Reading e-Science Centre Activity Report

K. Haines, J. Blower, D. Bretherton, A. Gemmell

11 February 2009

1. Executive Summary

This report summarizes the key achievements of the Reading e-Science Centre during the current contract period (November 2005 to the present day). The contract of £379k employed two full-time staff and produced 12 peer-reviewed papers, with 3 more in review. The ReSC has developed numerous open-source tools and services, which have been adopted world-wide by researchers and operational environmental data providers. A further £1.7M has been won from grants from research councils, European framework programmes, industry and other funding bodies, enabling the ReSC to employ a further 4 full-time staff in order to build upon its core work. The ReSC has been very active in engaging with many communities, including scientists, government agencies, industry, the general public and the University of Reading itself.

2. Table of Contents

3. Introduction

The Reading e-Science Centre (ReSC, http://www.resc.reading.ac.uk) was formed in November 2003 as an e-Science Centre of Excellence with the task of promoting and developing e-Science methods in the environmental science community. The initial funding of £151k over 2 years was provided jointly by NERC and EPSRC and paid for a Technical Director (Dr Jon Blower) and an Access Grid facility. During this period the ReSC was co-directed by Prof. Keith Haines (ESSC) and Prof. Rachel Harrison (School of Systems Engineering). This initial phase of activity was rated by EPSRC as "Tending to Outstanding" in review by EPSRC.

In 2004 ReSC successfully applied to NERC for a three-year contract extension to the value of £379k, mainly covering the salaries of the Technical Director and a second full-time staff member (Mr Dan Bretherton). This covered the period from 1 November 2005 to 31 October 2008, although a one-year no-cost extension is currently in place: therefore the current ReSC contract expires at the end of October 2009. This Report therefore describes ReSC activities from November 2005 to the present; many activities are still ongoing. Co-director Harrison has now left the University and so the ReSC is currently under the sole directorship of Prof. Haines.

The ReSC's activities complement those of the NIEeS: whereas NIEeS focuses on training, ReSC focuses on research.

4. Outputs

4.1 Publications

This section contains publications that represent both technical work that was funded by the current ReSC contract and scientific work that has been enabled by the tools and services that ReSC provide.

4.1.1. Peer-reviewed journal papers

1. Jon Blower, Keith Haines, Adit Santokhee, Chunlei Liu, **Godiva2: Interactive visualization of environmental data on the web**, *Phil. Trans. Roy. Soc. A*, **367**, 1035-9, 2009
2. Dan Bretherton, Jon Blower, Keith Haines, Greg Smith, **Running Climate Models on Grids using G-Rex**, *Phil. Trans. Roy. Soc. A*, **367**, 847-53, 2009
3. Keith Haines, Leon Hermanson, Chunlei Liu, Debbie Putt, Rowan Sutton, Alan Iwi, Doug Smith, **Decadal climate prediction (project GCEP)**, *Phil. Trans. Roy. Soc. A*, **367**, 925-37, 2009
4. Jason Holt, James Harle, Roger Proctor, Sylvain Michel, Mike Ashworth, Crispian Batstone, Icarus Allen, Robert Holmes, Tim Smyth, Keith Haines, Dan Bretherton, Gregory Smith, **Modelling the global coastal ocean**, *Phil. Trans. Roy. Soc. A*, **367**, 939-51, 2009
5. Keith Haines, Jon Blower, Chunlei Liu, Adit Santokhee, **Delivering NCOF operational marine data through the Internet**, *Journal of Operational Oceanography* 1(2) pp. 35-39, 2008

6. Lizzie Froude, **Storm tracking with remote data and distributed computing**, *Computers and Geosciences*, 34(11) pp. 1621-1630, 2008 (enabled by Reading Campus Grid)
7. Jon Blower, Andrew Harrison, Keith Haines (2006), **Styx Grid Services: Lightweight middleware for efficient scientific workflows**, *Scientific Programming*, 14, pp. 209-216
8. L. Bengtsson, K. I. Hodges, E. Roeckner (2006), **Storm Tracks and Climate Change**, *Journal of Climate* 19, pp. 3518-3543. *(Enabled by Reading Campus Grid)*

4.1.2. Journal papers under review

1. D. Lowe, A. Woolf, B. Lawrence, S. Pascoe, J.D. Blower, **Integrating the Climate Science Modelling Language with Geospatial Software and Services** (submitted to the International Journal of Digital Earth)
2. Thomas Loubrieu, Sylvie Pouliquen, Marc Spigai, Pierre Bahurel, Frédérique Blanc, Sophie Baudel, Giuseppe Manzella, Charlotte O'Kelly, Jon Blower, Chunlei Liu, **A distributed system for operational oceanography: the Mersea System** (submitted to the Journal of Operational Oceanography)
3. Alastair Gemmell, Greg Smith, Keith Haines, Jon Blower, **Validation of ocean model syntheses against hydrography using a new web application** (submitted to the Journal of Operational Oceanography).

