

# NERC EARTH OBSERVATION DATA CENTRE (NEODC) ANNUAL REPORT 2004/05

## **Contents**

1	OVI	ERVIEW	2
	1.1	NEODC Rationale	2
	1.2	Highlights of the Year	3
	1.2.	(A)ATSR Seamless Archive	3
	1.2.2	2 Appointment of CEDAR Programme Director	3
1.2.3 1.2.4		NEXTMap Britain	3
		EO Spring School	3
	1.2.	Centres of Excellence/New Staff Member	3
	1.2.0	Support of NERC Science Customers	3
	1.2.	BADC / NEODC Relationship	4
	1.3	Progress on Deliverables / Key Milestones	4
	1.3.	Deliverables	4
	1.3.2	2 Key Milestones	6
2	SCIEN	TIFIC AND TECHNICAL OUTCOMES	8
	1.4	Strategic Goals	8
	1.5	User Support, Operations, Science Support and Research	8
	1.5.	Update on User Support	8
	1.5.2	2 Update on Datasets	9
	1.5.3	Update on System Infrastructure	10
	1.5.4	Update on Services	10
	1.5.	Update on Liaison and Publicity	10
	1.5.0	Update on Collaboration	11
	1.5.	7 Other relevant information	12
2	FIN	ANCE OVERVIEW	13
	2.1	Expenditure for 2003/04.	13
	2.2	Budget for 2005/06	13
3	FUT	URE DEVELOPMENT & STRATEGIC FORWARD LOOK	15
	3.1	Update on Strategic Goals	15
	3.1.	Collaboration with NERC EO Centres of Excellence	15
	3.1.2 Mul	2 DMAG Project: Earth System Science: Enhanced Delivery and Visibility tiple Earth Observation Datasets	
	3.1.3	Future Funding Opportunities	16

	3.1.4	4 Opportunities for Improvement	16
4	App	endices	18
	4.1	Appendix 1 Projects supported	18
	4.2	Appendix 2 NEODC position paper 'Data Management for the EO Community'	19
	4.3	Appendix 3 NEODC Website access statistics for FY 2004-5	23

Authors: Dr Victoria Jay/Dr Matt Pritchard

Date: 21/04/2005

#### 1 OVERVIEW

#### 1.1 NEODC Rationale

NERC has acquired and generated a wealth of high-quality Earth observation (EO) data over the past three decades. The value of this rich and diverse time series of environmental observations is constantly increasing, as we seek to understand our dynamic and changing environment.

The role of the NEODC is, in line with NERC Data Policy, to ensure the responsible stewardship and distribution of this valuable resource, to give guidance on the availability and use of EO data and to coordinate the acquisition of new data resources. Such services are to be carried out in an efficient and cost-effective manner in response to requests from its customers in the environmental science community.

The NEODC is becoming increasingly proactive in its collaborations with relevant EO-related initiatives, both in the UK and worldwide, and strives to improve the service it can offer in connecting environmental scientists with the resources they need to carry out NERC's scientific objectives.

The NEODC Mission is:

"To deliver effective services to the NERC community in locating, accessing, interpreting and exploiting Earth Observation data and information, and to ensure the long-term integrity of EO datasets produced and acquired by NERC projects and programmes."

In order to achieve its mission, the NERC Earth Observation Data Centre:

- maintains a central archive and catalogue of NERC commercial satellite data and NERC airborne remotely-sensed data, accessible through the NEODC website at <a href="www.neodc.rl.ac.uk">www.neodc.rl.ac.uk</a>
- provides access to this data for NERC Centres and Surveys, the UK HEI community, NERC Directed Mode Programmes, NERC EO Centres of Excellence, and NERC-funded academics in accordance with the terms of the NERC Data Policy
- co-ordinates and supervises the archiving of all digital data and ancillary information relating to the annual flying campaigns of the NERC Airborne Research and Survey Facility
- maintains an informative web site containing a wide range of EO-related information resources
- works to ensure the professional curation, and ease of access for registered customers, to all EO data held on the NEODC archive
- has overseeing responsibility for NERC EO data held elsewhere within the community, ensuring it is managed and where possible made available to other NERC users
- continues to develop its infrastructure to improve the quality and scope of its data services to the scientific community
- acts as a contact and liaison point for communications on other national and international archiving/cataloguing initiatives relating to EO data
- provides policy and strategy input to NERC corporate data policy through the NERC Data Management Advisory Group.

A consultation is currently underway to define the roles that NEODC should perform to best meet the needs of the NERC EO community. Feedback will be collected during April 05 and it is anticipated that the NEODC mission will be reviewed with NERC during 2005.

