techniques and technologies to the larger data domains of IEM is an important need.

3.3. Model evaluation

Model evaluation combines quantitative and qualitative information about a modeling system’s appropriateness and effectiveness for the problem and ability to characterize the uncertainty of model predictions. Key attributes of model evaluation are transparency, refutability, and uncertainty quantification. Together, they establish the scientific veracity and stakeholder confidence/acceptance of an application and the information it produces. Transparency requires that all aspects of the application design and execution be accessible to facilitate understanding and reproducibility. Refutability requires a hypothesis-testing framework in which data are used in specific ways to test the model’s ability to simulate the system of interest. This may require the involvement of decision stakeholders because the ultimate test of a model is always its utility and usability by end users. Refutability is difficult for any model of an environmental system, and evaluating an IEM system is even more challenging.

Challenges related to uncertainty quantification of predictions in integrated modeling were experienced in recent climate change modeling (e.g., IPCC, 2007). In Table 5, we describe basic model evaluation methods developed for conventional modeling that have characteristics advantageous to IEMs. Several textbooks on these subjects have been produced in recent years (Menke, 2012; Aster et al., 2005 from geophysics; Beven, 2009; Clark et al., 2011; Hill and Tiedeman, 2007 from hydrology; Saltelli et al., 2008 from econometrics), and an active scientific community continues to explore new methods. Further, Matott et al. (2009) note that a great

---

3 http://www.geongrid.org/
4 http://cf-pcmdi.llnl.gov/
5 http://his.cuahsi.org/mastercvdata.html
deal of literature has been published on model evaluation ranging from introductory descriptions to uncertainty analyses to methodological applications. They cataloged 65 different model evaluation methods including data analysis, identification tools for applicability across seven thematic model evaluation applications. The stated purpose is to assist in environmental policy making across national boundaries. The spatial information considered includes seventeen topical and technical themes originating from numerous sources throughout the European Union. A web-based system where public information providers share environmental data and information. In implementing SEIS, the EEA is building on existing reporting systems and tools: the INSPIRE directive, Global Monitoring for Environment and Security (GMES), and (GEOSS).

### Table 4: Examples of data access initiatives of relevance to IEM.

<table>
<thead>
<tr>
<th>Initiative</th>
<th>Scope</th>
<th>Description</th>
</tr>
</thead>
</table>
| Global Earth Observation System of Systems (GEOSS)
| Infrastructure for Spatial Information in the European Community (INSPIRE)
| European Shared Environmental Information System (SEIS)
| Water Information Service for Europe (WISE)
| Environmental Resources Information Network (ERIN)
| Environmental Dataset Gateway (EDG)
| National Ecological Observation Network (NEON)
| US based Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) | Global   | An international effort to coordinate a comprehensive monitoring of the state of the earth, study large scale processes, and predict behavior of the earth system. |
| EU                                                              |           | The stated purpose is to assist in environmental policy making across national boundaries.                                                      |
| EU                                                              |           | A web based system where public information providers share environmental data and information. In implementing SEIS, the EEA is building on existing reporting systems and tools: the INSPIRE directive, Global Monitoring for Environment and Security (GMES), and (GEOSS). |
| EU                                                              |           | Gateway to information on European water issues. It comprises a wide range data and information collected by EU institutions to serve several stakeholders. |
| Australia                                                       |           | The objective is to organize environmental information from many sources and provide standards based tools for discovery, access, and use. The information includes maps, species distributions, documents and satellite imagery, and covers environmental themes ranging from endangered species to drought and pollution. |
| US                                                              |           | A gateway developed by the US EPA to web-based information and services. It enables data consumers to discover, view and access data sets, as well as geospatial tools. Users also have the ability to catalog and maintain their geospatial metadata contributions via the EPA Metadata Editor Tool. |
| US                                                              |           | Collection of data across the United States on the impacts of climate change, land use change and invasive species on natural resources and biodiversity. Designed to detect and enable forecasting of ecological change at continental scales over multiple decades. |
| US                                                              |           | Developing the Hydrologic Information System (HIS), an internet-based system that provides for sharing hydrologic time series data contributed by a wide range of providers, including the National Water Information System of the US Geological Survey. |

The objective is to organize environmental information from many sources and provide standards based tools for discovery, access, and use. The information includes maps, species distributions, documents and satellite imagery, and covers environmental themes ranging from endangered species to drought and pollution. Designed to detect and enable forecasting of ecological change at continental scales over multiple decades.

