Integrated Environmental Modelling – Achieving the Vision

R. V. MOORE[1] & A. G. HUGHES[2*]

[1]*British Geological Survey, Maclean Building, Crowmarsh Gifford, Wallingford, Oxon OX10 8BB, UK*

[2]*British Geological Survey, Kingsley Dunham Centre, Keyworth, Nottingham NG12 5GG, UK*

**Corresponding author (aghug@bgs.ac.uk)*

Abstract: Integrated Environmental Modelling (IEM) is a recent phenomenon which offers the opportunity to solve complex environmental problems. Whilst it has made great strides in recent years, there are still challenges to be met before IEM is universally accepted and used. This paper describes the current state of IEM and sets out a road map for achieving its full potential. A multi-disciplinary, multi-agency approach will be required whose main goals are to: (1) raise awareness and build confidence in IEM; (2) ensure availability and accessibility of IEM techniques, tools and standards; (3) establish a minimum set of standards; (4) build the IEM skills base; (5) establish an underpinning R & D programme; (6) co-ordinate and promote collaboration; and (7) foster IEM use by government, industry and the public. Once these goals have been achieved then IEM can be deployed to help resolve currently intractable environmental issues and the IEM methodology can be transferred to other fields.

Keywords: integrated environmental modelling; integrated modelling; environmental modelling; model linking; model coupling; strategy; roadmap;

INTRODUCTION

Purpose

The importance of understanding the world and all the events and activities within it as a set of interconnected interacting processes is now widely recognised. Technology is now in place that facilitates the linking of process simulation models, helping us to better understand and predict how the world will respond to events and management interventions. What is now needed is a strategy, an institutional infrastructure and resources to move that technology out of the research area into the domain of the early adopters. If those things could be put in place, then opportunities open up for finding sustainable solutions to present challenges and for developing new products and services.

In many spheres, simulation models have proved to be the most effective method of exploring processes, encapsulating our knowledge of them and predicting their response in real or imagined situations. This is true for all the physical sciences and applies to many social and economic sciences as well. To date most modelling development has taken place in relatively isolated, discipline specific groups; there has been little communication across traditional academic disciplines. By comparison with the investment in model development, little work has been undertaken on the complex problem of linking (or coupling[1]) model codes and their application[2] either within or across disciplinary boundaries; an activity generally referred to as 'integrated modelling' or in the context of this paper, 'integrated environmental modelling (IEM)'. Although much scepticism has been expressed about the viability of IEM, recently, a consensus on the need for IEM has emerged from a number of meetings between modellers, who either have an immediate need to study interacting processes or have seen the bigger picture and its opportunities (e.g. Moore and Hughes, 2012). Consequently, a number of draft roadmaps have been written to spur on progress in integrated modelling. Although these roadmaps have originated from different disciplines, they are remarkably

[1] The terms linking and coupling are synonymous and mean establishing a connection between two numerical models so that they can exchange data and, hence, simulate the interaction of the two processes. Generally, the link allows the model instances to request data from each other. The behaviour the requested model instance will therefore influence the behaviour of the requesting model instance.

[2] In this paper model codes are used to mean the software created to solve the mathematical models (i.e. MODFLOW) and model instance is the application of that software to a particular situation (i.e. London basin).

consistent in their conclusions. They all recognise that collaborations within and between disciplines and nation states will be required to bring about the necessary conditions for the new culture and technology to grow and flourish – environmental problems are seldom confined by national or disciplinary boundaries. An example roadmap developed against this background is the UK's *A Strategic Vision for UK e-Infrastructure* [BIS, 2011]. The report spans the aerospace and automotive industries, health and pharmaceuticals, entertainment and media, the bio-economy, weather and climate and basic research. It outlines the opportunity for scientific and industrial growth, the revolution in E-enabled science and innovation, and proposes a way forward that covers the actions required, a funding model and the communication channels between government and the participants. Being a master plan, it is inevitably set at a high level. There is, therefore, a need to translate these high-level aspirations into much more specific aims and objectives to be achieved in the environmental sector and then to propose how they are met and the benefits are applied to environmental modelling challenges. This paper sets out to meet that need for both the environmental sciences and their associated industries. It focuses particularly on integrated environmental modelling, identifying challenges and putting forward proposals for addressing them.

Context

The International Council for Science (ICSU) has recently defined the five "grand challenges" [ICSU, 2010] which need to be met if we, as custodians of the earth, are to manage our resources under increasing pressure from both population and economic growth and environmental change, i.e. land use change combined with a changing climate. These challenges revolve around the need to increase our capabilities in the fields of forecasting, observing, confining, responding and innovating and are as follows:

1. Improve the usefulness of forecasts of future environmental conditions and their consequences for people.
2. Develop, enhance and integrate the observation systems needed to manage global and regional environmental change.
3. Determine how to anticipate, recognize, avoid and manage disruptive global environmental change.
4. Determine what institutional, economic and behavioural changes can enable effective steps toward global sustainability.
5. Encourage innovation (coupled with sound mechanisms for evaluation) in developing technological, policy, and social responses to achieve global sustainability.

Box 1 lists some of the more eye-catching issues that are driving the need to understand the world as a system. It is not possible to answer any of them without that system wide understanding. The need for that understanding becomes even greater when other questions are considered, such as: "What can be done about it? What will be the impact? Is the solution sustainable? And how do we minimise the chance of unintended consequences?" Similar questions could have been provided from any other sector, for example health and transport: "What might be the impact of adding a third runway at London's Heathrow airport on health in the surrounding area?" What all these questions have in common is that often we have models of the individual processes involved but rarely do we have the means to link them together and answer the bigger question of how will they interact. Such complex questions are not the only drivers. Indeed, the real driver to solve the challenges of integrated modelling may come from everyday tasks that also require the same systems understanding of the world. Examples to consider might include determining whether a proposed discharge to a river, or a new groundwater abstraction, both of which could affect the river environment, should go ahead. Although simple by comparison to some of the questions in the side panel, they can still pose significant linkage challenges.

Many people can understand that system modelling might enable better decisions. However, many others are understandably sceptical as to whether integrated modelling can contribute meaningful systems understanding, other than in the most trivial cases. One reason for scepticism is that it is currently difficult to simulate even relatively simple natural processes reliably, that is to answer the question posed using the model instance to a suitably level of accuracy. Linking poorly constrained models simply compounds the problem; the real world is just too complex and chaotic for linking poorly constrained model instance to work. However, no alternative approach has been proposed. Accepting that the challenges ahead will be extremely demanding, some past experiences suggest that considerable progress can be made by following the integrated modelling path until an alternative can be found. One example is the development process from paper maps via digital maps to geographic information systems and Google maps, satellite navigation and location based services, all the latter now available through smart phones. In the 1950's, 60's and 70's, to all but the most farsighted, such as David Bickmore, the Director of the UK's Experimental Cartography Unit[3] (Coppock and Rhind, 1991), these things appeared equally unattainable for similar reasons. The digital mapping sceptics of the time had the added disadvantage of not having prior knowledge of the materials and IT revolution that we have witnessed. In spite of all the doubts at the beginning, the present mapping technologies have been created and today are taken for granted. It is highly likely that the development path ahead for integrated modelling will be very similar to that for mapping; so it is worth reviewing it briefly to see what it has to teach us and how the IEM development path might be shortened.

During the 1960's, when computing was becoming available to researchers, the first attempts at computerized mapping were concerned with automating the drawing of maps. For example problems such as capturing the line work from existing hand drawn paper maps, correcting for distortions in the paper, labelling each line with the pen width and colour for drawing it and then sending the information to a pen plotter for drawing needed solving. For many years, the resulting maps were far inferior to anything that a cartographer could achieve. For at least a decade or more the advances represented the triumph of hope over adversity; the benefits, financial or technical, if any were barely discernible. Problems which today seem inconsequential were major challenges. Examples were: how to place names on maps automatically without overwriting; how and when to simplify a motorway junction when drawing at smaller scales; and how to produce maps by computer that were as artistically pleasing to the eye as those drawn by hand. These and many similar problems were gradually solved in the course of thousands of MSc and PhD projects. During this phase, there was a growing awareness that the geometric data being amassed contained vast amounts information that could be exploited for all sorts of purposes beyond reproducing maps. The problem was that most of the coordinates related to the sheet of paper from which they had been captured; they had not been matched in any way to their position in the real world. Further, none of the data had been structured or labelled – the phrase at the time was 'spaghetti data'. There was no way of telling what a line represented, how it connected to other lines and, if it was a river or road, which way water or cars could pass along it.

