

---

## The environmental eScience revolution

R.J Gurney, N.S Badcock, E.A Garnett and H.C.J Godfray

*Phil. Trans. R. Soc. A* 2009 **367**, 803-813

doi: 10.1098/rsta.2008.0253

---

### References

[This article cites 22 articles](#)

<http://rsta.royalsocietypublishing.org/content/367/1890/803.full.html#ref-list-1>



This article is free to access

### Rapid response

[Respond to this article](#)

<http://rsta.royalsocietypublishing.org/letters/submit/roypta;367/1890/803>

### Subject collections

Articles on similar topics can be found in the following collections

[optics](#) (14 articles)

[climatology](#) (48 articles)

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

---

To subscribe to *Phil. Trans. R. Soc. A* go to:  
<http://rsta.royalsocietypublishing.org/subscriptions>

---

## INTRODUCTION

# The environmental eScience revolution

---

## 1. Introduction

Environmental science is in the midst of a revolution, catalysed by advances in the computer and information sciences. The ever increasing power of computers is widening the range of problems that can be explored by intensive computational analysis, while modern connectivity at both high and low bandwidths is allowing scientists to work together in new ways. The volume of data being generated by different scientific fields is increasing at an exponential or near-exponential rate, and this has spurred new ways to handle, analyse and visualize the torrent of information. As working scientists with institutional and normally personal webpages, we are more visible to the general public, and the Web offers new opportunities to engage with a general audience as well as to recruit the wider public to participate in scientific research. Without wanting to put forward a formal definition, we link all these activities together under the informal rubric of ‘eScience’.

This revolution in the environmental sciences comes at a time when the field is facing huge challenges of very high societal relevance. Predicting the magnitude of climate change requires an ability to model the Earth system at levels of precision and accuracy never previously attempted. Coping with the consequences of climate change, as well as the many other ways that man is changing the global and local environments, involves monitoring the environment, predicting possible changes using very advanced models and making complex decisions based on large amounts of information. Minimizing man’s footprint involves using the Earth’s resources sustainably, a challenge for physical scientists who seek to harness renewable energy resources while maximizing efficiency and reducing emissions, and a challenge to biological scientists involved in resource management. Moreover, these are not just science problems and any viable solution must also take into account economic, social and political realities: natural and social scientists have to learn to work together more often. And we share the planet with an immense number of other species. Some of this biodiversity provides economically significant ecosystem services and it is in our interests to maintain and protect it. But it is foolish to suppose that there is inevitably an economic argument to preserve natural habitats, and instead we should be aware of a duty to future generations to destroy as little as possible of the Earth’s biodiversity. It is our contention that eScience will play a critical role in addressing all these challenges, as well as improving the fundamental science base that underpins all good applied science.

One contribution of 24 to a Discussion Meeting Issue ‘The environmental eScience revolution’.

There have already been some truly impressive applications of eScience in environmental research, some of which are described in the papers that follow this introduction. But we believe that many people in the field are not yet using eScience to the extent that would be optimal for their research area. There are several reasons for this. First, many eScience developments require interdisciplinary collaboration between environmental and computer scientists, and interdisciplinarity is always difficult within the subject, institutional and funding constraints that people work under. Environmental scientists may just not get to meet computer scientists, or read (or understand) their research, and vice versa. Other environmental scientists may not be aware of techniques already in existence or there may be a financial or skill barrier preventing their take-up. Research agencies and scientific societies must play an important role in bringing communities together, and disseminating knowledge and expertise, one of the motivations for the Discussion Meeting that gave rise to the collected papers in this issue.

The environmental sciences are of course not alone in being transformed by eScience. As is now the stuff of legend, the World Wide Web originated as a means of fostering collaboration within the high-energy physics community at the CERN particle collider facility outside Geneva. Other areas of the physical sciences have driven this forward, such as sharing very large amounts of information in astronomy. In the biological sciences, the explosion in the quantity of information produced by molecular biology has seen the development of the whole new field of bioinformatics. Part of this subject includes the organization, retrieval and analysis of large datasets that may be of relatively simple structure (e.g. DNA sequences, the spatial location of atoms of protein) or very complex (annotated gene function described by a large structured set of terms). Bioinformatics also relies on advances in modern computer science to allow it to solve the structures of hugely complex molecules, or in systems biology to model the dynamics of the large number of molecule types interacting together within the cell. There are undoubtedly going to be important lessons to be learned for the environmental sciences from studying these cognate fields.