## 1.2 Highlights of the Year

## 1.2.1 (A)ATSR Seamless Archive

The major new development for the NEODC is the creation of a seamless archive of products from the (A)ATSR series of instruments (i.e. ATSR-1, ATSR-2 and AATSR). This will be a hugely important dataset, bringing together a unique time-series of observations which are invaluable to current scientific challenges. Demand for this dataset is likely to be high, and will significantly raise the profile of NEODC. NERC benefits greatly from NEODC's involvement in this project, since it secures access to this valuable dataset for all NERC users, at minimal cost to NERC.

## 1.2.2 Appointment of CEDAR Programme Director

Dr Bryan Lawrence was recently appointed to the position of Director, CCLRC Environmental Data Archival and Associated Research Programme (CEDAR), responsible for the development and evolution of the British Atmospheric Data Centre (BADC) and the NEODC. This new programme represents the amalgamation of the data curation activities of BADC and NEODC, the commitment to maintain and improve the relationship of the data centres with NERC and environmental science communities, and to make use of new technologies to further the scientific exploitation of the data archives.

## 1.2.3 NEXTMap Britain

The NEXTMap Britain digital terrain model by the British Geological Survey was purchased by NERC last year and is now available through NEODC, via the same "data browser" interface as used by BADC. The dataset has proved very popular and many new users have been registered as NEODC customers to enable them to access the NEXTMap data.

## 1.2.4 EO Spring School

The NEODC provided, for the 4<sup>th</sup> year in succession, an informative set of web pages, an efficient online application form system, and a post-event online resource for the NERC-sponsored Oxford University/RAL Spring School in Quantitative Earth Observation. This two-week course provides new environmental science researchers with a solid foundation in the techniques of quantitative EO, and is highly successful in encouraging both learning and interaction among the new generation of EO scientists.

#### 1.2.5 Centres of Excellence/New Staff Member

A new staff member, Dr Victoria Jay, was appointed in December 2004 and joined the NEODC from the RAL Remote Sensing Group. Victoria's role includes interaction with NERC Centres of Excellence, which so far is proving to be a much-needed and effective way to raise the profile of NEODC in the EO community, increase awareness of its activities, and gain feedback from the Centres on what priorities NEODC should have in the future.

## 1.2.6 Support of NERC Science Customers

The list of research project titles in Appendix 1 gives an example of the wide range of high-quality science that has been supported by the services and data provided by the NEODC over the past year. Note that, in many cases, particularly when information or advice has been provided to customers without the provision of data, no information was forthcoming from the customer regarding their source of funding. This is however indicative that the NEODC is reaching a wide section of the science community and is supporting environmental science both in the UK and internationally, which is of credit to NERC as an organisation.

## 1.2.7 BADC / NEODC Relationship

A significant number of functions of the British Atmospheric Data Centre (BADC) and the NEODC have been successfully merged. Regular, joint management meetings are held and progress has been made on sharing infrastructure. The data distribution systems have been integrated to make the ATSR UBT and NEXTMap datasets available and other datasets will be handled in this way in the future. Work will continue, to identify common functions of both data centres where effort can be combined, for example in areas such as project management, data storage and management of user databases. As such, the same group at RAL will be able to provide both Data Centres as services to NERC, but retain their individual identities and accountability to their respective communities and funding lines within NERC.

## 1.3 Progress on Deliverables / Key Milestones

#### 1.3.1 Deliverables

The following comments relate to the deliverables defined in the NEODC Technical Annex for 2004/05 which forms part of the NERC/CCLRC SLA for that year.

## 1.3.1.1 Delivery of EO Data Services

Successful delivery of data and information to customers is seen as a key part of the NEODC's service and this was achieved successfully, with no problems reported. Customer enquiries and data requests are now handled within the "Footprints" helpdesk system operated jointly with BADC and making use of a shared user database.

This year was busier than last year in terms of customer enquiries (approximately twice as many user enquiries logged). In particular, the NEXTMap dataset generated significant interest, with 603 separate user-tile-product downloads. Users also registered for ATSR data and are now able to download UBT products. The demand for these is likely to increase further as more data are added to the archive and with increased publicity. As NEODC merges with the BADC user registration system, more quantitative data access statistics will be available for next year.

#### 1.3.1.2 NEODC Website

Work on updating information pages, links and news item is ongoing, and further improvements to the website content and structure have been planned and implementation is underway. These improvements should make the pages more user-friendly and more focused on information which is useful to the UK EO community. Comments and requests (e.g. for links to other relevant datasets and web sites) from the community will be taken into account.

Some recent highlights of this activity are:

- New data pages for NEXTMap and ATSR were added.
- The BADC "data browser" is now being used successfully for delivering SHAC2000, NEXTMap and ATSR UBT data, with further datasets to be added to this system in due course.
- The website for the Oxford / RAL Spring School in Quantitative Earth Observation was updated with information as required both pre-and post-event, and was a successful supporting resource for this year's course.

#### 1.3.1.3 Physical Storage

As the user demands for new data have risen the NEODC has continued to increase its amount of online storage capacity. Jointly with BADC, two large new systems were purchased, providing over

6 TB of online storage space. This has mainly been used for the ATSR UBT archive. Other data assets continue to be held on the Atlas data store at RAL which provides secure back-up, and which is complemented by online storage at NEODC, enabling rapid access to data as it is requested.