### 3.4. Peer review

An important consideration discussed during the workshops is the challenge related to peer review of IEM integrated science as expressed through technology and applications. The interdisciplinary nature of the modeling challenges presents challenges to individual peer reviewers representing a particular science domain. Certainly, this level of review is necessary; however, there is concern whether it is sufficient with respect to integration issues, especially given the implicit manner in which integration issues are resolved and documented. Peer review of applications involving a wide array of stakeholders, each capable of varying degrees of understanding, also represents a challenge. Each application should be reviewed from the perspective of each stakeholder knowledge base and perspective. Finally, the peer review of implementation technologies represents a significant time resource requirement. Verifying that the science has been implemented correctly in software presents significant issues in an IEM world, where the technological implementations vary widely in terms of design, software/hardware, documentation, and testing. New means of ensuring the veracity of the science-based products and applications represents a prime challenge to the IEM science community.

### 4. IEM technology

Technology represents the primary means by which the science of IEM is expressed and applied.

In this section we present two key topics that emerged as critical technological drivers for IEM from the workshop discussions. First is a discussion of modeling frameworks and standards for IEM software design and implementation. This Section (4.1) provides background and context to inform the reader about the state-of-the-art, issues, and challenges associated with IEM modeling frameworks. It concludes with an argument for a universal standard for model integration that is compatible with framework-specific standards that already exist, but provides much needed interoperability across modeling frameworks. Second is a discussion of leveraging the World Wide Web for IEM. This Section (4.2) presents modern and visionary work using concepts such as Cloud-based computing and web services to achieve a higher level of functionality in next generation IEM modeling frameworks. The section argues that a key goal of the IEM community must be to more effectively leverage the Web for publication, discovery, access, and integration of IEM information and software in order to achieve the ambitious goals set by the IEM community. Fig. 5 presents the IEM technology roadmap, whose elements are discussed in the following sections.
Interoperability in this sense means the ability of different information technology components, systems, and software applications to communicate and exchange data accurately, effectively, and consistently, and to use the information that has been exchanged (Heubusch, 2006; IEEE, 1989). Thus, focusing on individual science models, the interoperability challenge is to enable communication and exchange of data between two scientific models that may be from different scientific domains. This problem is multifaceted, and we discussed the issues of coherence, complexity, semantics, and ontologies related to model integration as scientific goals for IEM in Section 2. Therefore in this section we focus the discussion on technological issues in achieving interoperability between scientific models.

Matott et al. (2009) note several software technology-based barriers to interoperability, such as different programming languages, compilers, and development platforms; inconsistent separation of system and model components (e.g., user and model interface code, executables, algorithmic code, execution management code, warning and error handling, and statistical functionality); and different input/output (I/O) file formats. Collectively these heterogeneities complicate the process of IEM because they limit the number and variety of tools available for integrated modeling and assessment. As a result, small communities have evolved, each with its own modeling framework and internal standards for integrating model components. Table 6 provides a list of some of the modeling frameworks that exist within the IEM community.

Many of the modeling frameworks listed in Table 6 have advanced over several years of development effort to become sophisticated tools. Many are also widely used tools within segments of the IEM community. The adoption of specific modeling frameworks within local communities is understandable and unlikely to change in the near future because maintaining local control over the user experience, in particular the design and implementation of the Graphical User Interface (GUI), is important for buy-in and effective use within specialized communities. The protocols and standards employed locally by individual frameworks to facilitate interoperability, while important within the framework itself, do not directly address the challenge of achieving interoperability across frameworks. This may be counterintuitive, but when one considers the vast variety of ways for achieving interoperability between models, it becomes reasonable that no two modeling frameworks have independently settled on the same standard for achieving interoperability. As stated earlier, this lack of cross-framework interoperability is a significant technical challenge facing the IEM community because, even though different frameworks may focus on different problem domains, the science (data, models, and methods) expressed within each framework is often the same. The frameworks themselves are therefore repetitive causing additional work in terms of code development and maintenance. But it is not only this additional work that is a cause for concern; more important is the fact that modeling frameworks, because they tend to focus on specific problem domains, do not always include the state-of-art scientific models for problem domains that are tangential to their own area of focus. For example, a groundwater model must include a way of modeling river hydraulics to provide a boundary condition to the subsurface environment, but this is not the primary focus of the groundwater model, so the tangential river hydraulics code is less likely to be kept up to date compared to the core groundwater code. Therefore the need for interoperability across frameworks is deep and far reaching and a solution to this problem would be a great benefit to the IEM vision.