A long and frustrating hiatus then followed. Considerable energy was expended on how to structure and label lines, and on developing algorithms to exploit those structures. At first these algorithms addressed what today will seem very simple problems, such as how to find out quickly which of a set of points or polygons, representing, for example, houses, lay within another polygon, representing, say, a district boundary. Against the computing power of the day, this was a difficult problem. As people and then commerce slowly began to see the opportunities, pressure for more sophisticated applications grew, together with the need to merge map data with other data. Among the earliest data sets to be merged were the environmental data required for flood estimation, but these were closely followed by planning, road and utility network characteristics and census data; today's

[3] Established in the Royal College of Art in 1969 - see http://www.casa.ucl.ac.uk/gistimeline/original/1960.html

satellite navigation systems depend upon knowing the speeds attainable along each stretch of road throughout the day, now in real time. The initial systems for managing and using these data were large, cumbersome and expensive (Goodchild, 1988). There were no standards and moving data from one system to another generally involved reformatting. Looking back, progress was grindingly slow for most of the time that it has taken to make the transition to today's systems. It is sobering to realise that that period now covers 50 years! The development and uptake of digital mapping followed the classical "Diffusion of innovation" (Rogers, 2003) which results in an exponential curve towards market saturation. However, what is possible now is not just due to the immediate digital mapping community. It has only become possible because of all the materials science, the physics, the chemical, the communications and the IT advances that have occurred in parallel (see for example Figure 2, Crampton, 2000). Also important has been the long hard slog of first transforming worlds maps into digital form as well as the development of standards (and standards organisations to support and maintain them) that have allowed digital maps to be brought together to create a global data set. It was when a critical mass of map data existed that progress began to accelerate and industry to invest. At one point, a real show stopper was the mistaken belief that enormous income could be generated from the sale of digital map data. This unfortunately coincided with the period in the 1980s and 1990s when many public undertakings, the main holders of map data, were privatised. It took several years for prices to fall to levels where user organisations were prepared to start buying and a proper market established itself. Several more years passed for prices to come down to where a mass market could open up, giving everyone access to the benefits of digital data. Here, price was not the only factor. Ease of use was central. Today it requires no more skill than that needed to rub a finger across a glass screen of a mobile to find the nearest restaurant for lunch. All the complexity involved in making that possible is hidden. This last point highlights a very important lesson. Most of the initial funding, which transformed a wild idea into a viable technology and digitised the base information, came from public sources. When, critical mass had been achieved, commerce joined in and created a whole new industry which has already and will continue to find new and innovative ways to exploit digital map data, most of them much simpler and far removed from the weightier applications which the pioneers had in mind. There are many lessons to be learned from this story and many more could have been added had there been space to give a more extended account. These have been incorporated in the proposals that follow later in the paper.

BACKGROUND

What is IEM?

Integrated environmental modelling (IEM) is becoming an increasingly valuable technique for understanding how processes interact and then using that knowledge to assess the likely outcomes of given scenarios. The detailed nature of IEM is evolving rapidly but an IEM toolkit can be visualised as a set of tools and modelling components. The modelling components include simulation models, databases, analysis tools and visualization tools among many others; the key feature of these components is that they are linkable, meaning that they can request data and send data from and to each other as they run. This is achieved by all the components adopting a standard interface for data exchange, for example, the OpenMI (~~REF~~Gregerson et al., 2007). The tools help developers and users undertake such tasks as making pre-existing models linkable, linking components to create an integrated model, maybe a decision support system, and running it.

It is convenient to think of IEM under four headings (based on the four suggested by Laniak et al., 2013):

- IT science
- Modelling science
- Organisation
- Applications

At the **IT science level**, in concept, IEM is simple; an integrated model is created by linking together models of individual or groups of processes (Moore and Hughes, 2012). The links allow data to pass between the models and the technical challenge is how best to achieve the transfer. Essentially, it is a pure IT problem whose solution should be independent of the data being passed. The problem is challenging because the requirement is to be able to link models based on different concepts, working at different scales, using different spatial and temporal resolutions and representations (including none at all) and sourced from different suppliers. In the simple case where the processes being simulated are sequential, the results of the first model can become the input of the next. All that is needed to affect the transfer is for the second model to be able to read the output of the first. More sophisticated approaches are required to handle situations where the processes are running in parallel and interacting time step by time step. This challenge can become greater if the models are running in different computing environments or in the high performance or cloud environments. At the time of writing, several approaches exist for solving the linkage problem and, in principle, most are very similar (e.g. OpenMI, Moore et al., 2005; CSDMS, Peckham et al., 2013; OMS, David et al., 2013). They work by providing each model with a standard interface through which the model can request and receive data or respond to requests.

Having developed a number of viable solutions, the IT science focus is now turning to the problem of transforming integrated modelling from a research tool which requires a high degree of skill to use into a tool that anyone can apply with ease and confidence. Can the processes of making models linkable, finding appropriate models, linking and running them be made invisible, as have the multitude of processes that take place when two zip codes are entered into a satellite navigation system, which then responds with maps and voice showing the way from point A to point B?

Now that it is possible to link models more easily than in the past, it is becoming easier to consider the opportunities that IEM creates for studying interacting processes. Integrated **modelling science** is concerned with the issues and opportunities that linking models of the same or different processes raises or creates, for instance:

- The validity of integrated modelling – is it valid to link (e.g. Lloyd et al., 2011) an economic model to an ecological model; two models at different scales; two models based on different conceptualisations of the world?
- How to approach a problem that requires an integrated approach? Experience shows it is very easy to focus on the mechanics of model integration and lose sight of the real problem. Voinov discusses how to conceptualise process interactions and how to calibrate and validate model chains in (Voinov et al., 2013).
- Can the propagation of uncertainty along a model chain be represented and assessed? (e.g. Bastin et al., 2013). How can uncertainty be modelled in a multidimensional web of interacting models with complex feedbacks, rather than a chain?
- How to establish and populate: catalogues for both model codes and applications of the model code – descriptions that allow people and machines to find and evaluate appropriate model codes and their instances (e.g. Harpham et al., 2013); controlled vocabularies of names of model inputs and outputs to reduce the chance of invalid links and for use in automated linking; ontologies (e.g. van Ittersum et al., 2008)
- How to exploit artificial intelligence and ontologies to reduce the chance of 'unintended consequences'. These are model instances that are linked inappropriately; see for example

Voinov and Shugart (2013) for the concept of "intergronsters". There are clearly limits to the human mind's ability to foresee all the possible consequences of a particular course of action. A limitation of current modelling is that is necessary to be able to define the scenario to be modelled before it can be modelled. If we can hold our knowledge of things and of processes in a structured manner, then it might be possible to construct algorithms that can search the knowledge base for potentially disadvantageous situations which we could then model in detail.

- How to reconcile the different objective functions used by different disciplines when evaluating similar interacting processes (e.g. those used to optimise the carrying capacities of sewers and rivers in relation to flooding)?
- How to analyse and present linked model results?
- How to record the details of linked model runs so that the results are reproducible and the lineage of all input data and model components is traceable?

The **organisational level** is concerned with creating the conditions for IEM to flourish. At present, much of our knowledge about earth system processes and their integrated modelling is concentrated in small groups spread across the world, each working on a particular problem or challenge. For some time there has been a shared belief that if some of these groups could be brought together huge opportunities for technical advancement and innovation could be created . This belief is based partly on observation of the advances over the last fifty years in other disciplines, especially geographic information system (GIS) development, and partly on consideration of the opportunities that would be created if a large array of linkable models spanning diverse processes from many disciplines became widely accessible.

One of the first groups to bring modellers together was the US Interagency Steering Committee on Multi-media Environmental Modeling (ISCMEM). This group however was confined to US federal agencies. Since 2002, there has been growing informal collaboration between the European Community (EC) countries, the United States (US) and Australia. The relationship has been actively encouraged by the EC and the UK Foreign and Commonwealth Office. It has focused on integrated modelling and, in particular, the standards necessary to enable models to be linked together, for example, the Open Modelling Interface (OpenMI) and several others.

A number of international meetings have been held in Europe and the US to establish whether the optimistic view of IEM's future of is widely shared; and, if so, raise awareness of the new opportunities among potential users, create the conditions for its use to expand and secure the funding to facilitate the change. These meetings arose out contacts between the OpenMI team, the United States Geological Survey (USGS), the US Army Corp of Engineers (USACE), the US Environment Protection Agency (EPA), the Community of Universities for the Advancement of Hydrological Sciences Inc. (CUAHSI), the Community Surface Dynamic Modeling System (CSDMS) and leaders such as Prof David Maidment of the University of Texas and Dan Ames of Idaho University. Following US EPA meetings in 2007 and 2008 to discuss a draft white paper on the role of integrated modelling in its regulatory work (Gaber et al., 2008), an embryo international Community of practice for IEM (CIEM) emerged.