## 1.3.1.4 Secure DSRS Secondary Archive

Provision of secure storage for the secondary DSRS AVHRR and SeaWiFs data archive has continued. This project, which makes small demands on the NEODC staff resources, continues to schedule. The matter of provision of additional fire-safe space for future deliveries of DSRS backup data was discussed both at the DSRS steering committee and with DMAG. A preliminary conclusion was that the ongoing need for the fire-safe backup (in addition to backups already made elsewhere) was not clear, and that this may not be continued once the existing safes become full. A final decision is still pending.

## 1.3.1.5 Search Interface / System Infrastructure

The search/browse interface for the remainder of NEODC data products was maintained on a "best effort" basis during this period, but with no major developments, although the existing system is still operating well to provide access to our commercial satellite and airborne data collections.

NEODC and BADC have merged many of their management and infrastructure functions, and now share a common user database. NEODC is serving 3 datasets (NEXTMap, ATSR and SHAC2000, with more to follow) via infrastructure shared with the BADC (Data Browser and FTP server), but "branded" as NEODC systems (i.e. the one shop with two doors approach).

A project (funded by DMAG) to archive and make available the NEXTMap dataset has now been completed, which included the development of an interactive map search for NEXTMap tiles.

#### **1.3.1.6 ATSR-1/2 UBT Archive**

The ATSR UBT products are now available via the NEODC website. User registration, dataset registration, search and browse facilities for this dataset, and a fact-sheet are provided. Data for 1995, 1996, 1997 and 2000 are available both via the new data browser and an FTP server shared with BADC. A detailed metadata catalogue provides the backend to a web-based search interface which currently holds all metadata for 1996 ATSR-2 products. Data for 1998, 1999 and 2001\* will follow shortly. (\*Note 2001 UBT data available to NEODC only covers period up to ERS-2 yaw-steering failure, i.e. mid-January 2001).

ATSR-2 UBTs have now been processed to produce preview images and corrections to product corner coordinates. As agreed between NERC, RAL and the Centre, there has been a hiatus (from November 2004 – March 2005) in ATSR-1 UBT production due to Hadley Centre requirements for ABT products (which cannot be produced in parallel with UBT products). ATSR-1 UBT production has now resumed and will be completed by August 2005. These data will be added to the online collection alongside ATSR-2 UBT products as they are produced. Preview image processing on ATSR-1 UBT products will follow.

#### 1.3.1.7 Interaction with NERC EO Centres of Excellence

The scoping study to determine the data curation and facilitation needs of the NERC EO Centres of Excellence was initiated in December 2004. Initially, data survey questionnaires were devised and sent out to all researchers in the six Centres (CTCD, COMET, CPOM, CASIX, CLASSIC, ESSC<sup>1</sup>). The survey aimed to determine the Centres' needs in terms of data support (provision of third-party

<sup>&</sup>lt;sup>1</sup> DARC data management has been addressed by BADC previously

data sets or other services) as well as issues related to the data generated by the projects. The main purpose is to consider data with long term importance and/or use to the wider scientific community.

A series of visits were undertaken to discuss data management issues with data managers, where Centres had appointed them, plus Directors and key researchers/data users and producers:

5 <sup>th</sup> January	CASIX data manager, SOC
7 <sup>th</sup> January	CTCD data manager, director and researchers, UCL
2 <sup>nd</sup> February	CTCD director, Sheffield
10 <sup>th</sup> February	CPOM director, UCL
23 <sup>rd</sup> February	CLASSIC director and researchers, Swansea
4 <sup>th</sup> March	CLASSIC researchers, CEH Wallingford
9 <sup>th</sup> March	ESSC director, Reading
10 <sup>th</sup> March	CPOM data manager, UCL
31st March	CASIX director, data manager and researchers, PML
19 <sup>th</sup> April	COMET director, Oxford

Data management plans were drafted incorporating the inputs received, and discussions on these are underway with the Centres.

The visits have also been extremely effective in raising the profile of NEODC and increasing awareness and visibility of NEODC, its role and its datasets. As a direct consequence of these contacts we have received several enquiries for data and more general EO-related information, and have been asked to become involved in hosting a number of new datasets which are considered important to the EO community (e.g. Swansea's globally processed AVHRR NDVI and biophysical parameter data).

A joint Centres meeting was arranged on 23 March in Reading, including representatives from ESSC, CLASSIC, CTCD, DARC, BADC and NEODC to discuss data issues common to the wider community. As a result of this meeting, a document has been drafted which intends to outline the present and future position of NEODC. This document has been sent out to the NERC EO community for feedback, and is attached (Appendix 2).