Table 6
Summary of methods used for model evaluation.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error evaluation and propagation</td>
<td>Starts with data used to construct the component and IEM models, and proceed to the analysis of predictions relative to observations. This is critical to IEMs because data availability and errors in one part of the system invariably affect uncertainties in other parts.</td>
</tr>
<tr>
<td>Error-based weighting</td>
<td>Critical to integrated use of data in IEMs; is complicated by need to synchronize weights among various disciplines since variations in the importance of processes and data may occur.</td>
</tr>
<tr>
<td>Sensitivity analysis</td>
<td>This can be conducted by combining computationally efficient local linear methods, efficient global screening methods, and computationally expensive global variance methods. Linear methods are attractive for IEMs with relatively few parameters because they may require only on the order of 10s of model runs to obtain useful results.</td>
</tr>
<tr>
<td>Alternative models</td>
<td>Important in IEM model development to assess effects of conceptual model uncertainty. Alternative conceptualizations often affect more than one aspect of the system. Integrated models allow consequences to be represented realistically throughout the system simulation (e.g., the hydrologic cycle, Clark et al., 2011). These are used to improve objectivity and reproducibility. The optimization process can identify the information content of a given set of observations and, along with sensitivity analysis methods, can identify important new data. For IEMs, this can be particularly challenging because of the number of linked/coupled component models involved.</td>
</tr>
<tr>
<td>Automated calibration methods</td>
<td>IEM models require both knowledge and uncertainty from different system components to be integrated into a unified expression of uncertainty quantification. Existing methods should be reviewed and applied considering the computational demands of the model and requirements of end users of model results.</td>
</tr>
<tr>
<td>Uncertainty Quantification</td>
<td>In IEM, it is important to test component models; integration requires additional tests throughout the IEM.</td>
</tr>
<tr>
<td>Model Tests</td>
<td>Should be undertaken against alternative data sets.</td>
</tr>
<tr>
<td>Post-Audits</td>
<td>Comparing model predictions to the observed results (a true post-audit) requires monitoring. IEMs provide important opportunities for post-audits because simulated results can affect resources important to large ecological systems and many people. These opportunities are not easily pursued however, as post-audits of IEMs can also be very difficult because of long delays in observing impacts, confounding variables, changes in forcing drivers etc.</td>
</tr>
<tr>
<td>Calibration and testing in data scarce conditions</td>
<td>The likelihood of having complete data sets for extensive IEM analysis is low. Building representative data sets and uncertainty analyses have to be performed in data-scarce conditions. The models and theories can suggest specific monitoring to collect the most important data to help decrease uncertainty and facilitate adaptive management strategies.</td>
</tr>
</tbody>
</table>

4.1. Modeling frameworks and standards for IEM software design and implementation

Conventional environmental modeling systems include science models, user interfaces, data analysis and visualization tools (including GIS), and calibration and optimization tools. Within this ecosystem of software tools required to perform IEM, there is a strong need to provide interoperability between tools to simplify and automate data transfer across applications. Not only is interoperability required across the tools used for environmental modeling, but at a deeper level interoperability is required between the individual science models used to address specific environmental concerns. It is this interoperability — between individual science models used in IEM — that has attracted much of the attention within the community because of the inherent challenges of properly translating and transferring knowledge between multiple science domains.
Given these technical challenges, we believe a near-term goal of the IEM community should be to work on protocols and standards that are appropriate to be elevated to a level higher than individual frameworks, thus facilitating access to a much wider inventory of models and components. This challenge is gaining recognition and progress toward increased interoperability is occurring. Workshop participants agreed that, to date, groups involved in modeling framework development have "discovered by doing" that the technological issues described above are common across frameworks and that the required functionality can be abstracted and standardized at a higher level, i.e., a global standard. For example, it is clear that a core set of properties that each model within a modeling system must follow exists. These properties include a structure that enables the modeling system to initialize, execute (e.g., step through time and update state), retrieve and provide data to other models, and close the model on demand through a standardized Application Programming Interface (API). Models that can provide these interface functions are able to provide their caller with fine-grained control of their functionality, which is a key step to achieving interoperability across frameworks. In effect, this process of standardization is one of separating framework functionality from the science components contained within them, rendering the components framework independent. If well designed, the standardization process can be done in a way that minimizes potential negative impacts including placing unreasonable burdens on scientists or inhibiting creativity due to the need to adhere to onerous standards.