To follow this up, a major meeting, the Summit on Integrated Environmental Modelling was held in Washington in December 2010 (Moore and Hughes, 2012). Its chief outcome was an IEM Roadmap (Laniak et al., 2012). More recently, the White House sought similar advances in modelling, public communication and transparency (US Government, 2012) and this is being advanced in the US water/environment arena by the EarthCube initiative amongst others (NSF, 2011). The environmental sector is not alone in recognising the importance of integrated approaches for taking forward its science and in finding sustainable solutions to society's challenges. For example, medical researchers have eloquently argued the case in two papers (Viconti and Clapworthy, 2011; VPH, 2009). As well as these discipline specific papers, the bigger picture covering the need to be able to bring together knowledge from all disciplines is covered in 'A Strategic Vision for UK e-Infrastructure' (BIS, 2011). This work has been monitored to see how successful it has been (BIS, 2013) and subsequently implemented by the UK's research councils (e.g. EPSRC, 2014).

IEM **applications** have the potential to deliver what modellers have long sought: 'plug and play' modelling and decision support system development. In practical terms this means that the modeller has access to a pool of linkable models and other modelling components such as GUI's. These components will be located and discoverable from anywhere in the world. Anybody will be able to contribute components and they will be available as open source or commercial code. For a component to usefully join the pool, it must be able or be capable of being made able to expose the variables it accepts or provides through one of the standard interfaces. There are already available a number of modelling platforms that allow users to browse available models, link them together and run the resulting composition. The linking process simply involves dragging each relevant output variable of one model onto the corresponding input of another. When the models compute values of each variable for a number of nodes within a model (e.g. end points of a river reach or cells in a grid), a second step maps the nodes of the two models onto each other. Larger models can involve many thousands of connection points. While the process of identifying and making geographical linkages is largely a manual exercise at present, work is underway to automate it. A water resources example might involve connecting a 1 or 2-D hydraulic model of a river network to an underlying groundwater model based on a 2 or 3-D grid. In most cases, the linkages can handle differences between models in time steps and units of measurement. Object oriented programming and open standards have greatly facilitated these linkages. Open standards in particular allow modellers working independently to produce components which they can be confident will link to other independently developed components. This has profound implications for the modelling market place, for science, and for the innovation community.

The attainment of a critical mass of modelling components and advances in the ease of use and reliability of modelling platforms (such as have been reached by GIS and office software) will be crucial to the take up of IEM. Once these are realized, then the current levels of time and effort required to link models and reformat data will largely disappear. More time will be available to study actual problems or issues and to find solutions. The Shell company has not been alone in discovering that its employees spend far more time reformatting data than in analysing them (NERC, 2009)!

Although the original reason for a component based approach to integrated modelling was to increase understanding of process interactions and apply it to real world problems, its first practical application has been by developers. Many of their applications have become large, cumbersome, pieces of code. They need a way of breaking down these systems into more manageable components. If they could do this, not only would maintenance and development become simpler but components could more easily be reassembled to create product variations or new products.

Science and societal challenges however remain the primary drivers. A first quick win, already beginning to be achieved, is the relative ease with which one modelling component can now be replaced by another in sensitivity testing. Bigger wins will accrue however if IEM can help bring better or timelier answers, for example to challenges of climate change or to incidents such as the Deepwater Horizon oil spill. The National Park Service Science Advisor who had to provide a quick response for the Park Service to the Deepwater Horizon tragedy, lamented that he had no model to help assess the situation immediately after the blowout occurred or to indicate the decisions most likely to mitigate the situation (from the Park Service's perspective). He needed an integrated model that could indicate how the oil was most likely to disperse, how it would behave when it reached the surface, which way it would go, what the environmental and economic damage might be should the oil reach the coast, and the likely effectiveness of different strategies.

From earlier integrated modelling exercises, it is known that modellers should expect to encounter and have to overcome a range of problems as new codes from different disciplines are linked together. For example, when a sewer model was linked to a river model to investigate the impact of sewer flows on river flooding and vice versa, it emerged that the criteria by which sewers were designed differed from those used to assess the adequacy of river channels. One was evaluated against a design storm while the other was assessed according to risk exceedance probabilities. Until these very different objectives were reconciled, meaningful conclusions were not possible.

Integrated modelling makes semantics increasingly important, especially as linkages are made across models/codes incorporating different disciplinary expertise or sourced from countries with different languages. Even within an area of disciplinary expertise, many terms used as variable names and the names of model inputs and outputs are often loosely defined. As numerical modelling is opened up to a broader range of scientists across the world, it will become increasingly important to have clear definitions of what each variable represents , so that modellers can take cognizance of codes and their linkages and simulations, and so that they can make appropriate judgements regarding the validity of their model constructs.

Most models conceptualise their view of the world at a particular scale. Most if not all of the current interfaces make no attempt to judge the scientific validity of a connection. This is the correct approach as scientists should be given the freedom to make their own judgement, and the technology doesn't yet exist to automatically make the links. However, the freedom to join anything to anything also leaves open the possibility of invalid connections (i.e. "Integronsters"; Voinov and Shugart, 2013). Linking models at different scales, whilst perfectly valid, can result in misleading results if the aggregation and disaggregation of data at the join is not carefully considered. It is important to establish good practice early.

There are clearly number of challenges to making IEM a useable tool, however a significant amount work has been already undertaken. The following section reviews the current state of the art.

STATE OF THE ART

Introduction

Integrated modelling is not new. The analogue flight simulators of the 1950's and 1960's were highly sophisticated integrated models of aeroplanes. It has been argued that some of the early digital models of environmental processes from the 1970's and 1980's could be called integrated if the code representing the different processes was separated out and placed in different subroutines. More recent examples of coupled models linking different processes include: linkages to the MODFLOW (Harbaugh, 2005) groundwater simulation code; such as the coupling (Sophocleous and Perkins, 2000) with the SWAT (Arnold et al, 1998) land surface model ; or the coupling with the HSPF rainfall runoff model (Donigian et al., 1995) to create IHM (Zhang et al., 2009); or the coupling with the PRMS watershed model to create GSFLOW (Markstrom et al., 2008; http://wwwbrr.cr.usgs.gov/projects/SW_MoWS/index.html).

The first large scale environmental integrated modelling exercise was the UK's £30M Land Ocean Interaction Study (LOIS) whose overall objective, set in the late 1980's, was 'To study and better understand coastal zone processes and their interactions in order to facilitate the development of sustainable management policies' (Wilkinson et al., 1997) and more specifically:

1 To measure the contemporary fluxes of materials (sediments, nutrients, contaminants) through the coastal zone.

2 To characterise key physical and biogeochemical processes that govern coastal morphologies and ecosystem functioning.

3 To describe the evolution of coastal ecosystems over the last 10,000 years in relation to climate and sea level change.

4 To develop linked land-ocean models to simulate the transport, change & fate of materials in the coastal zone as a basis for predicting changes over the next 50-100 years.

As will be explained below, the science at this time was severely limited by the available IT. Further, the environmental drivers were only just appearing and being recognised in the political world. During the 1990's and the start of the present century, public awareness changed dramatically and led to the need for much better understanding of the earth as a system, so that politicians and managers can find sustainable solutions to the emerging challenges so admirably summarised in the Belmont Forum report (ICSU, 2010).

Recent developments

There is a growing realization that it is neither practical nor useful to construct a single model encapsulating all the processes needed for decision-making and planning (Argent, 2006; US EPA, 2008). Not only are such large models extremely wasteful of resources, they are rarely reusable and frequently fail to make use of existing process models. These are often referred to as 'legacy models' -- the result of a huge, historic investment representing state-of-the-art modelling. Consequently, there are currently attempts to convert existing models into building blocks from which more complex models can be assembled (Warner et al., 2008; Barthel, et al., 2008; Argent et al., 2009). In today's IT terminology, these are referred to as 'components'; components that can be linked are called 'linkable components'. The term 'modelling component' now has a wider meaning that is not limited to models: it includes files, databases, analytical tools and visualization tools and, indeed, any component required to make up a modelling system. At the time of writing, although IEM technology is orders of magnitude simpler to apply than ten years ago, it is recognised that there is a long way to go before it matches what has been achieved in other technologies. Examples of attempts to streamline model integration are:

The US EPA, in conjunction with the US Nuclear Regulatory Commission, US Army Corps of Engineers, and US Department of Energy's Pacific Northwest National Laboratory, has been developing the Framework for Risk Analysis in Multimedia Environmental Systems (FRAMES-1) (2009) system to manage execution and data flow among multiple science modules. It uses a fixed file format system to exchange data between components. The Multimedia, Multi-pathway, Multi-receptor Risk Analysis system (FRAMES-3MRA) (2009) (Babendreier and Castleton, 2005) is an extension of FRAMES-1 and is based on an API and dictionary system to exchange data. 3MRA is a collection of 17 modules that describe the release, fate and transport, exposure, and risk (human and ecological) associated with contaminants deposited in various land-based waste management units (e.g., landfills, waste piles). The 17 models in 3MRA cannot be easily replaced. FRAMES-2 (Whelan et al., 2010) represents the best attributes of FRAMES-1 and FRAMES-3MRA and is designed to allow for easier registration and replacement of models and support components.