4.1.3. Peer-reviewed conference papers

Please note that in the computing community it is common to publish papers through conferences as well as journals.

1. Jon Blower, Frederique Blanc, Peter Cornillon, Steve Hankin, Thomas Loubrieu, **Underpinning technologies for oceanographic data sharing, visualization and analysis: Review and future outlook**, *GODAE Final Symposium, Nice*, 2008, invited presentation
2. Jon Blower, Alastair Gemmell, Keith Haines, Peter Kirsch, Nathan Cunningham, Andrew Fleming, Roy Lowry, **Sharing and visualizing environmental data using Virtual Globes**, *Proceedings of the UK e-Science All Hands Meeting*, 10-13 September 2007, pp 102-109, ISBN 978-0-9553988-3-4 (awarded "Best Paper" at the meeting)
3. Bryan Lawrence, Philip Kershaw, Jon Blower, **Practical access control using NDG security**, *Proceedings of the UK e-Science All Hands Meeting* 10-13 September 2007, pp 262-269, ISBN 978-0-9553988-3-4
4. Jon Blower, Keith Haines, **Building simple, easy-to-use Grids with Styx Grid Services and SSH**, *Proceedings of the IEEE e-Science Conference, Amsterdam, December 2006*
5. Jon Blower, Andrew Harrison, Keith Haines, **Styx Grid Services: Lightweight, easy-to-use middleware for scientific workflows**, *International Conference on Computer Science* 2006
6. Jon Blower, Keith Haines, Ed Llewellin, **Data streaming, workflow and firewall-friendly Grid Services with Styx**, *Proceedings of the UK e-Science All Hands Meeting* 19-22 September 2005

4.1.4. Conference abstracts

1. Jon Blower, Pauline Mak, John Caron, Ethan Davis, Adit Santokhee, **Visualizing Earth Science data using a Web Map Service interface to the THREDDS Data Server**, to be presented at EGU09, Vienna.

2. G. Aloisio, S. Fiore, M. Petitdidier, P. Fox, S. Denvil, J. Blower, H. Schwichtenberg, R. Barbera, **The Climate-G testbed: towards a large scale data sharing environment for climate change**, to be presented at EGU09, Vienna.
3. Alastair Gemmell, Jon Blower, Keith Haines, Martin Price, Keiran Millard, Quillon Harpham, **Integrating Distributed Physical and Biological Marine data using OGC Web Services**, *Eos Trans. AGU*, 89(52), Fall Meet. Suppl., Abstract IN52A-04
4. Alastair Gemmell, Jon Blower, Keith Haines, Greg Smith, **Using Virtual Globes and a Java web Application to Visualize and Compare Ocean Observations and Model Data**, *Eos Trans. AGU*, 88(52), Fall Meet. Suppl., Abstract IN42A-04
5. Gregory C. Smith, Keith Haines, Dan Lea, Jon Blower, Alastair Gemmell, Ben McDonald, **Reconstructing global ocean variability using a physically-based data assimilation method**, *Proceedings of the Royal Meteorological Society Conference*, 3-6 September 2007 (enabled by OceanDIVA software)
6. Jon Blower, Dan Bretherton, Keith Haines, Chunlei Liu, Adit Santokhee (2006), **Exploring Large Marine Datasets Using an Interactive Website and Google Earth**, *Eos Trans. AGU*, 87(52), Fall Meet. Suppl., Abstract ED23C-08
7. Jon Blower, Dan Bretherton, Keith Haines, Chunlei Liu, Chris Rawlings, Adit Santokhee, Ian Smith (2006), **Using Google Earth in Marine Research and Operational Decision Support**, *Eos Trans. AGU*, 87(52), Fall Meet. Suppl., Abstract IN43A-0903

4.1.5. Other publications

1. Alastair Gemmell, Greg Smith, Keith Haines, Jon Blower, **Evaluation of water masses in ocean synthesis products**, *CLIVAR Exchanges* 47 pp. 7-9, 2008 (enabled by OceanDIVA software)

4.2 Products and their exploitation

The ReSC has made available to the community a number of services and software tools. We have focussed on making these products available early, in order to build the community and gather suggestions for improvements.

4.2.1. ncWMS/Godiva2: online data visualization

ncWMS (http://ncwms.sf.net) is an **open-source** implementation of a Web Map Service (WMS), specifically designed to enable fast, interactive visualization of gridded environmental data, normally held in NetCDF files. (The WMS specification is an open standard of the Open Geospatial Consortium (OGC) for the sharing of map imagery over the Internet.) Godiva2 is the companion web interface to ncWMS: either component can be used individually, but it is most common for ncWMS and Godiva2 to be deployed as a combined system (e.g. http://www.reading.ac.uk/godiva2).