## 1.3.2 Key Milestones

The Key Milestones were defined in the NEODC Technical Annex which forms part of the NERC/CLRC SLA. Not all milestones were met this year, due to the unavoidable delay in appointment in additional staff, and the time spent on QMASS (by agreement<sup>2</sup>). The relevant milestone payments have been deferred to 2005/2006.

## 1.3.2.1 Quarterly Reports

Quarterly reports (covering NEODC Core and NEODC DMAG Enhancement) were delivered on schedule for the quarterly NERC SLA meetings.

<sup>&</sup>lt;sup>2</sup> This reprofiling and rescheduling was agreed because the work undertaken was judged to be strongly aligned with NERC objectives, high priority, of direct benefit to the NEODC Core and DMAG Enhancement work programmes and at no cost to NERC.

## 1.3.2.2 Annual Report 2003/4

The NEODC annual report (covering NEODC Core and NEODC DMAG Enhancement) for 2003/4 was completed and delivered.

**Note**: Due a failed recruitment and consequent period of understaffing 'NEODC Core' did not spend at the anticipated rate. Part of milestone payments Q3 and Q4 were deferred to the next financial year to allow the work programme to be carried forward with increased effort in 2005/6.

#### 1.3.2.3 Collaboration with EO Centres of Excellence

Initial contact was made with Centres of Excellence (CoEs) to discuss development of data management plans. Initial visits were made to the CoEs and draft data management plans have been produced. Completed data management plans are not yet available because discussions with the Centres are still underway (milestone Q4 has been deferred).

# 1.3.2.4 Enhanced Delivery and Visibility for Multiple Earth Observation Datasets (DMAG Enhancement)

- Web services for enhanced data delivery
  - Status: Rescheduled.
    - As reported earlier this year, it was decided to reprofile development activity on this component to take full advantage of the technology spin-offs from NEODC's involvement with the ESA Service Support Environment (via the "QMASS" collaborative project). Service development has now been reorganised to make use of the framework provided by ESA. This not only provides a supported environment in which to publish services, but, through our successful involvement with this project, has raised the profile of NEODC and improved our relationship with ESA and its ground segment and EO service activities. Furthermore, from NEODC's internal links with the NERC DataGrid (NDG) team, interest has already been shown by NDG regarding exploiting some of the architecture and interface developments involved.
    - Following the highly successful involvement with the ESA Service Support Environment, infrastructure technologies from this project will be re-used and work is now underway to design and implement 2 services (for ATSR->AATSR conversion, and NERC ARSF data geocorrection) to enhance access to these datasets.
- Improved metadata system.
  - Status: Work to extend the NERC DataGrid Metadata Model to handle EO metadata has now started, to be followed by the production of a generic metadata creation tool which can be adapted/sub-classed for each type of data to be handled.
- Groundwork for new datasets
  - Status: Work to begin in FY 2005/6
- Enhancement of the ATSR UBT archive
  - ATSR UBT data archive now operational and available via NEODC website
  - ATSR-2 preview image creation: complete
  - ATSR-1 preview image creation: now resumed as of April 2005. RAL was requested (after agreement between NERC and the Hadley Centre) to delay UBT production in order to allow the production of ABT products for Met Office/Hadley Centre work towards the UK contribution to the IPCC. ATSR-1 UBT production is now on track for completion in August 2005.

(Milestone payments Q3 and Q4 have been deferred)

## 1.3.2.5 NEXTMap Implementation

The project (funded by DMAG) to archive and make available for NERC users the NEXTMap digital terrain model dataset was completed successfully. Data are available to registered users according to the conditions of the End User License Agreement and metadata for the dataset are available both via the NEODC website and the NERC Metadata Gateway. All milestones satisfactorily completed.

## 2 SCIENTIFIC AND TECHNICAL OUTCOMES

## 1.4 Strategic Goals

A consultation is currently underway to define the roles that NEODC should perform to best meet the needs of the NERC EO community. Feedback will be collected during April 05 and it is anticipated that the NEODC strategic goals will be reviewed during 2005.

The NEODC primary strategic goals for the period 2002-2005 were defined as:

- **1.** Maintain and improve the acknowledged valuable services which the NEODC has previously delivered to the NERC scientific community
- 2. Provide enhanced data, metadata and information services through the NEODC website
- **3.** Develop automated search and retrieval systems to further improve the quality and timeliness of the NEODC EO data services
- **4.** Develop and implement the professional curation of the ATSR-1/2 archives and the future delivery of associated data product services for which there is increasing demand from the NERC community.

Work has continued to meet the strategic goals as outlined above, and with the full complement of staff resource now available (since December 2004), further progress can be made towards increasing NEODC's engagement with the EO community and improving the services it offers to a growing base of customers.

A new priority objective has been for the NEODC to collaborate with the NERC EO Centres of Excellence by supporting their requirements for data acquisition, curation, access and delivery as appropriate in meeting the wider needs of the NERC scientific community. Victoria Jay has taken on this activity with great enthusiasm.