The Open Modelling Interface and Environment (OpenMI, 2009) developed by a consortium of European private companies, research establishments and universities co-funded by the European Commission is a standard for model linkage (Moore, et al., 2005). The OpenMI Standard version 1.4 defined an interface that allows time-dependent models to exchange data at runtime; hence, OpenMI-compliant models can be run in parallel and share information at each time-step. It is, therefore, particularly appropriate for situations where it is necessary to simulate interacting processes, such as changes in river flow which increase nutrients which affect plant growth which, in turn, affects flow. It can handle feedback loops and iteration. It can link models based on different Modelling concepts. OpenMI is a generic standard and can be used to link models from different domains (hydraulics, hydrology, ecology, water quality, economics etc.), environments (atmospheric, freshwater, marine, terrestrial, urban, rural, etc.), scales and resolutions (spatial or temporal), platforms, or suppliers. It is not limited to linking models, but can also link any modelling components. OpenMI version 2.0 paves the way for linking models that run in a super-computing environment and models provided as web services. It can exchange a wider range of data types and simplifies the exchange process when models have either no spatial and or temporal dimensions, such as a terrain model. While version 1.4 only provided a ‘get values’ data option, version 2.0 also provides a ‘set values’ option to facilitate model optimization. The OpenMI version 2.0 is now an Open Geospatial Consortium (OGC) international standard.

The Common Component Architecture (CCA) is a product developed by the Department of Energy and Lawrence Livermore National Lab teams (Bernholdt et al., 2004) which targets high performance computers and complex models. The CCA supports parallel and distributed computing, as well as local high-performance connections between components, in a language-independent manner. The design places minimal requirements on components and facilitates integration of existing legacy code into the CCA environment by means of the Babel (2004) language interoperability tool, which currently supports C, C++, Fortran 77, Fortran 90/95, and Python. The CCA is being applied in a variety of disciplines, including combustion research, global climate simulation, and computational chemistry; it has also been adopted as the backbone in the Community Surface Dynamic Modelling System (CSDMS, csdms.colorado.edu; Peckham et al., 2013).

The Object Modelling System (OMS) was developed by the US Department of Agriculture (David et al., 2013; Kralisch et al., 2004; Ahuja et al., 2005). In contrast to FRAMES and some other systems, OMS requires modules to be rewritten in Java prior to insertion into the system library. Instead of just linking pre-existing blocks or components, OMS provides the tools and integrated framework to develop the components of an IEM in a coherent way.

Despite the evident need for IEM, it has yet to “take off” in the way its proponents hoped. The reasons include: (1) lack of convincing demonstrated added value provided by IEM; (2) lack of a critical mass of available and accessible linkable modelling components; and (3) difficulty of use and other barriers to entry. The latter include: having users with the right skills, lack of linkable models or linkable components, poor accountability on uncertainties and lack of necessary computer power. However, communities of practice are slowly emerging as a part of a number of IEM initiatives, which are attempting to address the challenges of IEM. Currently, the initiatives and their communities are relatively isolated because there is no umbrella organization to bring them together. Examples of these communities and initiatives include:

- OpenMI Association which makes the OpenMI standard freely available (www.openmi.org);
- CSDMS - Community Surface Dynamic Modelling System which “makes earth surface process models available, has computational resources for model simulations, and couples models that bridge critical process domains” (csdms.colorado.edu);
- CCMP - Chesapeake Community Modelling Program, dedicated to advancing the cause of accessible, open-source environmental models of the Chesapeake Bay in support of research & management efforts (ches.communitymodeling.org);
- ESMF – the Earth System Modelling Framework: software for building and coupling weather, climate, and related models (www.earthsystemmodeling.org)
- CHyMP – the Community Hydrologic Modelling Platform (www.cuahsi.org/chymp.html)

There are also communities designed to support individual models and software packages, examples include but not limited to:

- GRASS - free Geographic Information System (GIS) software used for geospatial data management and analysis, image processing, graphics/maps production, spatial modelling, and visualization (grass.fbk.eu);
- MapWindow - another GIS project that includes a free desktop geographic information system application with an extensible plugin architecture (www.mapwindow.org);

- ADCIRC - a system of computer programs for solving time dependent, free surface circulation and transport problems in two and three dimensions (www.unc.edu/ims/adcirc/);

There is now a need for an umbrella organization, maybe in two parts, one academic and one commercial, to coordinate these separate efforts. Collectively, they need to: raise awareness, promote interoperability between current *de facto* standards, lower the barriers to entry, and organise and fund a global research programme to accelerate innovation and the use of IEM.

What are the gaps?

As expected, the new ease of model linking has brought out new challenges many of which have yet to be addressed. An early discovery in the OpenMI-Life project was that joining models designed to operate separately can result in some or all of the models being pushed outside their design limits (~~ref)~~Sofiolea et al., 2009). For example, a hydraulic model normally used for flood studies was found to become unstable when the simulation was continued into a period of lower flows as required by the linked model. Hence when metadata standards are developed for describing models, it will be important that they indicate the assumptions made in developing the models and the circumstances they can be safely applied (Harpham and Danovaro, 2015).

As mentioned above, it was discovered in an OpenMI-Life case study on the Pinios River catchment (~~ref~~Makropoulos et al., 2010) that the design criteria used by different disciplines for same process could differ. Clearly it will be important for modellers to resolve such differences if they are able to draw any useful conclusion from the linked model. When existing models of an area are linked, the geographical extent and representation of the two models are often slightly different. At present, there is little understanding of the impact of these differences on the results. Similarly, the effects of linking models running at different time steps are not fully appreciated. Currently there are few validation checks ensuring that links between models are valid. This is intentional and leaves the modellers free to make any connection they wish. However, there are many operational contexts where it will be important that the risk of an invalid connection is kept a minimum. The starting point for addressing this problem will be the introduction of controlled vocabularies and later ontologies for input and output variables; the valuable preliminary work of the SEAMLESS project on ontologies for models will need to be greatly extended (van Ittersum et al., 2008).

As the Deepwater Horizon oil spill showed, there is the challenge of having a collection of linked models ready before a situation occurs (Machlis and McNutt, 2010). At present an impact has to have been anticipated before it can be modelled. Can Artificial Intelligence (AI) enable an ontology to be searched for a potential impact and be more successful than a normal human brain. Can AI be better at the identification and linkage of components and, if so, could AI be used to at least semi-automate the present manual model linking process? IEM skills are spread across different organisations, however, the communication between these groups is very poor; the resources of groups such as the Integrated Environmental Modelling Software and Systems society need to be increased so that the society can extend its reach. In the UK HR Wallingford has developed Fluid Earth (Harpham et al., 2013) for use internally and to open up the means of model linkage to other organisations. This is beginning to achieve a significant number of down loads especially in China. In the UK, it has been used on operational problems within the British Geological Survey (BGS) to help it solve advanced groundwater problems such as linking groundwater models with surface models to provide an integrated understanding of surface water driven groundwater flooding on the Oxford floodplain (Macdonald et al., 2012). It has also been used to holistically simulate groundwater processes in the Thames Basin (Mansour et al., 2013). Other commercial developers are producing similar platforms but each addressing a slightly different user need. These are the first signs of a market forming and need to be encouraged. However, despite this and other similar projects, e.g. those undertaken for OpenMI-Life - Safiolea et al. (2009), there is still a dearth of “showcase” projects which can be used to give others the confidence to join in. Demonstration projects probably represent the largest gap at the moment and will be a priority task in the strategy to be outlined later in this document.

One final point is the recognition of the human dimension of IEM and how complexity should be dealt with. Given that linked model instances will be more complex than the single model approaches of the past, then there needs to be a structured way to deal with complexity and in particular the “complexity paradox” (Oreskes, 2003). The latter being the innate desire of scientists to make their models more complex as a better understanding of natural systems develops. Along with this, there is an increasing need to involve stakeholders in this process and to undertake participatory modelling (e.g. Voinov et al., 2016). Any IEM development should be coupled with a critical analysis of the development of linked models and Glynn (2015) proposes a “re~~a~~d team” approach to ensure that IEM is applied appropriately. This uses separate teams which ensure that “groupthink” does not develop whereby an idea or approach is not properly challenged.

IEM has now reached the stage where it is ready to move out of the research arena and to be taken up by early adopters. As one modeller put it "IEM is a no brainer" (Jackson, pers comm). Whatever its current imperfections, for the moment, there does not appear to be an alternative. Therefore, we have to strive to overcome the imperfections, make it easy and safe to use and reduce the barriers to take up. In particular, it needs to be moved out into the market place where it can demonstrate its potential. It can then attract the investment that can enable it to form the basis of new industries comparable and exceeding those spawned by digital mapping and its related technologies.