Originally developed under the DEWS project (see section 6.1 below) the software has evolved over a number of years as an open source project, and has been very **widely adopted internationally**. In particular, it is used by the major European oceanography projects MERSEA (http://www.resc.rdg.ac.uk/mersea) and ECOOP (http://www.resc.rdg.ac.uk/ecoop), in which it is employed as a dynamic data visualization service that is compliant with the European INSPIRE directive. The software will also be used extensively in the upcoming MyOcean project (GMES Marine Core Service), fulfilling part of the project's obligations under INSPIRE and helping to move web-based visualization from a research activity onto a production-quality platform.

Other known adopters of the software include the NERC National Centre for Earth Observation (http://www.resc.rdg.ac.uk/nceo); Plymouth Marine Laboratory, as part of the INTERRISK project (http://www.npm.ac.uk/rsg/projects/interrisk); the NOAA National Geophysical Data Center; the US

Geological Survey; the US Navy, the INSEA project in the Netherlands (http://www.insea.info/); Deltares in the Netherlands; the Australian Institute of Marine Science (http://data.aims.gov.au/atlas/); the Institute for Marine Research in Norway (http://talos.nodc.no:8080/ncWMS/godiva2.html); Meteogalicia in Spain; ERDAS (http://www.erdas.com/, a US-based GIS company) and the Climate-G Data Distribution Centre in Italy (website not yet public).

It is very likely that there are many more users of which we are not specifically aware: the last release of the software has been **downloaded 570 times** (in source or binary format), as of February 2009.

The ReSC runs a Godiva2 server of its own as a demonstration site (http://www.reading.ac.uk/godiva2), displaying data from ocean forecast models, EO products and data assimilation projects. This serves an average of around **1000 map images per day**, to users on five continents.

ncWMS is now proven community software and is an important part of the ReSC's long-term strategy. A NERC knowledge exchange project (SHAVER), now coming to an end, has successfully integrated ncWMS with the popular THREDDS Data Server from Unidata, which is used throughout the world to share environmental data. The combined system (whose first release is imminent) will provide OPeNDAP, Web Coverage Service and Web Map Service interfaces to huge stores of environmental data, representing a big step forward in the adoption of this technology. Within the ESA-sponsored GENESI-DR project (http://www.genesi-dr.eu/), we are enhancing the capabilities of ncWMS to be able to display satellite swath data, which can then be compared with numerical model data, with many potential applications in data assimilation. Two further proposals to NERC are currently planned to develop and apply ncWMS further in the environmental science community.

4.2.2. G-Rex

ReSC has produced its own Grid solution for running large climate models on remote resources. G-Rex ("Grid Remote Execution") is pure Java grid middleware designed for remote execution of scientific applications and is the successor to JStyx (section 4.2.6 below). G-Rex overcomes two particular limitations in existing grid middleware associated with climate modelling: climate models typically produce very large volumes of data and are normally launched on local compute resources using large, complicated shell scripts. G-Rex is designed to be very lightweight to install and easy to incorporate into scientists' existing complex work-flow scripts, and it tackles the build up of large output on remote resources by transferring output from models back to the user during model runs and deleting files from the remote system as soon as they are no longer needed.

G-Rex is being used for running Proudman Laboratory coastal ocean models and also the NEMO global ocean model on cluster resources at Reading, POL, PML, NOCS and BAS in the framework of the NERC Cluster Grid (see section 4.2.8 below). G-Rex is still under development and is now being adapted to be suitable for use with existing HPC resources.

4.2.3. OceanDIVA

OceanDIVA (Ocean Data Intercomparison and Visualisation Application) is a Java Web Application (http://www.reading.ac.uk/oceandiva) which allows users to compare in-situ ocean temperature and salinity profiles with model output via a web interface. The computed differences between model data and observations can then be visualised either geospatially (in Google Earth) or statistically (probability density functions).

OceanDIVA has been used by marine group research staff at the UK Met Office, as well as here at the University of Reading to aid with model validation and data assimilation. It has also received interest and online usage from the CLIVAR Global Syntheses and Observations Panel (GSOP) as part of their efforts to model the past ocean state. CLIVAR (http://www.clivar.org/) is the World Climate Research Programme (WCRP) project that addresses Climate Variability and Predictability. In this case OceanDIVA has revealed some interesting differences in the way that global ocean syntheses, run at various institutes worldwide, reproduce certain features of interest of ocean water masses.

The ability to easily compare models and observations is a key to many of our projects. In the context of ECOOP (a large EU coastal oceanography project) the ReSC has developed a prototype online decision support system for an ecosystem health application (http://www.resc.reading.ac.uk/ecoop_ecosystem_portal) that allows users to assess the model's performance through comparisons with in situ observations, using a suite of OGC-compliant web services.

4.2.4. Java-CSML

From its research into the visualization of environmental data, the ReSC has recognized a need for greater harmonization of environmental data. A key limitation inhibiting the development of software enabling the sharing and analysis of environmental data is the plethora of file formats and metadata conventions employed by scientists and data providers. The Climate Science Modelling Language (CSML), developed by STFC in the context of the NERC Data Grid project, provides a common data model and representation for a very large proportion of environmental data, including numerical model output, in situ observations and EO products. We believe that greater adoption of CSML will lead to increased use of environmental data and will be particularly valuable for activities (such as data assimilation) that involve the intercomparison of multiple types of data.