Science may also learn from commerce and entertainment. Social networking sites have their origins in tools to enable scientists to work together, but today Facebook, MySpace and related sites are developing extraordinarily sophisticated software that is likely to be of great value to scientists when exchanging ideas. Similarly, the data retrieval and geographical information systems developed by Google and other search-engine companies are increasingly being used in scientific applications. Finally, companies working in the special effects and computer gaming industries are developing extraordinary tools for visualization and machine–human interactions. Applications are being seen in environmental sciences to visualize data and model output, and some examples are included in this issue.

Traditionally, there have been few interactions between computer scientists and environmental scientists. This issue arises from a Discussion Meeting held at the Royal Society in April 2008. The goals of the Discussion Meeting were to look back at recent progress in environmental eScience and in particular to explore new and exciting ways in which the subject might move forward. The Discussion Meeting was explicitly structured to be interdisciplinary, mixing, on the one hand, biological and physical environmental scientists and, on the other, environmental scientists and people working in computer and information sciences. It also

included speakers and participants from related fields such as molecular bioinformatics, and, in both the formal sessions and an evening programme involving demonstrations and short talks, people from industry and policy development as well as from the scientific community. The Discussion Meeting was sponsored by the UK Natural Environment Research Council (NERC) and by Microsoft Research, and was international in scope and ambition.

A further catalyst for this Discussion Meeting was the conclusion of the UK's environmental eScience programme coordinated by NERC. As NERC's Chief Executive, Alan Thorpe, describes in the Preface to this issue (Thorpe 2009), this was part of a larger British initiative on eScience. All of the successful NERC projects brought together computer and environmental scientists (46% of the principal and co-investigators were classified as computer scientists and 54% as environmental scientists). Of the 272 refereed papers that have been published so far, 37.5 per cent have appeared in the computer science literature. The programme can thus fairly claim to be interdisciplinary.

There have been parallel developments elsewhere in the world, engaging environmental and computer scientists together with varying degrees of success. A particularly similar initiative was launched in the USA by the National Science Foundation, called the Cyberinfrastructure Program. Like the UK programme, this has tried to bring together applications scientists, including environmental scientists, with computer scientists, and some of the resulting research is shown here. The most successful programmes are those which have balanced inputs from both environmental and computer scientists, so novel research can be done in both areas. Although not part of the same funding streams, there have also been some very significant developments in high-resolution global modelling, using some of the world's largest supercomputers, such as the Japanese Earth Simulator. Many of the dedicated programmes are coming to an end or finishing their first phase, and there is about to be a further set of major investments in supercomputing for environmental sciences. It is an opportune moment to review developments and look to the future.

Speakers at the Discussion Meeting were given a choice of contributing papers to this issue and not all the talks are represented here. We also commissioned some further articles to cover topics that would otherwise be missing. In addition to talks, the Discussion Meeting also included extended poster and demonstration sessions, and we invited a selection of the people responsible for them to contribute shorter papers to this issue. All eight consortia that were part of the NERC environmental eScience programme are represented here, and for ease of reference these are summarized in [table 1](#).

We think a number of general themes emerge from the papers collected here. The first of these is the critical importance of predicting complex nonlinear systems in the environmental sciences. Successful prediction first requires a fundamental understanding of the dynamics of the underlying environmental processes. In some cases, the dynamic process may be well understood and the challenge is working on large spatial scales, but in other cases there may be substantial uncertainty about the structure of the underlying dynamics. Once the process is clear, there are a series of computing issues involving developing the model code and in particular optimizing its performance, as well as obtaining the technical infrastructure to run the models, often using very high-performance computing. Once the model has been run, handling, analysing and