ROADMAP

Vision

The vision for IEM is a world where the user (who can be anyone from politician to a member of the public) can:

- articulate a problem which involves understanding and or predicting how many processes will interact in a given scenario, e.g.:
    - Will climate change alter the frequency and magnitude of flood damage across Europe?
    - Are the medical plans for treating X a threat to water supplies?
    - Will a switch to biofuels drive up the demand for water and hence energy (large amounts of energy are used in the delivery of water)?
- be guided to an appropriate set of modelling components (in some cases and for some users this can be hidden)
- be assisted to assemble the components (in some cases and for some users this can be hidden)
- run the composition (in some cases and for some users this can be hidden)
- analyse the results or be presented with an analysis
- be presented with the conclusion and if relevant information about the confidence that the user can place in that conclusion.

Mission

The mission of the IEM community is to

- Raise awareness of and build confidence in IEM by showing its utility to solve complex problems, so-called “stress test”;

- Improve the ability to understand and predict how processes, particularly environmental processes, interact;
- Create the opportunity for innovation, especially by making it possible to link models of processes from different disciplines and market sectors;
- Create new markets.

Strategy

The meetings, formal and informal, of the leaders of the IEM community over the last four years have been remarkable for the degree of consensus as to what needs to be done to move IEM from the present state of the art to a technology that is widely available and accessible. The essential elements of the strategy are:

a. Raise awareness and build confidence in IEM
b. Ensure availability and accessibility of IEM techniques, tools and standards
c. Establish a minimum set of standards
d. Build the skills base
e. Establish an underpinning R & D programme
f. Co-ordinate and promote collaboration
g. Increase IEM use by government, industry and the public

IMPLEMENTATION PLAN

Set out below is a plan whose objective is to carry through the main elements of the IEM strategy.

Raise awareness and build confidence in IEM

<u>Situation</u>

The leading IEM research and development organisations have begun the process of raising awareness of the benefits of IEM. The first audience has been their own managements and interestingly it has been within the commercial players that they have been most successful; indeed three are now routinely using the new technology internally. What is needed next is to raise awareness in the following areas:

- Governments
  - In the departments responsible for business development
  - In the departments responsible for policy development, i.e. the potential direct or indirect users and beneficiaries of IEM
  - In the departments responsible for overseeing and funding research
- Academia
  - The organisations responsible for developing research programmes
  - The organisations for developing the teaching curriculum
- Government agencies and local authorities
  - Those that undertake, provide or commission modelling and purchase modelling components and or services
- Consultants
  - Those to whom modelling studies are outsourced
- Utilities and other commercial users of models and modelling services

- Major investors

The major investors are placed last because experience suggests that they will want to see a strong demand that justifies the significant investment that will be required to take IEM from its present state to one comparable to that which now exists for GIS and location based services.

The audience will be extended later, eventually down to the public at large. It is important that at this early stage such resources as are available are focused on those who can bring about change.

It is anticipated that the various audiences may well require very different information packs ranging from case studies to business plans.

Actions

Task 1: identify and prioritise the information about IEM that each audience requires and the medium through which it will be most effectively communicated. Issues known to be of concern are:

- Integrated Environmental Modelling: What is it? What will it do for me and my organization? How will it add value?
- A wide range of demonstrations of added value illustrating: better science; better management answers (greater certainty; fewer unanticipated outcomes); increased opportunity for innovation; opportunity for wealth creation; reduced costs
- Will the ease of code reuse, reduced development time and increased efficiency produce a worthwhile return on investment?
- Will IEM results be accepted in a legal context?
- How will QA issues be addressed: transparency; audit trails; reproducibility of results?
- Reliability of IEM
- Support into the future
- Are the backers of IEM credible?
- What is the business plan for developing an IEM market?

Task 2: generate the information. It is expected that the highest priority requirement will be for a wide range of case studies/exemplars.

Task 3: disseminate the information

Ensure availability and accessibility of IEM techniques, tools and standards

Situation

Currently, only a first generation of IEM tools, standards and a small set of applications can be said to be available. They are only accessible to a small community of modellers with the skills to use them. Development of a supply chain and market is now needed to enable the delivery of IEM technologies to a broader set of end users. These users can be found in organisations faced with challenges whose resolution requires understanding of many interacting processes. In most cases these organisations are unaware of the advances in IEM technology and certainly do not have the in-house skills to apply them. The same is true of the consultants who provide them with knowledge and expertise, though they could come up to speed quickly if the necessary incentives were in place. The short term strategy is therefore for the developers of IEM to provide IEM support services to those needing to grow IEM expertise or use. These services range from training to undertaking IEM aspects of particular projects. Some may well follow the Environmental Systems Research Institute (ESRI) model in which ESRI makes their software available to the academic community at very advantageous prices. This pays significant dividends later when those students with their GIS skills move out into industry and understandably recommended buying ESRI software. Engagement with the academic community is important; only they have the resources to produce graduates in the numbers required. That said, it will be important to consider how remote learning can also be exploited to quicken the pace of growth.

The success of IEM is totally dependent on achieving interoperability and communication between people and between modelling components. In a global development exercise such as will be needed to get IEM off the ground, standards will be key to success. To ensure that standards do not become a barrier either to development or to entry to the market place, there is a consensus view that all the key IEM standards should be 'Open Standards'. For similar reasons the OpenMI Association and others have made their IEM tools and platforms open source and the strategy is to encourage others to do the same. In this way, the barriers to entry for new comers are kept low. All that is needed to join the IEM community is a PC and an internet connection.

For those who want to apply IEM for themselves, the cost of purchasing IEM applications can be a deterrent. One option to reduce the cost and encourage participation is to provide access through web services or the cloud. This is clearly only an option where the data volumes exchanged between components is low. There are many such instances, and as line speeds increase it will become feasible in more situations. The approach has the advantage of leaving software and data with the people in a position to maintain it.

Whilst ease of use has been transformed in the last decade, there is still a long way to go before the ease of use of IEM matches that of office software such as word processors, spreadsheets or GIS. However, ease of use has been identified as a major barrier to uptake and therefore a priority for action.

Actions

Task 1: Increase the availability of IEM by:

- Creating an open source and commercial IEM market place for:
    - IEM modelling components
    - IEM modelling tools
    - IEM modelling expertise
    - Related data
    - IEM compliance testing
    - Encouraging the development of new and upgraded models and applications as linkable components
    - Encouraging the development of 'adapters' that will allow modelling components using different data exchange interface to be linked and so increase the size of the pool of linkable models.

Task 2: Increase the accessibility of IEM to users by:

- Improving the ease of use of IEM tools and techniques by:
    - Developing the means of finding linkable models (see model meta data standard below)
    - Automating where practicable the process of making a model linkable
    - Providing debugging tools
    - Providing compliance testing tools

    - Minimising the opportunity for error in the coupling process through: vocabularies, ontologies and artificial intelligence – see also Standards below.
    - Providing adapters for spatial and temporal aggregation and disaggregation, unit conversions, interface mapping when data pass between models
    - Hiding processes and procedures irrelevant to end users
    - Providing a wide range of exemplars which those new to IEM can easily adapt to their specific problem
- Developing tools for calibration and validation
- Providing GUI's that allow all the results from a linked model to be seen together rather than individually through the GUI of each component
- Providing training – see building the skills base below

Establish a minimum set of standards

Situation

Standards for data exchange between models fall into two types. File format standards are used where models run sequentially and data exchange is affected by making the output file of one model the input file of the following model. WaterML (Taylor et al., 2012) is an example of such a standard. Where models run in parallel with data exchange happening time step by time step and the exchange takes place in memory, then interface standards are used.

To date standards work in IEM has been focused on developing interfaces for run time data exchange between models. A number of *de facto* standards have emerged, of which one, the Open Modelling Interface (OpenMI), is approved as an international standard by the Open Geospatial Consortium. These informal standards have emerged to serve different modelling communities for example climate, ocean modellers and water industry modellers. The OpenMI arose from the European Commission's programme of underpinning research in support of the Water Framework Directive and the implementation of its policy of an integrated approach to water management across Europe. None of these standards purports to be the ultimate answer to all modeller's needs and much further work will be required; though for the moment it is probably more important to grow the set of linkable models and for the time being create adapters that can link models following the different interface standards – see 'Ensure availability and accessibility of IEM techniques, tools and standards', Task 2 above.

No other IEM standards have yet emerged but the need for them is widely recognised. Now that interface standards are widely used within IEM, the next priority is for a minimum standard for the description of models and modelling components. The need to standardise the descriptions of linkable modelling components arises for several reasons:

- Models are developed across the world by independent groups. If their basic descriptions are published in a common format, it will greatly increase the chances of users, who are also dispersed, finding the components that most closely match their needs.
- The nearest equivalents to a model metadata standard are the metadata standards for describing datasets. However, these were conceived before the needs of model linking emerged. They will need to be built upon and extended to accommodate a description of the model or component process and descriptions the inputs and outputs. It will be important that these descriptions are structured for both human and machine searching.
- The model metadata standards will have a key role in automating the construction of model chains.