In order to increase adoption of CSML, the ReSC is working with STFC to develop a Java implementation, known as Java-CSML (a large proportion of environmental data-serving systems are developed in Java). Early prototypes of this software have been used in the ECOOP project (to compare in situ observations of the ocean with numerical models) and the GENESI-DR project (to compare satellite observations of atmospheric ozone with data-assimilated numerical models). Although this work is currently at an early stage, the software has already been adopted by NOAA, and there is strong interest from Ifremer and other parties. The ReSC's future strategy is to use Java-CSML to combine the capabilities of ncWMS and OceanDIVA to produce a web-based data visualization and intercomparison system that is capable of handling many different types of environmental data. This is the subject of a current proposal to NERC (in preparation).

The CSML data model is harmonized with the Observations and Measurements specification; this means that Java-CSML will be an important component of future data systems that expose environmental data through the new OGC Sensor Web Enablement suite of specifications.

4.2.5. GADS-WCS

Since 2003, the ReSC has developed and maintained the GADS (Grid Access Data Service) software, which provides SOAP Web Service access to gridded environmental data. In the DEWS project, the ReSC re-engineered GADS to be compatible with the OGC's Web Coverage Service (WCS) specification. The new system (known as GADS-WCS) was among the first such systems (if not the very first) to explore the application of WCS in a secure environment for providing access to large (GB to TB) scale datasets. As a result of this research, a number of modifications to the WCS specification were proposed, which are now being researched further in the context of the international "WCSplus" effort, which involves participation from Unidata, the UK Met Office and

other interested parties. Although the ReSC no longer actively maintains GADS-WCS, we continue to contribute to the ongoing discussions in this area, which have the ultimate goal of devising a means to serve large-volume environmental data in an OGC-compliant manner.

4.2.6. Styx Grid Services and jStyx

Following on from earlier research into the use of the Inferno operating system for creating distributed systems (in the previous ReSC grant period), the ReSC created a Java implementation of the Styx protocol (jStyx) for distributed systems (Styx is the communications protocol used by Inferno). The goal of this research was to enable large datasets (e.g. environmental numerical model output) to be streamed efficiently from machine to machine in a Grid workflow, a capability that was not addressed well by most Grid middleware at the time. This software was incorporated into the popular Taverna workflow system (http://taverna.sf.net), a product of the MyGrid e-Science project. This work is now being taken forward as G-Rex (section 4.2.2 above).

4.2.7. Reading Campus Grid

The University of Reading, in common with many other UK research-intensive universities, has established a Campus Grid infrastructure, giving University researchers access to computing resources that are free at the point of use (http://www.reading.ac.uk/internal/its/escience/its-eresearch-campusgrid.asp). At Reading, the Grid takes the form of a Condor pool (linked to the UK National Grid Service), which harnesses the spare computing power of desktop machines in the library and teaching laboratories. Although this has only recently become a production system, it has enabled a number of pieces of environmental research (e.g. Froude 2008, Bengtsson et al 2006). (Many problems of parameter-sweeping, Monte Carlo studies and data analysis are very well-suited to this type of resource.) The ReSC played a key role in the establishment of this facility by driving discussions, testing, and successfully lobbying for central funds to maintain the infrastructure. The Campus Grid is now developed and maintained by the University's IT Services and the School of Systems Engineering.

4.2.8. NERC Cluster Grid

The NERC Cluster Grid (NCG) is a **1600-processor** grid of High Performance Computing (HPC) clusters belonging to the following NERC research institutes: the Environmental Systems Science Centre (ESSC), the British Antarctic Survey, the National Oceanography Centre, the Proudman Oceanographic Laboratory and the Plymouth Marine Laboratory. The ReSC has led the development of the NCG, which uses G-Rex (section 4.2.2 above) for job submission and Ganglia for monitoring. The two climate models that are currently installed are NEMO and POLCOMS, which are exposed as G-Rex Web services (see above) on all the clusters in the Grid. The use of G-Rex for remote execution of climate models is supported by the NERC Cluster Compute Forum. Cluster administrators from the British Geological Survey and the Centre of Ecology and Hydrology have expressed an interest in adding their clusters to the Grid.