To move forward, the IEM community would be well served to look at examples of past work that offer more generic solutions for interoperability across modeling components. One example is the Common Component Architecture (Larson et al., 2004), which is a set of component and framework standards developed within high-performance, scientific computing. CCA-compliant components can be reused in any CCA-compliant framework (e.g., Caffeine, XCAT). CCA is used as the underlying architecture for modeling frameworks such as the CSDMS (Peckham et al., in press). Another example is from the OpenMI Association, which has proposed a global standard for exchanging data among linked models at run time (Moore and Tindall, 2005; Moore et al., 2005). The OpenMI standard has been the subject of many recent studies and movement of existing frameworks to accommodate the standard is occurring (Fotopoulos et al., 2010; Castronova and Goodall, 2010; Elag et al., 2011; Betrie et al., 2011; Janssen et al., 2011; Bulatewicz et al., 2010; Ewert et al., 2009; Reussneret al., 2009).

On a more general level, the challenge of interoperability present in IEM is similar to interoperability issues addressed by the World Wide Web, which would not be possible without broad agreement on standards for data and information (knowledge) storage and exchange (e.g., HTTP, HTML, etc.). The World Wide Web Consortium (W3C) and the Open Geospatial Consortium (OGC) are international consortia involving companies, government agencies and universities, committed to a consensus process for development of standards that empower development of a vast array of applications. The W3C pursues open standards related to every aspect of the web, from its basic architecture to the provision

---

Fig. 5. IEM technology roadmap.

---
of data and information storage and access services, to enabling web functionality on all manner of devices. One path forward for IEM is to embrace and build from standards and technologies for representing structured data and the Semantic Web technologies for representing knowledge and linking information sources (e.g., RDF, SPARQL, OWL, and SKOS). The OGC focuses on standards that facilitate access to and use of spatial information and related services (e.g., WFS, WPS, WCS, SOS, WMS), as will be discussed further in Section 4.2. Other standards relevant to IEM include the development of several science domain markup languages based on XML, including the GeographyML, the Earth ScienceML, the NetCDFML, and the Systems BiologyML. Examples of ontology applications based on OWL include the Semantic Web for Earth and Environmental Terminology (SWEET), the Environment Ontology (EnvO), and the Exposure Ontology (ExO). Each of these applications focuses on a particular science domain. The challenge for the IEM community is to bring these concepts and standards to bear on IEM systems in an effort to establish a unifying publishing capability for software that facilitates discovery and utilization of individual IEM components and systems independent of the source of their development.

4.2. Leveraging the World Wide Web for IEM

A second theme that emerged from the workshops concerns the leveraging of the World Wide Web for building next generation IEM modeling systems. Clear momentum exists in the broader information technology domain toward storing data and tools in the Cloud. Much work has been done in IEM communities to create Web-based analysis tools and portals (e.g., Booth et al., 2011), expose large databases as web services (e.g., Goodall et al., 2008), and for creating workflows to coordinate data flow between databases, analysis tools, and models (e.g., Granell et al., 2010; Kepler, 2012). Recent work has focused on service-oriented and resource-oriented paradigms for organizing model software architectures suggesting that model frameworks themselves could be integrating computational and data resources that are distributed across the Web (Goodall et al., 2011; Nativi et al., 2013; Granell et al., 2013). Commercial investment in cloud-based computing resources, which allows users to rent computing resources, opens new doors for dynamically-scaling compute intensive tasks or web applications with temporarily high demands, as discussed later in this section. Collectively, these Web-based initiatives offer potentially transformative changes to the technological approaches available to the IEM community, but much work is needed to understand how to effectively and efficiently leverage these approaches for specific applications within IEM.

A key challenge to the IEM community for achieving the full potential of the web is to advance our understanding of how to optimize data and operations between traditional personal computer (PC) environments and remote computers on the Web. This is because, while the Web is a promising tool that could be better leveraged in IEM applications, there remain issues that must be addressed. For example, IEM applications often require the use of large volumes of data, and moving these data sets effectively and efficiently over the Web is challenging. Second, in some cases IEM workflows may require dynamic and complex interaction between components of the workflow. An example might be coupled model components within a workflow that have a time-step dependent feedback loop for data exchanges. The challenge facing the IEM community is how to allow for such functionality while still maintaining sufficient reliability and serviceability of the IEM software systems.