- The model metadata standards will have a key role in reducing the chances of making invalid links between models.

  A foreseeable problem in model linking is that even within disciplines the interpretation of the terms used for input and output variables is often highly context specific. When interpretation across disciplines and across natural language boundaries the opportunity for misunderstanding and hence invalid connections between models is considerable. For situations where it is critically important that such misunderstanding does not occur, it will be essential that clear definitions of input and output variables are provided, ideally in both human and machine readable forms and that such variables are assigned globally unique identifiers.

Actions

Task 1: define the purpose of and functions to be supported by a model metadata standard.

Task 2: identify the minimum sets of requirements for the description of a modelling component.

Task 3: draft a model metadata standard following the OGC standards template.

Task 4: define the purpose of a controlled vocabulary for model component input and output variables

Task 5: adopt, adapt or conceive a design for a controlled vocabulary appropriate to the needs of IEM.

Task 6: Research the use of ontologies and artificial intelligence as future options to replace controlled vocabularies and for underpinning improved ways of searching for models and building model chains.

Build the skills base

Situation

When viewed on a world scale and against for example, GIS, IEM skills are concentrated in a relatively small set of groups almost all in the research sphere. For IEM to become widely taken up within a reasonable time span, say 5 – 10 years, all modern channels of communication need to be exploited to disseminate this new knowledge and skills into the wider community. At the time of writing, these could include:

- Inclusion in appropriate university courses at undergraduate and postgraduate level as specialist modules
- E-learning courses
- Summer schools
- Commercial training packages

The process could be stimulated if the IEM product developers were to make versions of their software available under academic licences. It would be similarly helpful to the course developers if any existing teaching aids could be made available to the universities.

Actions

Task 1: identify a set of universities interested in developing IEM courses.

Task 2: identify IEM developers willing to make available software and training material.

Task 3: Develop and deliver course material.

Task 4: Develop equivalent e-learning course material and publish it on the web.

Establish an underpinning R & D programme

Situation

While there has been very significant research investment in the design and development of particular models, there has never been a research programme specific to the issues that arise when two or more are brought together. These could be technical, commercial or legal.

Operational trials for IEM are revealing a large number of often small but important problems that need to be addressed and handled by IEM tools and applications. A typical example might be instability at the point where the models have been joined. Another might be the validity of linking two models whose geographical representations of the river system might not match exactly. Major issues are how to automate the process of finding and linking modelling components. At present, someone must have thought of a potential impact before it can be modelled and assessed. Could artificial intelligence improve on our present ability in this field? These problems need to be catalogued and made the subject of a number of research programmes to be carried out at whichever is the appropriate level – MSc, PhD or postgraduate; national, major EC programme within, say, Horizon 2020 or Belmont Forum funded.

The unknowns are not confined to the technical arena. Although the commercial and legal problems that arise when products from different suppliers are brought to together are not new, they are new for model developers and vendors. No one is yet sure whether linking models will create a set of problems for which there as yet are no precedents for managing.

When the development of IEM began to accelerate around 2000, there was a long period of uncertainty as to who was the 'end user' for the work being undertaken. Although there is a feeling that there is now a better understanding of how an IEM market place could operate, this has never been subjected to professional market research. If the parallels with digital mapping are valid, then the potential future market for IEM products and expertise is very large. It seems logical therefore that there should be series of first exploratory and then more detailed research exercises to assess the market and to recommend how it can be grown, and the investment required to realise it, found.

Actions

Task 1: set up a process for cataloguing known IEM problems.

Task 2: find suitable funding agencies.

Task 3: draft and issue calls for proposals.

Task 4: ensure results are published.

Co-ordination and promotion of collaboration

Situation

It has become clear that there are "islands of excellence" emerging all around the world related to integrated environmental modelling (IEM). For example, in Europe, the OpenMI Association is championing the use of the OpenMI standard for linking models at run-time. In the US, FRAMES is being developed by the US Environmental Protection Agency (US EPA), Common Component Architecture (CCA) by the Community Surface Dynamics Modeling System (CSDMS), and Object Modeling System (OMS) by the US Department for Agriculture (USDA). The complete list is too long to enumerate here; however, the shared objective of all these initiatives is to enable us to better understand and predict the wider implications of environmental events and their management. Various meetings have been held to bring these communities together:

- Environmental Software Systems Compatibility and Linkage Workshop (March 2000)
- Integrated Modeling for Integrated Environmental Decision Making (January 2007; US EPA, 2008)

- Collaborative Approaches to Integrated Modeling: Better Integration for Better Decision-Making (December 2008)
- iEMSs 2010 Conference: Workshop: The Future of Science and Technology of Integrated Modeling (July 2010, Voinov et al., 2010)
- Washington Summit December 2010 that brought these parts of the community together (Moore and Hughes, 2012).
- Arlington Workshop: March 2012 developing a roadmap implementation plan.

From these meetings and others it is clear that there is a nascent community dedicated to using IEM and addressing its challenges.

Action

Task 1: Ensure that the community meets regularly by arranging relevant workshops, conferences or special sessions within existing conferences

Task 2: Promote a Community of Practice for researchers within IEM

Task 3: Utilise existing organisations such as the OGC to provide the forum for meetings

Task 4: Create a business club to allow commercial organisations to communicate

Grow take up by government, industry and the public

Situation

The first tranche of models and modelling tools have reached the stage of being 'near market ready' and suitable for first application by early adopters. They are out of the research phase and have been tested under operational conditions, for example, by the OpenMI-Life project (OpenMI, 2009). As has been explained, there is as yet little IEM experience outside the research and development community, therefore the proposed approach to growing the uptake of IEM is as follows:

1. **Select**, with the aid of end users who have problems that could benefit from the application of IEM, a set of problems
2. **Develop** solutions for those problems based on IEM; the solutions may take the form of products (probably software) or services or a combination of the two.
3. Undertake the **first application** of those products and services with the end users
4. **Package** the solutions in ways that will allow them to be sold in different countries or climatic regions or in support of different manifestations of the same underlying problem, e.g. an agricultural flood estimation model could be repackaged as an urban flooding estimation model as the underlying equations are essentially the same.
5. Explore how the IEM products and services can be extended across and **replicated** in other market sectors, e.g. energy, transport and health which have very similar requirements for integrated modelling.

Actions

Task 1: Commercialisation: identify and fill the gaps in the market

Task 2: Solutions: provide “off the shelf” products which can help decision-makers solve their problems

Task 3: Packaging: bundle up the compositions so that they can be more easily used

Task 4: Dissemination: promote or market the IEM products and solutions that are available.

SUMMARY AND CONCLUSIONS

IEM is essential to help solve the many problems that require an understanding of multiple interacting processes. Many such problems are described in the Belmont Forum "grand challenges" and include providing improved forecasts of future conditions within the environment, particularly with respect to global environmental change. IEM can help suggest best steps to mitigate this change.

IEM technology still requires maturation. This can be considered similar to the path that digital mapping took from its initial hesitant steps in the 1970s. This industry is now a fully-fledged field of commerce with applications that were undreamt of when the first attempts rolled off the printers. To ensure that IEM achieves the same status, the availability of linkable components needs to be improved and their interoperability along with the ability of linking them easily needs to be addressed. Additionally, the description of components along with their ease of discovery through metadata needs to be improved.

The main strands of any strategy to solve the problems of IEM are to raise awareness and build confidence in IEM; ensure availability and accessibility of IEM techniques, tools and standards; establish a minimum set of standards; build the skills base; establish an underpinning R & D programme; co-ordinate and promote collaboration; and, finally, increase use by government, industry and the public. Without the latter, then the demand for IEM will not exist. Once stakeholders and potential users start demanding the use of IEM then this will provide the "pull" for the gaps in IEM to be filled.

An important consideration is "what will success look like?" This can be imagined as a world where IEM solutions are used almost unknowingly by decision-makers of all types, from householders to governmental level. Where complex decisions are solved with its assistance and that resources are much better managed.

ACKNOWLEDGEMENTS

There are many and various people to acknowledge who have contributed to the meetings that have fed into this paper. They are almost too numerous to mention. Some individuals have been mentioned in the text, and included in the references below. However, many remain un-named, for which the authors apologize.

The contribution of one anonymous reviewer and Pierre Glynn (USGS) is gratefully acknowledged. Addressing their comments improved the paper. This paper is published with the permission of the Executive Director, BGS (NERC).