In making it easier for collaborating scientists to use each others' clusters, the Grid benefits NERC because it allows idle processors to be used productively, thereby increasing utilisation rates. The Ganglia Web Frontend (http://www.resc.reading.ac.uk/ganglia) can be used to view historical utilisation rates in order to assess the impact of increased remote execution by external users. Science projects can benefit from the NERC Cluster Grid too, because it enables scientists to easily run models that are not installed or too large to run at their own institutes. The main beneficiaries so far have been the GCOMS (POLCOMS and ERSEM models) and MarQUEST (NEMO and NEMO + PlankTOM5 bio. model) projects and the DRAKKAR consortium (NEMO model). Several scientists at ESSC frequently use the Grid to run NEMO on remote clusters, particularly the ¼° model, which is too large to run on ESSC's cluster. Members of the GCOMS project at POL are

planning a large series of POLCOMS-ERSEM model runs, which will utilise idle processor time on a range of different clusters in the Grid. Scientists at the Centre for Polar Observation and Monitoring recently expressed an interest in running NEMO on the Grid due to a lack of suitable resources at their own institute.

4.2.9. Access Grid

The Access Grid is an advanced web-based videoconferencing system. The ReSC has supported an Access Grid node since 2003 and makes this available to the University of Reading for holding meetings with remote partners, reducing the need to travel. Our Access Grid is in regular use for supervision of remote PhD students, project meetings and the remote teaching of MSc courses. Our original choice of using a commercial Access Grid solution (from IOCOM, then known as Insors) has proven to be a wise one, since our commercial system appears to require much less maintenance and resource than self-built systems at other institutions.

4.2.10. ESSC Cluster

The ESSC Cluster is a High Performance Computing cluster containing 64 processors, which is designed for efficient, parallel execution of large models, including HadCM3, NEMO and POLCOMS. It is still the most powerful cluster at Reading University. The cluster's data storage and analysis servers have a total capacity of more than 21 Terabytes, with enough processing power to allow data analysis tools to process the large data sets typically produced by models executing on the cluster. Bought 3 years ago as part of the GCEP project the cluster was used to initiate the NERC cluster Grid. Now scientists at collaborating NERC institutes can easily run the NEMO and POLCOMS models on any of these clusters with little extra effort. The cluster is also used as a Condor pool, where potential enhancements to the Reading Campus Grid (see above) have been tested. The cluster will soon be linked to the Reading Campus Grid, which in turn will allow it to be part of the UK National Grid Service.

4.3 Data collections

On behalf of the Met Office, the ReSC maintains a ca. 3TB repository of ocean model analysis and forecast data, which it makes available to the scientific community through OPeNDAP and Web Map Services. This arrangement has made it possible for the ReSC to become involved in major European oceanography programmes (MERSEA, ECOOP and MyOcean) and allows Met Office data to be used more widely in science. The services that the ReSC has developed to share this data are now being adopted by the Met Office and other European oceanography institutes in high-profile GMES programmes.

The ReSC also exposes a number of research datasets (mainly ocean model output) through its Web Map Service interface in order to share outputs visually with collaborators. Through the Godiva2 portal (section 4.2.1 above), research outputs can be viewed from anywhere in the world. This facility is used routinely in scientific collaboration with international partners.

5. Achievements as measured against original objectives

Our original plan (see proposal) was to encourage the use of e-Science techniques in the environmental science community through the development of a Grid toolkit for environmental scientists, based upon the CDAT library for climate data analysis. Over the course of the research, we decided that our objectives were best achieved through an **open standards approach**, which avoids the problem of becoming confined to a particular tool, and enables engagement with a wider community. By changing our focus from desktop tools to **web-based applications** we have been able to reach a wider audience than was originally envisaged. Our ambitions in the field of **data**

visualization, sharing and analysis have been realized through our work on ncWMS/Godiva2 (section 4.2.1), OceanDIVA (section 4.2.3) and these capabilities will be developed further in our ongoing work on Java-CSML (section 4.2.4). This approach to environmental data handling shares much in common with the NERC Data Grid, with whom we have collaborated closely. In the field of **Grid computing**, we have developed the G-Rex software (section 4.2.2), which makes Grids more user-friendly for climate scientists and modellers, and developed accessible Grid infrastructure, including the NERC Cluster Grid (section 4.2.8) and the Reading Campus Grid (section 4.2.7).

5.1 Awards

"Best paper" 2007 UK e-Science meeting

"Website of the month" OGC newsletter, January 2008

6. Added value

6.1 Grants won

The ReSC has won **over £1.7M** of additional grants to build on the work funded by the ReSC grant itself. Many of these are large international consortium bids. These funds have allowed us to hire four full-time staff (Dr Alastair Gemmell, Dr Chunlei Liu, Dr Lizzie Froude and Mr Adit Santokhee).

Name of project	**Funding body**	**Dates**	**Value to ReSC ***
DEWS (Delivering Environmental Web Services, total value £2.2M)	DTI (now TSB)	July 2005 – July 2007	£90k
SHAVER (Sharing and Visualizing Large-Volume Environmental Data)	NERC	Jan 2008 – Apr 2009	£104k
NERC Data Portals (total value £460k)	NERC	Oct 2007 – Mar 2008	£20k
GCOM	NERC	Feb 2006 - Sept 2009	£95k
GCEP	NERC	Sept 2005 - Apr 2009	£276k
ECOOP	EU	Feb 2007 - Oct 2010	£84k
GENESI-DR (total value 4 million euros)	ESA	Jan 2008 – Dec 2009	38k
MyOcean (total value 55 million euros)	EU	Jan 2009 – Dec 2011	£130k
PSRE	DIUS	Oct 2006 - March 2009	£233k
e-Research South (total value £1.5 million)	EPSRC	Jun 2008 - May 2013	£500k
GUTS (GOCE User Toolbox Specification)	ESA	Jan 2006 – Jan 2007	£40k
2-year PDRA	Schlumberger	Jan 2007 – Dec 2008	£144k

* Some values are approximate due to exchange rate fluctuations.