In addition, following the direction by EOEG and the NERC Data Management that the management and infrastructure functions of the NEODC and BADC at RAL should be merged, significant progress has been made towards this objective. Much generic functionality is now shared, and plans are in place for even closer cooperation, while the two data centres continue to retain their own individual identity to their own community.

## 1.5 User Support, Operations, Science Support and Research

## 1.5.1 Update on User Support

The NEODC provided data and information services across the range of the environmental science disciplines:

Science: Atmos. Earth Marine Terr.& Fresh. Earth Obs. Sci. Bas. Arch. Polar No. of Projects: 0 29 4 2 9 4 0

The majority of individual requests for data during this financial year were satisfied from the existing NEODC data archives. Most of these requests related to the NEODC satellite data archive.

It is noted that the NEODC is beginning to field an increasing number of enquiries regarding general data availability, i.e. not restricted to its own data archives, and is increasingly acting as a source of information about EO and remote sensing in general.

No enquiries for data in support of commercial applications could be satisfied by the NEODC because either we lacked the specified datasets or for reasons of copyright restrictions.

No significant data quality problems were recorded during the year and no formal/informal complaints regarding the quality and delivery of the NEODC services during 2004/05 have been received. In fact, there has often been correspondence complimenting the NEODC on the quality and timeliness of its data and information services.

## 1.5.2 Update on Datasets

The present NEODC data holdings comprise commercial satellite datasets acquired in support of environmental research since 1972; the majority of these data comprise the Landsat series together with smaller holdings of SPOT, Radarsat, ERS-1/2 SAR, AVHRR and Ikonos imagery over the UK and worldwide. The NEODC holds complete sets of the satellite imagery used by NERC to create the UK Landcover Maps for 1996 and 2000. From 2004/05 there has been a significant increase in the size of the NEODC satellite archive since the Data Centre accepted responsibility for the complete ATSR-1/2 data archives (~40Tbytes). This will increase further (80 TB over the following two years) with the addition of the AATSR archive and the conversion of ATSR-1 and 2 into AATSR format, creating a seamless 15-year (A)ATSR data archive.

The present NEODC data holdings also comprise the total archive of airborne data - digital multi-spectral imagery and photography - acquired by the NERC Airborne Research and Survey Facility (ARSF) since 1982.

Specific datasets added to the NEODC archive during 2004/05 included:

- NEXTMap Britain high-resolution digital elevation dataset
- ATSR-1/-2 UBT product archive
- ARSF data covering part of the 2004 flying campaign

The NEODC also holds a secure secondary off-line archive of the total AVHRR and SeaWiFs imagery acquired by the NERC Dundee Satellite Receiving Station (DSRS); this secondary archive is augmented quarterly each year.

Significant projected future additions to the NEODC archive include:

- NERC ARSF campaign data
  - o Remaining data from 2003 flying campaign
  - o Remaining data from 2004 flying campaign (including routine LiDAR and digital photography acquisition)
- (A)ATSR seamless archive
  - o AATSR L1B, L2 and Meteo products (historical and ongoing)
  - o ATSR-1/-2 in AATSR data format (L1B, L2 and Meteo products)

The majority of the NEODC data holdings are probably unique in the UK in the context of their local UK geographical coverage and the period of data coverage.

Certainly the total archive of the NERC Airborne Research and Survey Facility datasets – comprising Airborne Thematic Mapper (ATM), Compact Airborne Spectrographic Imager (CASI) and aerial photography is a unique collection both in the context of the data characteristics and their temporal coverage. Similarly the complete satellite imagery of the UK - predominantly Landsat and SPOT - which formed the basis of the UK Landcover Maps created by NERC for 1996 and 2000 is almost certainly unique.

It is possible to duplicate much of the commercial satellite imagery acquired over the preceding three decades by NERC, and lodged with the NEODC, by purchase from other sources. It is not possible to duplicate any of the NERC airborne datasets which are the sole property of NERC and for which NEODC acts as primary repository.

## 1.5.3 Update on System Infrastructure

As the price of disk space has continued to fall over the year, and with the move to closer operational harmony with the BADC (and thus sharing of generic resources such as disk space), it is intended that as much ATSR UBT data as possible will be loaded online, reducing the need to access tape data. Over the coming months, more network-attached storage servers will be purchased and populated with data as soon as possible. NEODC will continue to purchase additional storage with BADC as our joint needs increase. It is intended to move towards a system where services provided to both data centres from within the BADC/NEODC group are shared to reduce any unnecessary duplication and resources are used as efficiently as possible (e.g. finances, charging, centrally provided resources, etc).

The majority of the NEODC's previous airborne and satellite data holdings are now stored online, with the Atlas DataStore at RAL as secure backup.

The "Footprints" helpdesk software, also used by the BADC, has continued to provide an efficient means of logging and handling enquiries by email and telephone.