Moving forward, the community should be aware of the different ways in which the Web can be leveraged when building IEM software systems. In one scenario, the entire IEM solution

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHEM</td>
<td>Flexible landscape model structures, easily modified or extended for different goals.</td>
<td>Voinov et al. (2004)</td>
</tr>
<tr>
<td>OMS</td>
<td>Provides the ability to construct models and applications from a set of components.</td>
<td>David et al. (2002, 2013) and Ahuja et al., 2005</td>
</tr>
<tr>
<td>MIMOSA</td>
<td>A model simulation platform for building conceptual models and running the simulations.</td>
<td>MÜller (2010)</td>
</tr>
<tr>
<td>FRAMES</td>
<td>A modeling system that includes a collection of models as well as data retrieval and analysis tools.</td>
<td>Johnston et al. (2011) and Babendreier and Castleton (2005)</td>
</tr>
<tr>
<td>SHEDS</td>
<td>A modeling system for simulating human activity patterns and related chemical exposures.</td>
<td>Zartarian et al. (2012)</td>
</tr>
<tr>
<td>ARAMS</td>
<td>Used to estimate impacts and risks associated with military relevant contaminates.</td>
<td>Dorch et al. (2007)</td>
</tr>
<tr>
<td>GENIE</td>
<td>For calculating radiation dose and risk from radionuclides released to the environment.</td>
<td>Napier (2007)</td>
</tr>
<tr>
<td>IWBS</td>
<td>Integrates a collection of water resource models (watersheds, rivers, lakes, estuaries) to support decision makers.</td>
<td>Thurman et al. (2004)</td>
</tr>
<tr>
<td>AMBER</td>
<td>Designed to explore the benefits of making scientific modeling tools available on the internet.</td>
<td>Quintessa (2012)</td>
</tr>
<tr>
<td>GoldSim</td>
<td>For dynamically modeling complex systems including but not limited to IEM systems.</td>
<td>GoldSim (2012)</td>
</tr>
<tr>
<td>GMS/WMS/MS</td>
<td>Groundwater, watershed, and surface water modeling systems.</td>
<td>Aquaveo (2012)</td>
</tr>
<tr>
<td>BASINS</td>
<td>Integrates modeling and assessment tools with national watershed data using a GIS</td>
<td>EPA (2001)</td>
</tr>
<tr>
<td>ESMF</td>
<td>For building climate and weather prediction models as interlinked components.</td>
<td>Hill et al. (2004)</td>
</tr>
<tr>
<td>CSDMS</td>
<td>Component-based modeling framework targeting the earth surface dynamics community.</td>
<td>Peckham (2010)</td>
</tr>
<tr>
<td>SEAMLESS</td>
<td>Integrated framework for linking models, data, and indicators in support of environmental, economic and social analysis for agricultural systems.</td>
<td>van Ittersum et al. (2008)</td>
</tr>
<tr>
<td>HydroModeler</td>
<td>Integrated modeling environment plug-in to CUAHSI HydroDesktop application and built on OpenMI standard.</td>
<td>Castronova et al. (2013)</td>
</tr>
<tr>
<td>ARIES</td>
<td>Tool for assessing and validating ecosystem services in decision-making.</td>
<td>ARIES (2012)</td>
</tr>
<tr>
<td>EvoLand/ENVISION</td>
<td>Regional planning and environmental assessment tool; spatially explicit and multi-agent based.</td>
<td>Bolte et al. (2006) and <a href="http://envision.bioe.vt.edu/">http://envision.bioe.vt.edu/</a></td>
</tr>
</tbody>
</table>

**Table 6**

A sampling of the modeling frameworks with connections to the IEM community.
might be hosted on a single web server and the user would then interact with the application through a Web browser. This is the typical solution and has been widely leveraged for providing data, visualization tools, and basic analysis tools in the IEM community. A second scenario is for the IEM solution to be a Desktop application that has built-in capabilities for leveraging remote data or processing resources directly through the Desktop application itself. In this case, the remote resources would ideally be made available to the Desktop application as web services using a public and well-described API. An example of this approach is the CUAHSI HydroDesktop application that provides access to remote data archives made available using the CUAHSI WaterOneFlow web service (Ames et al., 2012; Tarboton et al., 2009). In HydroDesktop, the functionality for both searching and downloading data is executed on remote servers that have their own databases. The software architecture involves a network of servers, with each server having its own database and software stack that allow it to be a node within the network (Horsburgh et al., 2009; Horsburgh et al., 2011). The user executes data search and download tools directly from the HydroDesktop application instead of through a Web browser. Because the web services remote servers have been built using a standardized API, it is possible to ingest data from any server that adopts the API into a local database on the machine running HydroDesktop. This local data is then available for performing data analysis, visualization, and modeling activities.