6.2 Community engagement

In addition to interactions within the ReSC's portfolio of projects, the ReSC has been very active in engaging with a number of communities in academia, government agencies and industry. This has been done mainly through organizing and participating in workshops, often giving invited presentations.

6.2.1. Engagement with the environmental science community

- Jointly organized a workshop with NIEeS on the use of geobrowsers in environmental science (http://www.niees.ac.uk/events/GoogleEarth/index.shtml, 2-3 April 2007). This attracted 80 delegates, including representatives from industrial software providers (ESRI, Google), government agencies (NASA) and international universities (University of Alaska).
- QUEST: The ReSC participated in two workshops on the use of web-based visualization systems for the QUEST programme (July 2007 and July 2008), sharing its expertise in interactive visualization of environmental data.
- NERC Portals project: The ReSC participated in a NERC Knowledge Exchange project that developed a new web portal interface to the NERC Data Grid infrastructure (http://www.edp.nerc.ac.uk/), in collaboration with BODC, BADC, BAS, CEH, and Defra's Central Science Laboratory (CSL). The lessons learned from this project are providing input to NERC's ongoing information strategy.
- Dr Blower gave an invited presentation at an Oceans2025 workshop at POL in October 2008, discussing different approaches to developing data visualization websites.
- Dr Blower is currently on the Advisory Board for the ongoing NERC Data Grid Medium Sized Initiative project.
- Although the ReSC is not explicitly funded under the National Centre for Earth Observation, it is providing an online data visualization service (http://behemoth.nerc-essc.ac.uk/ncWMS/nceo.html). This will encourage the NCEO partners to expose their data and make it more readily available.
- See section 4.2.8 above for details on interactions with the NERC High Performance Computing community.
- Dr Blower will give an invited presentation at the upcoming NERC Data Management workshop (16-17 February 2009) on the use of virtual globes and web-based visualization tools in environmental science.
- Dr Gemmell will give a presentation on "An open approach to ocean data visualization" at a meeting of the Royal Meteorological Society in London on the 18th of February.

6.2.2. Engagement with government agencies

The ReSC enjoys a close relationship with the UK Met Office, with whom it has collaborated on many projects (including DEWS, MERSEA, ECOOP, MyOcean and a PSRE project). These projects have led to collaborations with other European operational met-ocean forecasting and research institutes, notably Ifremer in France.

Other specific interactions with government agencies include:

- Dr Blower gave an invited presentation on virtual globes at Defra in the "public engagement" session of their EO Week, November 2007.

- Dr Blower gave a presentation on the ncWMS/Godiva2 system at workshop at ECMWF in November 2008. The aim of the workshop was to explore the use of OGC standards in meteorology, and the ReSC is preparing a proposal to NERC to allow it to address some of the issues that were raised.

6.2.3. Engagement with industry

The ReSC has ongoing collaborations with BMT who partly sponsor the ReSC Director Prof. Haines. The recent acquisition of ARGOSS (www.argoss.nl) to the BMT group has led to new collaboration opportunities in areas such as atmospheric and ocean forecasting and in use of ESA satellite data for services development. We are actively exploring the potential for collaboration here and have jointly submitted an expression of interest for an ESA tender.

Dr Lizzie Froude was a NERC e-Science PhD student within ReSC using Grid computing to study storm tracks. She graduated in 2007 and for the last 2 years has been sponsored by Schlumberger.

The DTI-sponsored DEWS project (section 6.1 above) involved BMT, IBM and the software company Lost Wax (www.lostwax.com). ReSC and Lost Wax have continued collaborative links in the NERC Data Portals project.

ReSC's work in open geospatial data standards has led to engagement with providers of GIS software, including ESRI, Google and Microsoft, through many workshops including the ReSC/NIeES-organized Geobrowsers workshop, Defra's EO week (see above) and other events at the University of Nottingham and ECMWF. There is continuous interaction with these partners through OGC technical committee discussions.

Dr Blower gave an invited presentation to a Knowledge Transfer workshop of the British Association of Remote Sensing Companies (BARSC) in May 2007 at the DTI, London.

Dr Gemmell and colleagues from the Met Office involved in the PSRE project gave a demonstration of OceanDIVA at Oceanology International – a major marine technology trade fair – in London in March 2008.