GLOSSARY

Term	Meaning	Further information
ADCIRC	ADCIRC is a system of computer programs for solving time dependent, free surface circulation and transport problems in two and three dimensions	adcirc.org
BIS	UK Government's Department of Business, Innovation and Skills	www.gov.uk/government/organisations/department-for-business-innovation-skills
BGS	British Geological Survey	www.bgs.ac.uk
CCA	Common Component Architecture	www.cca-forum.org
CCMP	Chesapeake Community Modelling Program	
CHyMP	Community Hydrologic Modelling Platform	
CSDMS	Community Surface Dynamic Modelling System	csdms.colorado.edu
CUASHI	Community of Universities for the Advancement of Hydrological Sciences Inc.	www.cuashi.org
EarthCube	EarthCube is a community-led cyberinfrastructure initiative for the geosciences.	earthcube.org
E-infrastructure	This refers to the ecosystem of resources that allows distributed collaboration and computation, large-scale simulation and analysis, and fast access to (large) data collections (well organized according to accepted standards and with rich metadata), analytical and visualization services and facilities	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/249474/bis-13-1178-e-infrastructure-the-ecosystem-for-innovation-one-year-on.pdf
EPSRC	Engineering and Physical Sciences Research Council	www.epsrc.ac.uk
ESMF	Earth System Modelling Framework	www.earthsystemmodeling.org
ESRI	Environmental Systems Research Institute	www.esri.com
FRAMES	Framework for Risk Analysis in Multimedia Environmental Systems	
GIS	Geographic Information System	www.esri.com/what-is-gis
GRASS	Geographic Resources Analysis Support System	grass.osgeo.org/grass64
GSFLOW	USGS coupled groundwater and surface-water flow model	water.usgs.gov/ogw/gsflow/
HSPF	USGS Hydrological Simulation Program - FORTRAN	water.usgs.gov/software/HSPF/
ICSU	International Council of Science	www.icsu.org
ISCMEM	Interagency Steering Committee on Multi-media Environmental Modeling	
IHM	Penn State Integrated Hydrologic Model	www.pihm.psu.edu
LOIS	Land Ocean Interaction Study - NERC funded research project	www.bodc.ac.uk/projects/uk/lois/

MapWindow	GIS project that includes a free desktop geographic information system application with an extensible plugin architecture	www.mapwindow.org
MODFLOW	USGS Groundwater model code	water.usgs.gov/ogw/modflow/
NERC	Natural Environment research Council	www.nerc.ac.uk
NSF	National science Foundation	www.nsf.gov
OGC	Open Geospatial Consortium	www.opengeospatial.org
OpenMI	Open Modelling Interface	www.openmi.org
OMS	Object Modelling System	http://nrrc.ars.usda.gov/ModelFrameworks/ObjectModelingSystem.aspx
PRMS	USGS Precipitation-Runoff Modeling System	wwwbrr.cr.usgs.gov/projects/SW_MoWS/PRMS.html
SWAT	Soil & Water Assessment Tool	swat.tamu.edu/
USACE	United States Army Corps of Engineers	www.usace.army.mil
US EPA	United States Environmental Protection Agency	www.epa.gov
USGS	United States Geological Survey	www.usgs.gov
VPH	Virtual Physiological Human	www.vph-institute.org

REFERENCES

Ahuja, L.R., Ascough II, J.C., David, O., 2005. Developing natural resource models using the object Modelling system: feasibility and challenges. Advances in Geosciences 4, 29e36. http://hal.archives-ouvertes.fr/docs/00/29/68/06/PDF/adgeo-4-29-2005.pdf.

Argent, R. M., Voinov, A., Maxwell, T., Cuddy, S. M., Rahman, J. M., Seaton, S., Vertessy, R. A., and Braddock, R. D., 2006. "Comparing modelling frameworks – A workshop approach." Environmental Modelling & Software, 21(7), 895-910.

Argent, R.M., Perraud, J.-M., Rahman, J.M., Grayson, R.B., Podger, G.M., 2009. A new approach to water quality modelling and environmental decision support systems. Environmental Modelling & Software 24 (7), 809e818.

Arnold, J. G. and Srinivasan, R. and Muttiah, R. S. and Williams, J. R., 1998. LARGE AREA HYDROLOGIC MODELING AND ASSESSMENT PART I: MODEL DEVELOPMENT. JAWRA Journal of the American Water Resources Association.vol. 34(1).

Babel, 2004. Lawrence Livermore National Laboratory homepage http://www.llnl.gov/CASC/components/babel.html

Babendreier, J.E., Castleton, K.J., 2005. Investigating uncertainty and sensitivity in integrated, multimedia environmental models: tools for FRAMESe3MRA. Environmental Modelling & Software 20 (8), 1043e1055.

Barthel, R., Janisch, S., Schwarz, N., Trifkovic, A., Nickel, D., Schulz, C., Mauser, W., 2008. An integrated modelling framework for simulating regional-scale actor responses to global change in the water domain. Environmental Modelling & Software 23 (9), 1095e1121.

Bastin L,, Cornford D., Jones R., Heuvelink G.M.B., Pebesma E., Stasch C., Nativi S., Mazzetti P., Williams M.. 2013. Managing uncertainty in integrated environmental modelling: The UncertWeb framework. Environmental Modelling & Software, 39. Pp 116-134.

Bernholdt, D.E., B.A. Allan, R. Armstrong, F. Bertrand, K. Chiu, T.L. Dahlgren, K., Damevski, W.R. Elwasif, T.G.W. Epperly, M. Govindaraju, D.S. Katz, J.A. Kohl, M., Krishnan, G. Kumfert, J.W. Larson, S. Lefantzi, M.J. Lewis, A.D. Malony, L.C., McInnes, J. Nieplocha, B. Norris, S.G. Parker, J. Ray, S. Shende, T.L. Windus and S. Zhou, 2004..A component architecture for high performance scientific computing, International Journal of High Performance Computing Applications https://e-reportsext.llnl.gov/pdf/314847.pdf ACTS Collection Special Issue, 75 pp.

BIS, 2011, A Strategic Vision for UK e-Infrastructure: A roadmap for the development and use of advanced computing, data and networks. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/32499/12-517-strategic-vision-for-uk-e-infrastructure.pdf (accessed 4th April 2014)

BIS, 2013. E-infrastructure: The ecosystem for innovation One year on, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/249474/bis-13-1178-e-infrastructure-the-ecosystem-for-innovation-one-year-on.pdf (accessed 3rd March 2016)

Coppock J.T. & Rhind D.W., 1991, "The History of GIS". In Maguire D.J., Goodchild M.F., and Rhind D.W. (editors) Geographical Information Systems : Principles and Applications, Volume 1, pages 21-43.

DAVID, O., ASCOUGH II, J C., LLOYD, W., GREEN, T R., ROJAS, K W., LEAVESLEY, G H. & AHUJA, L R. 2013. A software engineering perspective on environmental modeling framework design: The Object Modeling System. Environmental Modelling and Software, 39 201-213.

Donigian, A. S., Jr.; Bicknell, B. R.; Imhoff, J. C., 1995, Hydrological Simulation Program - Fortran (HSPF). In Computer models of watershed hydrology. (Ed. Singh, V. P.), pp 395-442. ISBN0-918334-91-8

EPSRC, 2014. E-infrastructure roadmap, https://www.epsrc.ac.uk/newsevents/pubs/e-infrastructure-roadmap/ (accessed 3rd March 2016)

~~University of Osnabrück, Germany. http://www.iemss.org/iemss2004/pdf/integratedmodelling/kralusin.pdf, 6 pp.~~

FRAMES, 2009 http://www.epa.gov/ATHENS/research/Modelling/3mra.html.

Gaber N., Laniak G., and Linker L. 2008. Integrated Modeling for Integrated Environmental Decision Making. EPA100/R-08/010

Glynn, P. D. (2015). Integrated Environmental Modelling: human decisions, human challenges. Geological Society, London, Special Publications, 408, SP408-9.

Goodchild, M. F. (1988). Stepping over the line: technological constraints and the new cartography. The American Cartographer, 15(3), 311-319.

Gregersen, J. B., Gijsbers, P. J. A., & Westen, S. J. P. (2007). OpenMI: Open modelling interface. Journal of hydroinformatics, 9(3), 175-191.

Harbaugh A.W., 2005. MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model—the Ground-Water Flow Process. U.S. Geological Survey Techniques and Methods 6-A16

Harpham Q., Cleverly P. and Kelly D., 2013. The FluidEarth 2 Implementation of OpenMI 2.0, Journal of Hydroinformatics. doi:10.2166/hydro.2013.190

Harpham, Q., & Danovaro, E. (2015). Towards standard metadata to support models and interfaces in a hydro-meteorological model chain. Journal of Hydroinformatics, 17(2), 260-274.ICSU, 2010. Earth System Science for Global Sustainability: The Grand Challenges. International Council for Science, Paris. ISBN: 978-0-930357-73-3

Jackson, C. FluidEarth meeting, 2014

Kralisch, S., Krause, P., David, O., 2004. Using the Object Modelling System for hydrological model development and application. In: Proceedings of the iEMSs 2004 International Conference. University of Osnabrück, Germany. http://www.iemss.org/iemss2004/pdf/integratedmodelling/kralusin.pdf, 6 pp.