## 1.5.4 Update on Services

Web access statistics for the NEODC web server <a href="www.neodc.rl.ac.uk">www.neodc.rl.ac.uk</a> for 2004/5 show a similar level of activity as the previous year. Around 2,000 individual "visits" (see Appendix 3) took place each month, representing a total of nearly 580,000 individual HTTP requests ("hits") - see Appendix 1 for definition of "visits" and "hits".

Data delivery to registered users via the web-based download facility continued successfully, in addition although an "NEODC-branded" version of the BADC data browser has been implemented. This provides self-service access (via a system of access control linked to a shared user database) to the various NEODC data collections, and is now used on three datasets: NEXTMap and ATSR UBT and SHAC2000.

## 1.5.5 Update on Liaison and Publicity

- Steve Donegan attended the NERC EO Conference in Plymouth in June 2004 and gave a poster presentation promoting the activities and datasets of the NEODC
- Steve Donegan attended the 2004 ARSF Workshop (joint with RSPSoc 2004) in Aberdeen in September, and gave an oral presentation promoting the activities and datasets of the NEODC
- Steve Donegan attended the 2004 AGM of the Geological Remote Sensing Group in London in December 2004 and gave a poster presentation promoting the activities and datasets of the NEODC.

 Victoria Jay attended the first meeting of NCAVEO, the NERC Network for Calibration and Validation of EO data (January 2005) where she met EO researchers from Centres of Excellence and the wider EO community.

## 1.5.6 Update on Collaboration

Matt Pritchard (MJP) attended the regular meetings of the NERC Data Management Advisory Group (DMAG).

MJP continued involvement in the evaluation of the MapsDirect UK mapping service which underwent a 12-month trial by NERC, ending 31/12/2004.

MJP and Steve Donegan (SJD) formed part of a consortium with Plymouth Marine Laboratory and the British Antarctic Survey, and provided input to a bid to the NERC EO Enabling Fund for expediting access to the Rothera AVHRR archive, which was successful. NEODC was allocated 30 days of staff time to carry out metadata creation and preparation of a web interface in early 2005.

MJP provided input to a NERC consortium bid for the validation and provision of ice thickness products from Cryosat, led by CPOM/UCL, which was successful. NEODC was allocated £15k to undertake metadata creation and web interface development work, due to start 2007/8.

SJW and MJP worked with a consortium of EO-related companies (NPA Group, Comsine Ltd, and RSI UK) led by QinetiQ for the development of services as part of ESA's Service Support Environment SSE. NEODC services were developed to enable access to parts of its Landsat-7 and ATSR UBT collections. These are now incorporated into the ESA SSE Portal, <a href="http://services.eoportal.org">http://services.eoportal.org</a> for a demonstration period. Technology used is already contributing to future improvements to NEODC internal infrastructure. Following the success of this project, the technologies and infrastructure used will form the basis of the data delivery services to be provided under the DMAG Enhancement project.

MJP continued collaboration with QinetiQ and Comsine Ltd regarding the use of NEODC metadata in a BNSC ICP2 project to demonstrate desktop GIS access to z39.50 catalogues for searching for EO data, and the use of an OGC-compliant Web Coverage Server for data delivery. This project has now been completed and NEODC z39.50 server was used successfully. NEODC now has a working OCG WCS which can be used to demonstrate delivery of some of its Landsat-7 data.

MJP and Victoria Jay (VLJ) worked with a consortium led by Edinburgh University and including the Met Office, University of Leicester, RAL and SOC to prepare a bid in response to NERC's call for knowledge transfer funding to exploit the archive of (A)ATSR data for a climate-quality global sea surface temperature record. NEODC's involvement in this "ATSR-Climate" project would be to facilitate the data processing, and to archive and make available to the community, the outputs of the project.

MJP has assisted with work to prepare the successful bid to DEFRA/Space ConneXions for AATSR Archiving. This major (~£0.5M) project to create a seamless archive of (A)ATSR data products is now underway with over 80 TB of data to be made available via the NEODC archive.

System development and programming work is underway on the EOEF-funded project to develop the conversion tool from ATSR-1/2 UBT format to AATSR L1b format, which is a crucial part of the proposed seamless (A)ATSR archive.

MJP has been involved with a project initiated by ESA to copy the ATSR data set to UKPAF as part of the "Oxygen" concept and ESA's wish to have its own parallel (A)ATSR archive for international users. ATSR-2 UBT products from 1995, 1996, 1997 and 2000 have now been supplied, and procedures are in place to begin the large-scale copying of the remaining ATSR-2 data and associated preview imagery and metadata, to be followed by ATSR-1 UBT products when available.

VLJ has cooperated with NCAVEO (NERC Network for Calibration and Validation of EO data) to define the potential role of NEODC in this network. As a result, it is anticipated that NEODC will host a database (of instruments, field sites and cal/val laboratories) and a shared workspace for NCAVEO participants to exchange documents and preliminary data files. NEODC has been invited to become a partner in the network and have a representative on the NCAVEO Steering Committee. This should help the incorporation of strategic issues (e.g. compatibility with NERC Data Grid and e-Science) into the future plans for the network, and raise the profile of NEODC in the EO community.