This approach is in the spirit of ‘Cloud Computing’, although recent investment by the public and private sectors has grown the Cloud computing concept into a powerful new paradigm for building Web systems. A key idea of the cloud computing concept is that a user can rent computer resources (processing, storage, software) from a vendor rather than buying and maintaining their own computing resources. Commercial-based cloud services are offered from a range of providers, e.g., Amazon, Google, and Rackspace. The advantage is that a user has access to a theoretically unlimited amount of computing resources to accomplish a task. Therefore applications can more dynamically scale as more or less resources are needed over time.

Governmental agencies have begun exploring cloud-computing and offering tools as Web resources made available through the Cloud. The British Geological Survey has initiated the Environmental Virtual Laboratory project and is exploring the provisioning of data services, web-enabled environmental models, and a suite of on-line local community tools in the spirit of the Cloud paradigm of software as a service. The US Department of Agriculture Natural Resources Conservation Service is developing the Cloud Services Innovation Platform to offer data and modeling services for use in the field, and the US Environmental Protection Agency has developed the WATERS program that provides services that perform various data services and related analyses like watershed delineation. These applications suggest a certain momentum in the community toward moving more of the tasks needed to support IEM, such as running environmental simulation models or large databases, to remote computer servers on the Cloud rather than on PCs.

The approach of using cloud computing and web services has several advantages over traditional modeling approaches, including the potential to greatly reduce the cost and time required for the development of an IEM solution through the reuse of modeling components. Each component in the modeling chain can be rapidly distributed to the community and sharing of services can be a catalyst for building a stronger integrated environmental modeling community. Furthermore, the process of designing services would break down the different aspects of environmental modeling into a set of interoperable services that can be dynamically configured to create custom solutions to environmental problems. Another advantage is that moving resource intensive tasks to servers opens the possibility of performing sophisticated IEM operations on mobile devices as well. Such tools could have large benefit in engaging stakeholders, and for this reason the US Environmental Protection Agency issued a challenge for the public to develop environmental applications for mobile devices and the British Geological Survey has developed the iGeology application that combines GPS functionality with informational databases related to geology across Britain.

Despite access to these new tools and resources available in the Web-domain, the fundamental challenge of establishing common standards for data and modeling services outlined in Sections 2.1 (science content) and 4.1 (technology) are only amplified as one attempts to capture the potential of the web for providing interoperability across a wide community of end users. The IEM community would be well served by identifying how existing science and technology standards can be leveraged, integrated, and extended to serve the broader needs and interests of IEM. As stated several times in this paper these standards must be guided through an approval process that includes sufficient representation from the community. Without such effort, the IEM tools built for the Web will suffer from the same interoperability challenges faced by modeling frameworks built for Desktop environments.

5. IEM community

In many fields and sectors, openness, collaboration, sharing, and social learning have been shown to drive innovation and growth (Tapscott and Williams, 2006). These behavioral attitudes and characteristics are often expressed through formal communities of practice (Lave and Wenger, 1991). Structured community processes can reduce duplication of efforts and increase leveraging of resources and overall efficiency. IEM is transdisciplinary and, as such, involves a “community”. The members of the community include the full array of decision and science stakeholders described earlier. During the workshops, discussions of community focused on establishing and promoting a community of practice, IEM education, and a web-based community center. The resulting roadmap of IEM community activities is presented in Fig. 6 with related workshop discussions of the specific topics in the following sections.

5.1. Communities of practice

The early phases of IEM development were conducted by disparate groups working on science and applications (e.g., EPA, 1992; Onishi et al., 1985; Whelan and Nicholson, 2002; Whelan et al., 1986; Yu et al., 1993), mostly independent of one another. Results of these efforts were disseminated through traditional outlets, such as conference presentations, technical reports, journal articles, and websites. These groups advanced the science and application of IEM, but their efforts and products typically were neither coordinated nor compatible. The challenge is to foster and promote participation in a coordinated manner across the full community, which is often easier said than done (Voinov and Bousquet, 2010).
In recent years, a community of practice approach has emerged as a paradigm within IEM. Several formal groups relevant to the IEM community have formed. Table 7 lists several of these groups and describes the focus and scope of their activities. These groups have formed around particular IEM sub-domains, such as multimedia modeling, surface dynamics models, earth systems modeling, and hydrology. Operationally, these groups act as communities of practice growing their knowledge base and developing solutions to common problems together.