Table 1. NERC environmental eScience consortia represented by papers in this issue (further information on these projects can be found at <http://www.nerc.ac.uk/research/programmes/escience>).

| principal investigator | NERC environmental eScience project   | related papers in this issue   |
|------------------------|---|--|
| Dr M. R. Allen         | <a href="http://climateprediction.net">climateprediction.net</a> : distributed computing for global climate research  | Frame (2009)   |
| Prof. M. T. Dove       | eMinerals: applications of grid-enabled science to understand the environment from the molecular level  | Frame <i>et al.</i> (2009), Salje <i>et al.</i> (2009), Walker <i>et al.</i> (2009) and White <i>et al.</i> (2009) |
| Prof. K. Haines        | GODIVA: grid for ocean diagnostics, interactive visualization and analysis  | Blower <i>et al.</i> (2009)  |
| Prof. T. Lenton        | GENIEfy: creating a grid-enabled integrated Earth system modelling framework for the community (Earth system model for long-term and palaeoclimate studies) | Lenton <i>et al.</i> (2009)  |
| Dr B. Lawrence         | the NERC DataGrid and associated toolkit  | Latham <i>et al.</i> (2009) and Lawrence <i>et al.</i> (2009)  |
| Dr M. J. Scoble        | CATE: creating a taxonomic eScience   | Clark <i>et al.</i> (2009)   |
| Dr J. Holt             | GCOM: global coastal-ocean modelling  | Holt <i>et al.</i> (2009)  |
| Prof. K. Haines        | GCEPS: grid for coupled ensemble prediction studies   | Haines <i>et al.</i> (2009), Hermanson & Sutton (2009) and Putt <i>et al.</i> (2009)                               |

visualizing the data can present further major eScience challenges, as does obtaining the high-quality datasets required to initialize and validate the models.

A second theme that we believe emerges from the collected papers is the need frequently to overcome a different type of hurdle. Much scientific information, including the raw scientific literature itself, is heterogeneous and in a state that cannot be easily machine read or digitized. Getting these data into a form that can be used more efficiently can be expensive and time consuming, and modern advances in automatic data extraction are going to be important to take advantage of the observations, though not sufficient to eScience-enable these fields. Once digitized, developments in semantic Web and related technologies will be valuable in accessing and understanding the information. However, in some areas of the environmental sciences, perhaps particularly those without a strong quantitative and technological tradition, cultural change may be as important as new techniques in introducing eScience.

In the remainder of this introduction, we provide an overview of the papers in the issue arranged under broad thematic categories. We also mention some areas of research not covered by the papers in this issue. Research in the environmental sciences not only has been the user of new techniques developed by computer scientists but also has led to innovations of new tools and methodologies, so that the interactions between environmental and computer scientists have been exactly that, interactions that have led to

developments in both fields. The new technologies are primarily in four areas, and are reflected in the papers in this issue. These are the ability to handle very large amounts of data and model output; to construct, organize and analyse datasets that are large, but not enormous, but which pose particular challenges owing to their heterogeneity; to facilitate ever more complex models and multiple model runs; and to enable scientists to work together across disciplines.

## 2. Climate and weather modelling

eScience has been at the forefront of developments in our ability to predict changes to our climate, and several of the papers in this issue illustrate the latest developments in climate prediction. Climate prediction involves very large global models coupling ocean and atmosphere with the land surface and cryosphere; many processes have to be included, such as changes in atmospheric chemistry or in land cover. This is extremely challenging computationally, and many of the fastest supercomputers are used in making the predictions. There is a trade-off between (i) model realism and explicitly including processes that often involve high spatial resolution, (ii) longer time integrations and (iii) running many versions of a model with slightly different parameter values or boundary conditions in order to quantify the uncertainties in predictions. There is insufficient computing power available to improve all three simultaneously. Papers here represent a series of different ways of addressing this trade-off. [Slingo \*et al.\* \(2009\)](#) and [Washington \*et al.\* \(2009\)](#) both report on weather-resolving models at very high spatial resolution, and on adaptive meshing techniques to allow similar high-resolution modelling in the ocean. [Holt \*et al.\* \(2009\)](#) describe similar techniques applied to ocean models alone. [Slingo \*et al.\* \(2009\)](#) and [Washington \*et al.\* \(2009\)](#) also lay down the challenge of how to represent biological processes in global models using innovative mathematical techniques. [Slingo \*et al.\* \(2009\)](#) in particular make a persuasive case for the need for petascale computing to understand properly the nonlinear scaling required to represent both physical and biological processes adequately at global scales. There is also the need for advanced techniques to handle the complex modelling environment, as well as the very large volume of results, and one approach to help make this easier is illustrated by [Bretherton \*et al.\* \(2009\)](#).