#### 1.5.7 Other relevant information

SJD's and MJP's availabilities have been shared between NEODC duties and the collaborative project with QinetiQ on services for ESA's Service Support Environment. As a result, the main thrust of the development work under the DMAG enhancement project was put on hold for some of this year, but has now resumed in collaboration with the NERC DataGrid team. NERC will benefit significantly from NEODC's involvement in this ESA project, since the technology developed and demonstrated as part of this project are directly transferable to the delivery of enhanced services to NEODC users, and as such will now be incorporated into the DMAG enhancement project.

#### 2 FINANCE OVERVIEW

**Note:** Funding for the DMAG project is, for this financial year, included in the figures below. From April 2005, financial reporting was reorganised internally so that monies received by NEODC from DMAG are managed separately from the EO funding line, and separate 2005/2006 Technical Annexes have been prepared for EO and DMAG. The intention is that future reports on progress against DMAG funding and milestones will be reported to the NERC Data Management Coordinator.

## 2.1 Expenditure for 2003/04

The full cash cost budget for operating the NEODC in Financial Year 2004/05 was £172k from the NERC Earth Observation Directorate plus, from DMAG £105.5k.

Due to the partial redirection of efforts of MJP and SJD to complete the QMASS project, and the late appointment of the new NEODC staff member, it became clear that NEODC had not spent at the anticipated rate for the year as a whole, so funds were reprofiled to enable the programmes of work to be carried forward into the next financial year.

The actual spend against the EO budget of £172.0k in 2004/05 was £113.5k comprising expenditure as follows:-

- £48.5k for staff at CCLRC/RAL
- £23.1k recurrent
- £35.3k overheads (CCLRC/RAL)
- £6.5k capital

(£51k deferred to 2005/2006: milestones delayed as described in 1.3.2.2 and 1.3.2.3)

The actual spend against the DMAG budget of £105.5k in 2004/2005 was £52.7k, broken down as follows:

- £28.9k for staff at CCLRC/RAL
- £2.6k recurrent
- £21.1k overheads (CCLRC/RAL)

(£43k deferred to 2005/2006: milestones delayed as described in 1.3.2.4; £10k committed to hardware purchase early in 05/06)

# 2.2 Budget for 2005/06

The budget allocation for the NEODC from the NERC Earth Observation Directorate for 2004/05 is as follows:

Core activities £103.0k

New post (S. White replacement) £65.0k

Carry over from 04/05 £51.0k

Total from Earth Observation £219.0k

In addition, the following budget is allocated to NEODC from DMAG for 2005/6

DMAG Enhancement project: £69.5k Carry over from 04/05 £43.0k **Total from DMAG** 

£112.5k

## 3 FUTURE DEVELOPMENT & STRATEGIC FORWARD LOOK

## 3.1 Update on Strategic Goals

#### 3.1.1 Collaboration with NERC EO Centres of Excellence

Since the recruitment of Dr Victoria Jay who joined NEODC on 1<sup>st</sup> December 2004, collaboration work with the NERC EO Centres of Excellence has been underway. Following a survey and series of meetings, data management plans have been drafted which address the data curation and facilitation needs of the centres. These will be completed in discussion with the centres' Directors and researchers and work will begin, where resources allow, to support the centres with data access, distribution and archiving.

# 3.1.2 DMAG Project: Earth System Science: Enhanced Delivery and Visibility for Multiple Earth Observation Datasets

The NEODC was awarded £225k funding over 3 years for this project which started in October 2003. The overall objectives of this project are:-

- Developing a flexible framework for providing services alongside NEODC datasets. For example, there will soon be a software tool to convert from ATSR UBT format to AATSR format. This conversion routine will need to be "wrapped" in the necessary interfaces to enable it to be available "on demand" via the NEODC website. The framework developed here will enable this and other similar services to be provided, using generic Web Service technology to wrap existing software tools. Ever closer interaction with the NERC DataGrid is a likely and necessary by-product of this work.
- Developing an improved and extensible system for the automated creation of metadata for EO data products.
- Groundwork for hosting upcoming datasets, such as Airborne LIDAR, MODIS, MERIS, SeaWiFS and AATSR.
- Enhancement of the ATSR UBT archive by the creation of preview images and location maps, and additional helpdesk support.