6.2.4. International engagement

As described in the sections above, many of ReSC's activities involve international collaborations and engagement. Particular highlights include:

- The use of ReSC's ncWMS software in European operational oceanography programmes and other institutes in Europe, the US and Australia.
- Dr Blower gave an invited presentation and paper to the Final Symposium of the Global Ocean Data Assimilation Experiment (GODAE) in Nice in November 2008.
- As a result of this, Dr Blower has been invited to present a report on new data-sharing technology to the Global Ocean Observing System (GOOS) Scientific Steering Committee in Perth, Australia in February 2009.
- In the context of the SHAVER Knowledge Exchange project, the ReSC is working with Unidata in the US and the University of Tasmania to integrate the ncWMS software with the THREDDS Data Server.
- The ReSC will participate in the Climate Change Integration Plugfest at the FOSS4G conference in Sydney, Australia in October 2009. This will demonstrate how geospatial data standards can be used to integrate a variety of data to reveal new information. The audience for this activity will be very wide.

- Mr. Bretherton gave a talk about the NERC Cluster Grid at the 3rd Grid and e-Collaboration Workshop, which was hosted by ESA at ESRIN in Frascati. The workshop was attended by members of the European e-research community from industry and academia.
- Mr. Bretherton and Prof. Haines contributed to the GOCE User Toolbox Specification project (GUTS), an ESA funded collaboration involving several European countries. ReSC was responsible for the toolbox architectural design work package, and contributions were also made to the development of algorithms for processing GOCE (Gravity field and steady-state Ocean Circulation Explorer) satellite data.

6.2.5. Engagement within the University of Reading

The original ReSC grant (2003-05) gave the ReSC an explicit role to promote e-Science within the University of Reading. Since the start of the present NERC contract in November 2005, the ReSC has had no official resource to work within the University; nevertheless we have continued this engagement in order to encourage the University to provide central investment into e-Science/e-Research.

The ReSC formed part of a University "e-Research advisory group" who successfully lobbied the University to provide central funds to employ an e-Research officer (Dr David Spence) in IT Services, funded and managed jointly with Oxford University and enabling greater interaction between the two universities. The University is currently considering an additional e-Research support appointment.

The ReSC is currently developing new collaborations across the university under the EPSRC sponsored e-Research South consortium. Examples include;

- Grid enabling Walker Institute hydrological models to enable much larger ensembles for water resource management and prediction
- ReSC has been testing the new ThamesBlue IBM supercomputer that is being operated by the ACET Group in the School of Systems Engineering. The performance of the NEMO model on the IBM architecture has been evaluated by comparing against smaller clusters in the NERC Cluster Grid.

6.2.6. Public engagement

The ReSC's work has been exposed to the public through a number of publications and press releases. See http://www.resc.reading.ac.uk/publicity.php for more details and links.

- The ReSC's Godiva2 website was featured as the "Website of the Month" in the January 2008 edition of the Open Geospatial Consortium's newsletter.
- The ReSC contributed a number of articles to the "glossy brochure" published to present the key outputs of the NERC e-Science programme.
- ReSC work on Virtual Globes featured in the Summer 2007 edition of NERC's "Planet Earth" magazine ("Environmental scientists jump on Google Earth").
- NewScientistTech published an article entitled "Virtual Earths let researchers 'mash up' data" in May 2007, featuring ReSC demonstrations of the use of Google Earth in combining environmental data.
- A press release entitled "e-Science makes weather forecasts available for search and rescue" resulted in articles in CCNews, NASA Earth Observatory News and FS World, in 2005.

- Dr Blower gave an invited presentation on virtual globes to the Public Engagement Seminar during Defra's EO Week in November 2007.
- Prof Haines gave an invited presentation at the London Technology Network (www.ltnetwork.org) meeting in Feb 2008 on Climate change, presenting GCEP work

6.3 Education

In addition to its role in organizing and presenting at a number of workshops to a variety of audiences (see sections above), the ReSC has engaged in educating students and other members of the scientific community. Dr Blower has served on the advisory committee of three PhD students and is the internal examiner for two. Dr Blower has also co-supervised a number of Masters students in collaboration with the University's Computer Science department.

The ReSC also publishes educational material on its web and Wiki site, including a number of tutorials. Tutorials on deploying simple web services and the use of the Inferno operating system are amongst the most popular pages on our site, receiving several hits per day.

6.4 Enhanced efficiency

A key aim of e-Science is to increase the efficiency of scientific research. The ncWMS/Godiva2 (section 4.2.1) and OceanDIVA (section 4.2.3) systems both give very fast and easy-to-use methods for visualizing and intercomparing different datasets. These systems make the diagnosis of numerical models much more efficient, by providing quick visual assessments of the model's output and its performance against observations. The G-Rex system (section 4.2.2) and the NERC Cluster Grid infrastructure (section 4.2.8) greatly simplify the task of running complex models on remote clusters, making it easier for scientists to access powerful resources, and increasing the usage of expensive cluster systems that might otherwise lie idle. The use of Access Grid technology to hold meetings reduces the need for travel, saving time, money and reducing carbon emissions.