[Frame \(2009\)](#) uses a set of extremely large ensembles of a climate model to investigate the level of uncertainty in climate predictions. The model used is simplified so it can be run on ordinary desktop computers. In conjunction with the BBC [climateprediction.net](http://climateprediction.net) programme, many ensembles were created using personal computers belonging to schools, businesses and members of the general public on every continent in the world, including Antarctica. Not only has this been excellent for engaging the public in environmental science, but it has also led to improved understanding of the range of possible anthropogenic warming. An outstanding question is how these limited ensembles relate to the latest high-resolution models, because the relationship is likely to be nonlinear. [Lenton \*et al.\* \(2009\)](#) also use a simpler model to run large ensembles over long time periods, to examine the conditions under which the North Atlantic thermohaline circulation will collapse, in order to see if an



early warning system can be designed to warn of a collapse. The potential for such an early warning system appears limited, but in itself this is a useful result; again, improved understanding will come from relating the results from the hierarchy of models with different resolutions and physical realism to each other.

Weather forecasting and seasonal climate prediction are initial-value problems, while long-term climate prediction is a boundary-value problem, and the computing challenges are somewhat different. This difference is well illustrated by [Droegemeier \(2009\)](#), who shows a set of examples that bring together observations and advanced models in real time to improve predictions. Along with many other papers in this issue, he uses computing clusters to speed up modelling. He also shows how advanced visualization and an easy-to-use interface are important for the efficient use both of the novel tools and of the results. [Froude \(2009\)](#) uses techniques to track storms in weather forecast models to show that, while storm tracks are well reproduced, the speed and intensity of storms are less well modelled, because the vertical structure, specifically storm tilt, is not modelled well enough over the oceans. This result points to the need for improved satellite retrievals of the atmosphere over the oceans. Weather forecast models usually make predictions of the atmosphere only a few days ahead, but we can now run fully coupled models of both atmosphere and ocean that can make predictions with some skill over seasons if we have sufficiently good knowledge of the initial state of the atmosphere and oceans. Both [Hermanson & Sutton \(2009\)](#) and [Putt \*et al.\* \(2009\)](#) use a coupled model, HadCM3, to investigate whether such a fully coupled model can be used to make useful seasonal predictions, using an ensemble approach of running the model several times to understand the errors in the initial conditions. [Hermanson & Sutton \(2009\)](#) look at ocean heat anomalies, and show that some aspects of the model carry predictive skill 2 years ahead, while [Putt \*et al.\* \(2009\)](#) examine seasonal snow cover in El Niño events and also find that there is predictability into the second year of snow cover. [Haines \*et al.\* \(2009\)](#) have written a system that allows the Hadley Centre decadal climate prediction system to be run across a compute grid, so that a far larger ensemble prediction can be made. The results show an impressive improvement in prediction skill and show how predictability depends on the knowledge of the initial state, so work stating average predictive skill in a model may be misleading. There is much more seasonal predictive skill in a fully coupled model of atmosphere and ocean, including for instance the meridional overturning circulation, whereas models that only include sea surface temperature and other variables driven largely by atmospheric processes are shown to be inherently less predictable, and so have less seasonal predictive skill.