More recently, communities have begun to form at levels higher than the sub-domain. For example, the U.S. National Science Foundation’s EarthCube23 initiative aims to support the development of community-guided cyberinfrastructure to integrate data and information for knowledge management across the geosciences by fostering community collaboration. In this initiative, community groups, consortia, researchers, and educators share ideas, introduce concepts, and find and develop collaborative efforts focused on solving issues common to all. The International Environmental Modeling and Software Society (iEMSs) was formed to develop and use environmental modeling and software tools to advance the science and improve decision making, promote contacts among physical, social and natural scientists, economists and software developers from different countries, improve the cooperation between the sciences and decision makers/advisors on environmental matters, and exchange information in the field of environmental modeling and software among scientific and educational organizations and private enterprises. iEMSs sponsors a biennial conference and focuses attention on several areas of importance to IEM. Another approach to addressing broader participation is being pursued by the Community of Practice for Integrated Environmental Modeling (CIEM). CIEM has formed as a community of communities with several goals in mind including formalization of the discipline of IEM, linking of the growing number of sub-domain communities, and development and promotion of best practices and standards at the global scale. As part of its strategy, CIEM has developed the iemHUB24 web portal to 1) enhance IEM learning and education (establish best practices and produce and share educational tools), 2) leverage IEM solutions (make them accessible and reusable), 3) facilitate scientific collaboration, and 4) allow efficient use of resources (limit duplication in technology development). Engaging this level of community has proven to be quite challenging. Organizations or
groups have a mission, which defines its work, needs, and priorities. There are many fundamental differences in the types of inter- or intra-organizational responsibilities, including regulatory/ enforcement, resource management, scientific research/monitoring, education and outreach, issue advocacy, community engagement, commercial or private sector development, etc. Even when there is a joint interest among several organizations or groups (e.g., working together to develop or apply an IEM), differences in priorities or alignment of perspectives can create barriers to effective collaborations. Organizations can be the greatest facilitators or barriers to a community of practice approach. For the IEM community of communities concept to be successful, organizations must come to terms with the need to collaborate in a joint effort to develop and promote the best practices and standards that will enable efficient sharing of the myriad of valuable IEM science products being produced.

5.2. IEM education and use

Workshop participants consistently expressed the need for a specific focus on education. To promote and improve IEM within the stakeholder community, a symbiotic relationship needs to exist between academic, government, non-government institutions and the general public. Academic institutions train the next generation of scientists and engineers, so developing an appreciation of and skill sets that deal with multi-dimensional problems, allows for a more holistic and systematic understanding. Academic curricula could be designed that not only develops IEM science and tools (e.g., modeling frameworks) but are also structured to clearly view issues and challenges that remain to be addressed. With the perspective afforded from this view, we present a roadmap to provide direction and context for the continued advancement of IEM.

IEM provides a science-based structure to assimilate and organize multidisciplinary knowledge. It provides a means to apply this knowledge to explain, explore, and predict environmental-system response to natural and human-induced stressors. Its structure serves as a unifying vehicle of communication among stakeholders holding diverse perspectives, values, and priorities. It serves the decision makers’ needs to understand the dynamic workings of systems involving social, economic, and environmental components, compare impacts among decision scenarios, and respond to unintended consequences.

In all of the workshop discussions leading up to the roadmap presented in this paper, there were several omnipresent themes that related to how we think about complex problems and how we should conduct the science and application of IEM. These are not new ideas; most have been part of the modeling conversation and literature for quite some time. The intent of sharing them here is to state that they remain not only relevant but critically important to the future value and acceptance of IEM. They should be explicitly considered and applied to IEM activities articulated in the roadmap. They are:

- **Systems thinking**: Intrinsic to solutions of complex problems is the idea that an appropriate decision (i.e., one that is science-based, cost effective, socially responsible, adaptive, and sustainable) requires a system framework and approach. All activities should reflect awareness of the larger system into which they fit.
- **Stakeholder involvement**: Ensuring appropriate stakeholder participation, assimilating the range of stakeholder perspectives, contributions and needs, and developing a consensus understanding of the problem, decision goals, conceptualized system, and solutions must be viewed as an essential ingredient for the conduct of IEM from individual components to complete decision support systems and applications.
- **Community development**: Sponsoring, nurturing, and participating in a global community that transcends individual groups and organizations will facilitate learning, sharing (knowledge and tools), and communication.
- **Openness**: Openness is a combination of transparency, cooperation, and collaboration. Openly sharing the products of individual research and development efforts will allow a wider access to and enable innovation with respect to IEM science, technology, and applications.
- **Reusable products**: Community-wide acceptance and use of globally recognized best practices and standards in the design...
and implementation of software-based science products is fundamental to long term IEM value and acceptance.

- **Investment**: Virtually any effort to develop a portion of an IEM solution for a problem has value beyond its original need. To realize this value (e.g., make it available to the larger community) requires an investment beyond that necessary for the problem at hand. This investment must be shared among those organizations that sponsor and fund IEM.

In developing the IEM roadmap, we organized the discipline of IEM as a landscape containing a set of interdependent elements including applications, science, technology, and community. IEM applications are the stakeholder community’s methods for selecting, organizing, integrating, and processing the combination of environmental, social, and economic information needed to inform decisions and policies related to the environment. The science of IEM provides the knowledge and integrative strategies that support and serve applications and related decisions. Technology represents the primary means by which the science of IEM is expressed and applied. Integrated modeling systems are constructed and executed on a variety of platforms serving research, applications, and education. Community reflects the fundamental nature of modern complex problems (i.e., problems affect communities and are solved by communities). In an equally important way, the community of IEM practitioners plays a fundamental role in the research and development of IEM science and technology. The roadmap presented in this paper is organized by IEM landscape element and includes a series of activities that represent a holistic approach to addressing the issues and challenges discussed throughout the workshops and summarized in this paper. Here we summarize the major activity areas for each element, whose details are captured in Figs. 2, 3, 5 and 6.

### 6.1. IEM applications

The roadmap related to IEM applications focuses on three principal activity areas. First, with respect to stakeholder involvement, roadmap activities include further development of methods and guidelines for elucidation and integration of diverse knowledge bases, perspectives, values, and priorities. Secondly, activities are focused on the use of IEM in a fully adaptive decision and policy formulation context. And finally, activities are focused on the identification and promotion of best practices, the use of applications as a tool for education, and the ability to reuse applications.

### 6.2. IEM science

The roadmap related to IEM science focuses on four principal activity areas. The first area involves developing awareness and guidelines related to holistic systems thinking and the design of integrated modeling systems for coherence and complexity. A second area of activity includes advancing the design of data monitoring studies to reflect the needs of IEM for cross-disciplinary data for model setup and evaluation. In the area of model evaluation, activities include the development of systems-levels methods for calibration and sensitivity and uncertainty characterization. Finally, peer review of complex IEM systems and their applications require attention.

### 6.3. IEM technology

There are three principal areas of activity for IEM technology. First, activities are included that focus on the development of protocols and standards for software design and implementation for reuse and interoperability. Secondly, activities are included that focus on building tools to enable automated discovery and utilization of IEM components and systems. The final technology activity area focuses on methods for further exploiting the World Wide Web and related technologies.

### 6.4. IEM community

A principal activity related to IEM community includes the articulation of the IEM science domain and its relationship to contributing disciplines. Additional activities include further developing and energizing a global community of practice for IEM, establishing IEM as a formal academic discipline, and encouraging funding organizations to coordinate funding efforts related to IEM.

Finally, discussions have recently begun concerning the implementation of this roadmap and activities. A key aspect of the implementation is that solutions to common issues and challenges reflect community-wide participation and acceptance. As such, implementation of the roadmap faces several challenges, principle among them is the need to transcend individual problem needs and organizational mandates and pursue solutions to core issues and challenges of IEM (i.e., the roadmap), as a well connected, cooperative, and collaborative global community. We encourage all IEM practitioners and stakeholders to contribute to this global awareness and effort.

### 7. Disclaimer

The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the affiliated organizations, except the U.S. Geological Survey. The use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. This manuscript is published with the permission of the Executive Director of the British Geological Survey (Natural Environmental Research Council).

### Acknowledgments

We wish to acknowledge the contribution to the numerous workshop participants who shared their knowledge and perspective concerning IEM. The research described in this paper was in part carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

### References