Laniak, Gerard F.; Olchin, Gabriel; Goodall, Jonathan; Voinov, Alexey; Hill, Mary; Glynn, Pierre; Whelan, Gene; Geller, Gary; Quinn, Nigel; Blind, Michiel; Peckham, Scott; Reaney, Sim; Gaber, Noha; Kennedy, Robert; Hughes, Andrew. 2013 Integrated environmental modeling : a vision and roadmap for the future. Environmental Modelling & Software, 39. 3-23. 10.1016/j.envsoft.2012.09.006

LLOYD, W., DAVID, O., ASCOUGH II, J C., ROJAS, K W., CARLSON, J R., LEAVESLEY, G H., KRAUSE, P., GREEN, T R. & AHUJA, L R. 2011. Environmental modeling framework invasiveness: Analysis and implications. Environmental Modelling & Software 26, 1240-1250.

Macdonald, D.; Dixon, A.; Newell, A.; Hallaways, A.. 2012 Groundwater flooding within an urbanised flood plain. Journal of Flood Risk Management, 5 (1). 68-80. 10.1111/j.1753-318X.2011.01127.x

Machlis, G. E., & McNutt, M. K. (2010). Scenario-building for the Deepwater Horizon oil spill. Science, 329(5995), 1018-1019.

Makropoulos, C., Safiolea, E., Baki, S., Douka, E., Stamou, A., & Mimikou, M. (2010, July). An integrated, multi-modelling approach for the assessment of water quality: lessons from the Pinios River case in Greece. In Proc. International Congress on Environmental Modelling and Software (iEMSs 2010), Ottawa, Canada.

Mansour, Majdi; Mackay, Jonathan; Abesser, Corinna; Williams, Ann; Wang, Lei; Bricker, Stephanie; Jackson, Christopher. 2013 Integrated Environmental Modeling applied at the basin scale: linking different types of models using the OpenMI standard to improve simulation of groundwater processes in the Thames Basin, UK. In: MODFLOW and More 2013: Translating Science into Practice, Colorado, USA, 2-5 June 2013. (Unpublished) - http://nora.nerc.ac.uk/501789/

Markstrom, S.L., Niswonger, R.G., Regan, R.S., Prudic, D.E., and Barlow, P.M., 2008, GSFLOW-Coupled Ground-water and Surface-water FLOW model based on the integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005): U.S. Geological Survey Techniques and Methods 6-D1, 240 p.

Moore, Roger; Hughes, Andrew; (Eds) with contributions from Gaber, Noha; Geller, Gary; Glynn, Pierre; Laniak, Gerry; Voinov, Alexey; Whelan, Gene. 2012 International Summit on Integrated Environmental Modeling. US Environmental Protection Agency, 62pp. (OR/12/087) (Unpublished) - http://nora.nerc.ac.uk/21007/

Moore R., P. Gijsbers, D. Fortune, J. Gergersen and M. Blind, 2005. OpenMI Document Series: Part A Scope for the OpenMI (Version 1.0), HarmonIT

Martin K. van Ittersum, Frank Ewert, Thomas Heckelei, Jacques Wery, Johanna Alkan Olsson, Erling Andersen, Irina Bezlepkina, Floor Brouwer, Marcello Donatelli, Guillermo Flichman, Lennart Olsson, Andrea E. Rizzoli, Tamme van der Wal, Jan Erik Wien, Joost Wolf, Integrated assessment of agricultural systems – A component-based framework for the European Union (SEAMLESS), Agricultural Systems, Volume 96, Issues 1–3, March 2008, Pages 150-165, ISSN 0308-521X, http://dx.doi.org/10.1016/j.agsy.2007.07.009. (http://www.sciencedirect.com/science/article/pii/S0308521X07000893)

Natural Environment Research Council. 2009. Natural Environmental Research Council Science Information Strategy. (Swindon: Natural Environment Research Council)

NSF, 2011. The "Earth Cube" - Towards a National Data Infrastructure for Earth System Science http://www.nsf.gov/pubs/2011/nsf11065/nsf11065.jsp?org=NSF (accessed 4th April 2014)

OpenMI, 2009. The Open-MI life project website http://www.openmi-life.org/, 2009.

Oreskes, N. 2003. The role of quantitative models in science, In: Canham, C.D., Cole, J.J. & Lauenroth, W.K. (eds.), Models in Ecosystem Science, 13-31, Princeton: Princeton University Press.

PECKHAM, S D., HUTTON, E W H. & NORRIS, B. 2013. A component-based approach to integrated modeling in the geosciences: The design of CSDMS. Computers and Geosciences, 53, 3-12.

Rogers, Everett (16 August 2003). Diffusion of Innovations, 5th Edition. Simon and Schuster. ISBN 978-0-7432-5823-4.

Safiolea, E.; Makropoulos, C.; Mimikou, M, 2009. Remove from marked Records Benefits and challenges in integrated water resources modeling using OpenMI. Integrating water systems. Proceedings of the Tenth International Conference on Computing and Control for the Water Industry, CCWI 2009 - 'Integrating Water Systems', Sheffield, UK, 1-3 September 2009 2010 pp. 481-484

Sophocleous M. and Perkins S.P., 2000. Methodology and application of combined watershed and ground-water models in Kansas. Journal of Hydrology. 236 pp185–201.

Taylor P., Cox S. and Walker G., 2012. Adapting standard information models for water data. 10th International Conference in Hydroinformatics, HIC2012, Hamburg, Germany.Voinov, A. A., C. DeLuca, R. R. Hood, S. Peckham, C. R. Sherwood, and J. P. M. Syvitski, 2010. A Community Approach to Earth Systems Modelling, Eos Trans. AGU, 91(13), doi:10.1029/2010EO130001.

US EPA 2008, Integrated Modeling for Integrated Environmental Decision Making. EPA100/R-08/010. Washington, DC. Office of the Science Advisor

US Government, 2012. http://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government.html?goback=.gde_3427896_member_120023144 (accessed 4th April 2014)

Viceconti M and Clapworthy G, 2011. VPH-FET Research Roadmap: Advanced Technologies for the Future of the Virtual Physiological Human.

VPH, 2009. A vision and a strategy for the Virtual Physiological Human in 2010 and beyond. http://www.vph-institute.org/upload/vph-vision-strategy-submitted-141209-4_519244d49f91e.pdf (accessed 4th April 2014)

Voinov, A. A., C. DeLuca, R. R. Hood, S. Peckham, C. R. Sherwood, and J. P. M. Syvitski, 2010. A Community Approach to Earth Systems Modeling, Eos Trans. AGU, 91(13), doi:10.1029/2010EO130001.

Voinov, A., Kolagani, N., McCall, M. K., Glynn, P. D., Kragt, M. E., Ostermann, F. O., ... & Ramu, P. (2016). Modelling with stakeholders–Next generation. Environmental Modelling & Software, 77, 196-220.

Voinov A. and Shugart H.H., 2013, 'Integronsters', integral and integrated modeling. Environmental Modelling & Software 39 (1) pp149-158.

Warner, J.C., Perlin, N., Skyllingstad, E.D., 2008. Using the Model Coupling Toolkit to couple earth system models. Environmental Modelling & Software 23 (10e11), 1240e1249.

Whelan, G., M.E. Tryby, M.A. Pelton, J.A. Soller, and K.J. Castleton. 2010. "Using an Integrated, Multi-disciplinary Framework to Support Quantitative Microbial Risk Assessments."

Wilkinson, W. B., Leeks, G. J. L., Morris, A., & Walling, D. E. (1997). Rivers and coastal research in the Land Ocean Interaction Study. Science of the Total Environment, 194, 5-14.

Zhang, Jing, Ross, Mark A., Geurink, Jeffrey S. and Gong, Huili, 2009. Modelling vadose zone moisture dynamics in shallow water table settings with the Integrated Hydrologic Model: field-scale testing and application. Water and Environment Journal, vol. 24 (1), pp 9-20

Box 1: Challenges requiring IEM

- What is the risk to infrastructure of multiple natural disasters?
- What would be the impact of a “Carrington” type space weather event on electrical distribution systems and civil society?
- What would be the impact of leakage from an oil and gas well in UK waters on the national economy, coastal and marine biodiversity and the well-being of the population affected?
- How will climate change affect:
    - the global distribution of malaria?
    - the incidence of road and rail closures due to landslides?
    - the frequency of drought conditions and hence security of water supply and biological diversity?
    - the number of insurance claims for properties lost to inundation and cliff erosion?
- How economically viable will it be to store CO2 in a geological formation under the North Sea?
- What impact will a volcanic ash cloud from an Icelandic volcano have on civil aviation and subsequent economic losses for a country?
- Rainfall can trigger eruptions….?
- …eruptions can affect hydrology?
- How does spatial variability of rainfall interact with geology and affect flooding?