### 3. Modelling the biosphere

While modelling physical processes has many challenges, in many cases the underlying physics is relatively well understood. Thus, the Navier–Stokes equations lie at the heart of most atmospheric and ocean circulation models. The challenge comes in approximating very nonlinear processes across a wide range of space and time scales. Theoretical ecologists usually face different problems. First, the equations governing population (and evolutionary) dynamics are very

variable, and for any particular system are often only partially understood. Second, even non-spatial ecological numbers often have high dimensionality, involving numerous interacting species, with complex age structure and time lags. Third, in common with physical environmental models, obtaining the data to parametrize and validate the model can be very difficult, especially as many processes occur over long time scales. There are many examples of the analysis of biological dynamics in the field, but, not surprisingly, these tend to involve relatively simple interacting communities, rather than very complex systems such as whole communities.

Similar problems are faced in other fields of biology, not least in understanding cellular and molecular dynamics in the (relatively) new field of system biology. Recent advances in computer science and related fields of mathematics such as complexity theory may be helpful here, and cross-disciplinary interactions across the spectrum of biology need to be encouraged. At the Discussion Meeting, Stephen Pacala (Princeton) talked about some of the most sophisticated population models of large communities of interacting organisms. His group has developed population models of the growth of temperate woodlands that mechanistically include the main ecological interactions between different tree individuals and populations: competition for limiting resources, in particular light. The models can explain current patterns in tree distribution and relative abundance, and can also explain past patterns revealed by analysis of fossil pollen. As well as giving insights into ecological processes, the models can also be used to ask questions relevant to climate change, for example, 'Can boreal forests help to explain gaps in the global carbon budget?', 'Will elevated CO<sub>2</sub> levels increase tree growth?' and 'Will this on balance remove carbon from the atmosphere?'. Extending this work to other ecosystems, particularly tropical ecosystems, will require new modelling and computation approaches to overcome the problem of much higher species richness.

Purely biological models alone cannot explain major biogeochemical processes; approaches are required that integrate both ecological models and land use patterns with physical processes. Work in this area is most advanced with studies of the carbon cycle, where a combination of modelling with remote sensing and assimilation of data from eddy flux tower networks is starting to explain major land-based fluxes, although there are considerable worries that not all processes are well described owing to the biases often shown by the models when compared with observations. Such projects analysing heterogeneous data sources have particularly benefited from eScience research, and are even starting to use mathematical techniques of data assimilation, though there is a need to develop new techniques that allow the bias to be studied more explicitly. A different aspect of coupling physical and biological processes is discussed here by [Holt \*et al.\* \(2009\)](#), who describe a new modelling system that attempts to couple a physical model of the coastal ocean with a biological model of the same areas. The model appears to reproduce areas of very high primary productivity, but validation of the model output is extremely challenging. The physical model is easier to validate, and appears to be working well. Obtaining and validating the necessary biological data will be very difficult, particularly as the relevant satellite data cannot always give the necessary biological detail.



#### 4. Environmental bioinformatics

What is sometimes called the ‘bioinformatics crisis’ describes the avalanche of molecular data that began when automatic sequencing machines were introduced in the late 1980s and continued with the sequencing of the first large genomes in the 1990s. Genomics was followed by transcriptomics, proteomics, metabolomics, the postfix ‘-omics’ indicating high-throughput acquisition of different classes of molecular information. The molecular biology community understood from the start the importance of eScience in organizing these data, and today several highly sophisticated Web portals for accessing and downloading this information exist. One of the pioneers of molecular bioinformatics, Michael Ashburner from Cambridge, gave a talk entitled ‘It’s the semantics ...’ at the Discussion Meeting. He described how a structured ontology or controlled vocabulary had been set up to help organize molecular genetic information. This described, on the one hand, gene function and, on the other, all the ancillary information required to interpret gene function, from morphology to ecology. Efforts had been made to coordinate terminology across different sequencing projects, and more broadly with other biomedical databases.

The molecular biology community shares with environmental biologists the problem of classifying large heterogeneous datasets. Compared with environmental biologists this field’s take-up of eScience has been much faster. One reason for this is that biomedical work is comparatively well funded, but the molecular biology community has also had the advantage that ‘-omics’ data are often relatively new and have gone ‘straight to digital’. In environmental biology many valuable data are older and in paper form, and hence expensive to digitize. Though there are projects working on structured ontologies in ecology, for example, these have yet to enter the field’s mainstream.