7. Structure and management model

The key ReSC personnel are the Director Keith Haines, the Technical Director Jon Blower, Web Services developer Alastair Gemmell and Grid Project Manager Dan Bretherton. ReSC has a flat management model. Most projects are undertaken with in house scientists working on other NERC or EU projects in oceanography and Earth Observation science who then play a key role in testing and evaluating the software that is developed by ReSC. This has been very effective and has led to a number of joint publications involving both environmental and computer science authorship.

The Centre has been effective at Knowledge Transfer, in a number of areas. It has won specific NERC KT grants in the past and is planning new applications in this area. It has also worked closely with the Met Office in supporting EU programs in Operational oceanography.

The Centre will be sustainable at least for the next 4 years since winning an EPSRC e-Research Platform grant in 2008 in Partnership with Oxford and Southampton e-Science Centres. Through this funding we aim to play a more regional role in encouraging e-Research which will go beyond the NERC remit.

The ReSC is also planning to get involved with the NERC Technology Theme through a Proof of concept project to provide seamless access to Model, Satellite and in situ observational data within a service that allows direct and easy comparisons between these data within a GIS framework

The Technical Director of ReSC now has a permanent post at Reading University and is taking the role of PI on most new projects run by ReSC. Jon is also on the NERC peer review college to provide e-Science expertise.

8. Conclusions and recommendations

The ReSC's work has focussed on interacting closely with the environmental science community to develop and apply new e-Science tools. It has successfully applied its research into data visualization and Grid computing and its outputs have been adopted world-wide. It has been very successful in winning new funds and has engaged closely with a number of communities outside academia, including government agencies, industry and standards bodies. In the long term the ReSC will continue to develop its capabilities in the area of data visualization and intercomparison, in close collaboration with NERC's emerging data infrastructure resulting from the NERC Data Grid project. Through the high-profile "e-Research South" consortium the ReSC will continue to encourage the uptake of e-Science methods in a number of new communities. The ReSC's key outputs will be sustained in the long term through further grants, building upon the communities that have been established. Through close engagement with the University of Reading the ReSC has helped to ensure that a sustainable infrastructure is available to support local e-Research.

Our experience has shown that software tools and e-infrastructures (both for data and computing) can be extremely valuable to scientists. E-infrastructures are commonly maintained through JISC or through central University initiatives, but there is currently **no mechanism for long-term support for software tools**. Many of the software outputs of the NERC e-Science programme could continue to be useful for many years, particularly if they are maintained as open-source projects (we have found that the open-source model is very useful in helping to build a community that can contribute to the development effort). In addition, almost all environmental scientists use computers daily and a very large proportion of these will, at some time or other, create software tools for their own use or for their colleagues. We believe that a mechanism is needed to ensure the longevity of these tools (where appropriate) and avoid duplication of effort. This could be achieved in many ways:

- By educating the community about good practice in software development.
- By encouraging the community to release popular pieces of software as open-source projects and providing the necessary support.
- By educating the Peer Review College and reviewers about the importance of software and e-Science. (We are pleased to note that NERC has already taken steps in this direction, although time will no doubt be required for this to percolate through the system.)
- By publicising available funding streams for gaining support for software (e.g. the NERC Technology theme and the TSB).
- By recognizing that publication in the scientific literature is not the only mechanism for publicising new technology and encouraging its adoption. (The review cycle in a typical journal is too slow to be useful for disseminating fast-moving technology.) Researchers should be rewarded equally for producing other kinds of output. This requires quite a large culture change in the scientific community.
- By establishing a mechanism for the community to exchange its experiences and its tools: for example, one might envisage a social networking site such as MyExperiment that allows scientists to collaborate easily on writing tools for the common good.
- By recognizing the importance of professional software engineers in the scientific process. For example, the DEWS, NERC Data Portals and NERC Data Grid project have all benefitted from the employment of professional software developers, who can produce very good quality software, leaving the scientists to concentrate on the application of the tools.

We believe that continued interaction with the academic computer science community is valuable, particularly if good people with strong technical skills can be employed to use their talents for the benefit of environmental science (this is why the ReSC has become involved in MSc Computer Science programmes at Reading). However it must be recognized that a computer science researcher's skills and objectives are different from those of a professional software engineer. Computer scientists will not be satisfied simply developing tools on the instruction of scientists; they wish to perform research of their own, and it is very difficult to perform science using "bleeding-edge" technology. This is largely a social problem, revealing the importance of a greater understanding on both sides of the relative roles of science and technology.

We conclude that computing technology will become an increasingly-important part of NERC research, in order to ensure that its data holdings are used to their best advantage and that suitable infrastructures are available to process the ever-increasing volumes of data required by modern research. We encourage NERC to consider how to establish and maintain the necessary tools and services, in collaboration with technical experts in academia and beyond; this will require cultural change in addition to technological innovation.