The area of environmental biology that has perhaps made most progress is taxonomy and systematics. So far most of the digitization has been concentrated on collections in major museums and herbaria. An organization called Biodiversity Information Standards (often referred to by its old acronym, TDWG) has pioneered a structured vocabulary to describe taxonomic concepts, geographical information and morphology. There is also much interest in how taxonomy itself, the field invented by Linnaeus to solve an eighteenth-century bioinformatics crisis, might move to the Web. [Clark \*et al.\* \(2009\)](#) in this issue describe a pilot project that explored how the taxonomy of two representative groups of plants and animals might be moved completely to the Web. More importantly, the project considered how taxonomy itself might be conducted on the Web, so that the site would evolve and reflect current taxonomic thinking. A further feature of the project is to allow multiple taxonomic hypotheses from different researchers, but to present a current consensus classification for a broader audience. Taxonomy is an area of biology whose outputs can be used by a very large audience, including non-professional scientists, and also an area where informed amateurs can still make major contributions. A project to facilitate this is the Encyclopedia of Life, the brain-child of E. O. Wilson, which was launched in early 2008. Its chief executive, Jim Edwards, described the project and illustrated how it would call on existing eScience expertise, but would also develop new techniques in particular to engage the broadest audience. Roderick Page from Glasgow also spoke on environmental informatics, and in

particular described the unique place that Linnaean binominal names have in retrieving information about living organisms. He explored not only how these can be used to harvest information from diverse sources through the Web, but also how the adoption of novel ways of indexing data using digital unique identifiers could improve and simplify this.

## 5. Mineral simulations

Numerous problems in environmental science involve scaling across many spatial and temporal scales that are not possible using traditional experimental approaches. An extreme version of this involves many chemical reactions in the environment, which occur at the molecular scale but which have impacts regionally or even globally, as is shown in much atmospheric chemistry or environmental pollution research. Atomistic simulations are used increasingly to predict changes in behaviour under changed and extreme conditions where experiments are difficult to perform. In environmental problems, such simulations are additionally difficult owing to the need to carry out ensemble predictions, which leads to advances both in the computations themselves and in the methods for handling and visualizing the results. [Salje \*et al.\* \(2009\)](#) describe an atomistic simulation modelling system that has been developed, and applied to a range of problems, including understanding the fate of dioxins in soils and how effective different glasses are in encapsulating radioactive waste for long-term storage. [Martin \*et al.\* \(2009\)](#) also discuss new tools for comparing different chemical reactions, giving the proper provenance of the observations and model results so that results can be put in their proper contexts. They have applied this to problems in atmospheric chemistry, although they identify that the burden of additional recording required at the time experiments are done is a significant barrier to use of the system even if, over the life of the project, there is a saving of time and more effective use of resources.

## 6. Advanced data handling

Many of the papers describe advanced data handling and visualization as well as new scientific results. Several concentrate on data handling and visualization, because the need for systems that are easy to use and attuned to problems in environmental sciences will greatly reduce the barriers to using the new technologies. [Hall \*et al.\* \(2009\)](#) give a perspective on the development of Web technologies, and look forward to further developments. The paper brings together many of the underlying themes of both the Discussion Meeting and this issue, namely that many of the new developments are catalysed by allowing different groups to work together. The authors identify the development of semantic methods as important for organizing information interactively across research fields. This theme is picked up also in [Lawrence \*et al.\* \(2009\)](#), who describe a system to give easy access to heterogeneous environmental information. They describe an information taxonomy to help with this. The challenge will be to migrate this interesting development into the wider environmental community, and allow it to evolve based on how users actually use the system. [Latham \*et al.\* \(2009\)](#) describe some of the developments in more detail, including how the search vocabulary was developed.

Dozier & Frew (2009) discuss the importance of provenance in making hydrological forecasts based on diverse snow data. The system described allows confidence to be put in the diverse types of data that are used, both for research and operationally. Blower *et al.* (2009) describe an interactive visualization and data access website that allows users to explore extremely large datasets without having to install new software or become very proficient in a particular software system. It is based on open international standards, and has been widely adopted internationally already.

eScience allows many more conditions to be explored than traditional experimental approaches, and so the discipline naturally faces a data deluge. Perhaps some of the greatest advances in data handling are associated with designing the simulation system itself and associated toolkits for data and information handling. Tools have been developed to manage large heterogeneous datasets and provide ways to visualize the data (White *et al.* 2009). In conjunction with this, Walker *et al.* (2009) exploited grid computing to enable better management of the vast data and metadata produced by the mineral models. Integral to eScience is collaboration across disciplines and between sites. Frame *et al.* (2009) describe the use of new eScience tools to support virtual collaborations, including using Web 2.0 social networking tools to share information and data and to document ideas and the collaboration process.

## 7. Conclusion

The impact of novel and powerful computing, the connectivity of the Internet and other scientific grids, and innovations such as the Semantic Web, are only now beginning to have a major impact in the environmental sciences. The papers in this issue represent a snapshot of some of the areas where change is happening rapidly. The science described here has already been recognized by several international awards, and several of the projects have had a major impact outside their specialist community. Our impression is that the most novel and innovative advances have come where interdisciplinary collaborations have really been made to work. This will be a challenge to sustain, both due to academic disciplinary pressures and because it is not a natural marriage for many funding agencies. Whichever scientific funding agencies manage this are likely to reap scientific rewards disproportionate to the funds expended as the revolution in environmental science unfolds. We are excited by the potential hinted at in this issue, and certainly look to the future of environmental eScience with great interest, optimism and anticipation.

We would like to thank all the participants at the Discussion Meeting, the speakers, the authors of the papers in this issue and the other contributors, especially Microsoft Research for sponsoring the evening demonstration session at the Discussion Meeting, many Royal Society staff for administrative help, and Michelle Truman for her editorial support.

R.J. Gurney<sup>1</sup>, N.S. Badcock<sup>2,\*</sup>, E.A. Garnett<sup>2</sup> and H.C.J. Godfray<sup>3</sup>

<sup>1</sup>*ESSC, University of Reading, Reading RG6 6AL, UK*

<sup>2</sup>*NERC, Swindon SN2 1EU, UK*

*E-mail address: nsba@nerc.ac.uk*

<sup>3</sup>*Department of Zoology, University of Oxford, Oxford OX1 3PS, UK*

*\*Author for correspondence.*

## References

- Blower, J. D., Haines, K., Santokhee, A. & Liu, C. L. 2009 GODIVA2: interactive visualization of environmental data on the Web. *Phil. Trans. R. Soc. A* **367**, 1035–1039. (doi:10.1098/rsta.2008.0180)
- Bretherton, D. A., Blower, J. D., Haines, K. & Smith, G. C. 2009 Running climate models on grids using G-REX. *Phil. Trans. R. Soc. A* **367**, 847–853. (doi:10.1098/rsta.2008.0179)
- Clark, B. R., Godfray, H. C. J., Kitching, I. J., Mayo, S. J. & Scoble, M. J. 2009 Taxonomy as an eScience. *Phil. Trans. R. Soc. A* **367**, 953–966. (doi:10.1098/rsta.2008.0190)
- Dozier, J. & Frew, J. 2009 Computational provenance in hydrologic science: a snow-mapping example. *Phil. Trans. R. Soc. A* **367**, 1021–1033. (doi:10.1098/rsta.2008.0187)
- Droegemeier, K. K. 2009 Transforming the sensing and numerical prediction of high-impact local weather through dynamic adaptation. *Phil. Trans. R. Soc. A* **367**, 885–904. (doi:10.1098/rsta.2008.0211)
- Frame D. J. *et al.* 2009 The *climateprediction.net* BBC climate change experiment: design of the coupled model ensemble. *Phil. Trans. R. Soc. A* **367**, 855–870. (doi:10.1098/rsta.2008.0240)
- Frame, I., Austen, K. F., Calleja, M., Dove, M. T., White, T. O. H. & Wilson, D. J. 2009 New tools to support collaboration and virtual organizations. *Phil. Trans. R. Soc. A* **367**, 1051–1056. (doi:10.1098/rsta.2008.0193)
- Froude, L. 2009 Storm tracking and eScience. *Phil. Trans. R. Soc. A* **367**, 905–911. (doi:10.1098/rsta.2008.0183)
- Haines, K., Hermanson, L., Liu, C., Putt, D., Sutton, R., Iwi, A. & Smith, D. 2009 Decadal climate prediction (project GCEP). *Phil. Trans. R. Soc. A* **367**, 925–937. (doi:10.1098/rsta.2008.0178)
- Hall, W., De Roure, D. & Shadbolt, N. 2009 The evolution of the Web and implications for eResearch. *Phil. Trans. R. Soc. A* **367**, 991–1001. (doi:10.1098/rsta.2008.0252)
- Hermanson, L. & Sutton, R. T. 2009 Climate predictability in the second year. *Phil. Trans. R. Soc. A* **367**, 913–916. (doi:10.1098/rsta.2008.0181)
- Holt, J. *et al.* 2009 Modelling the global coastal ocean. *Phil. Trans. R. Soc. A* **367**, 939–951. (doi:10.1098/rsta.2008.0210)
- Latham, S. E. *et al.* 2009 The NERC DataGrid services. *Phil. Trans. R. Soc. A* **367**, 1015–1019. (doi:10.1098/rsta.2008.0238)
- Lawrence, B. N., Lowry, R., Miller, P., Snaith, H. & Woolf, A. 2009 Information in environmental data grids. *Phil. Trans. R. Soc. A* **367**, 1003–1014. (doi:10.1098/rsta.2008.0237)
- Lenton, T. M., Myerscough, R. J., Marsh, R., Livina, V. N., Price, A. R., Cox, S. J. & the GENIE team. 2009 Using GENIE to study a tipping point in the climate system. *Phil. Trans. R. Soc. A* **367**, 871–884. (doi:10.1098/rsta.2008.0171)
- Martin, C., Haji, M., Jimack, P., Pilling, M. & Dew, P. 2009 Semantically enhanced provenance capture for chamber model development with a master chemical mechanism. *Phil. Trans. R. Soc. A* **367**, 987–990. (doi:10.1098/rsta.2008.0168)
- Putt, D., Haines, K., Gurney, R. & Liu, C. 2009 Land surface anomaly simulations and predictions with a climate model: an El Niño Southern Oscillation case study. *Phil. Trans. R. Soc. A* **367**, 917–923. (doi:10.1098/rsta.2008.0182)
- Salje, E. *et al.* 2009 eScience for molecular-scale simulations and the eMinerals project. *Phil. Trans. R. Soc. A* **367**, 967–985. (doi:10.1098/rsta.2008.0195)
- Slingo, J., Bates, K., Nikiforakis, N., Piggott, M., Roberts, M., Shaffrey, L., Stevens, I., Vidale, P. L. & Weller, H. 2009 Developing the next-generation climate system models: challenges and achievements. *Phil. Trans. R. Soc. A* **367**, 815–831. (doi:10.1098/rsta.2008.0207)
- Thorpe, A. 2009 Preface. *Phil. Trans. R. Soc. A* **367**, 801–802. (doi:10.1098/rsta.2008.0254)
- Walker, A. M., Bruin, R. P., Dove, M. T., White, T. O. H., van Dam, K. K. & Tyler, R. P. 2009 Integrating computing, data and collaboration grids: the RMCS tool. *Phil. Trans. R. Soc. A* **367**, 1047–1050. (doi:10.1098/rsta.2008.0159)
- Washington, W., Buja, L. & Craig, A. 2009 The computational future for climate and Earth system models: on the path to petaflop and beyond. *Phil. Trans. R. Soc. A* **367**, 833–846. (doi:10.1098/rsta.2008.0219)
- White, T. O. H., Bruin, R. P., Chiang, G.-T., Dove, M. T., Tyler, R. P. & Walker, A. M. 2009 Lessons in scientific data interoperability: XML and the eMinerals project. *Phil. Trans. R. Soc. A* **367**, 1041–1046. (doi:10.1098/rsta.2008.0175)