#### 3.1.2.1 (A)ATSR Seamless Archive

The major new development for the NEODC is the DEFRA-funded project to create a seamless archive of products from the (A)ATSR series of instruments, with the following scope:

- To produce and maintain an archive of consolidated AATSR and ATSR GBTR (Level lb) global 1 km data, the Level 2 averaged and gridded products, and Level 2 Meteo products.
- To supply a consolidated set of the (A)ATSR Level 2 Meteo products to the Hadley Centre.
- Provide (A)ATSR products to DEFRA and NERC scientific users

This bid was successful and work has begun in April 2005. This will be a hugely important dataset, bringing together a unique time-series of observations needed to address current scientific challenges. As part of this project, agreement has been reached to allow access for NERC researchers to this valuable data set, which will greatly benefit the NERC community at minimal cost to NERC (by its funding of a key component: the ATSR-to-AATSR conversion utility).

## 3.1.2.2 Rothera AVHRR Archive

NEODC is involved with BAS on a project led by PML, funded by a recent NERC EO Enabling Fund grant to provide access to ARIES AVHRR data from Rothera base in Antarctica. This work is now underway to be completed in the course of 2005.

## 3.1.3 Future Funding Opportunities

## 3.1.3.1 Cryosat: Archiving of Ice Thickness Products

The successful Cryosat Consortium Bid, led by CPOM/UCL includes a work package for NEODC to provide ongoing data management, and a web interface to the archive of validated ice thickness products from Cryosat.

## 3.1.3.2 EO Centres of Excellence Data Management

New work has involved the creation of formal data management plans with the NERC EO Centres of Excellence, and signals the start of a new phase of greater engagement with NEODC's scientific community. A large-scale exercise to scope the requirements of the Centres, and to a certain extent the wider EO community, has almost been completed, and a number of data management activities have emerged as priorities. Depending on the scope of these activities, some can be carried out with existing core funding, but to address the more ambitious requirements it is likely that bids for additional funding will be made to NERC.

#### 3.1.3.3 ATSR-Climate

A bid has been prepared and submitted led by the University of Edinburgh in response to NERC's call for knowledge transfer funding. The aim is to exploit the archive of (A)ATSR data for a climate quality global sea surface temperature record. NEODC's involvement would be to facilitate the data processing, and to archive the output data from the project and make it available to the community.

#### 3.1.3.4 ESA interface

The NEODC hopes to take on a role as portal to ESA data for the UK EO community. Ideally, this would involve delegated registration of users for ESA datasets and acting as a go-between to smooth out access and delivery problems (perhaps using the BADC – UKMO relationship as a model). A dialogue between NERC/NEODC and ESA is being initiated to investigate possibilities in this area and initial indications are that ESA has an interest in pursuing a collaboration with NEODC.

## 3.1.3.5 Other Projects

The NEODC now has a range of new exciting projects to work on for the NERC community, listed above, and will be concentrating on achieving success with these over the next years. The engagement with the EO community has already been increased significantly and will continue to do so in the coming year. Following feedback from this interaction it is intended that the efforts of NEODC can be directed to better meet the community's needs, with respect to data acquisition, provision, communication, interface and liaison with other data providers.

## 3.1.4 Opportunities for Improvement

## 3.1.4.1 Engagement with the Scientific Community

The NEODC has significantly increased its engagement with its scientific customers during the past few months. The interaction with the EO Centres of Excellence was the first step along this road,

which will be coupled with greater publicity and communication with the scientific community. It is clear that it is important to be seen to be not only a valuable data archive but a coordinating body for the management and exploitation of EO data.

Following the meeting in Reading in March with representatives from the Centres of Excellence, a position document outlining the required roles and functions of a data centre serving the EO community has been drafted and sent out via the NERC EO mailing list (attached as appendix). Feedback on this will be incorporated into an outline plan for NEODC, and depending on the scope of activities required outside those already existing and covered by core funding, a proposal will be submitted to NERC.

## 3.1.4.2 Handling of Enquiries

The introduction of the "Footprints" helpdesk software has improved the way enquiries are logged and handled internally, enabling distribution of enquiries to a larger team of NEODC/BADC staff. Full integration with the BADC user database and address book is due shortly and will add full functionality to this system. A further improvement which will allow extraction of data access statistics from the system will follow.

#### 3.1.4.3 IT infrastructure

The IT infrastructure of the NEODC has improved considerably and will continue to benefit from increased sharing of resources with the BADC. Further online storage capacity will be acquired as necessary and, where possible, resources will be managed centrally within NEODC/BADC, to reduce duplication of effort.

## 3.1.4.4 Physical data storage and cataloguing

There is still much work to be done in arranging easily-accessible physical storage and metadata descriptions for many datasets that are currently held by the NEODC. With the increased staff resource now available, it is intended that the backlog of data which is stored, but not properly catalogued and therefore "visible", will start to be cleared. More efficient ingestion procedures will be put in place that will ensure that metadata creation keeps pace with physical storage.

#### 3.1.4.5 Staff resources

As mentioned in earlier sections there were a number of reasons why NEODC underspent over the last financial year as a whole. The payments which have been deferred to 2005/06 will be put to use to increase staff resource within NEODC to address the opportunities for improvement mentioned above.

## 4 Appendices

## 4.1 Appendix 1 Projects supported

Study of Iron Age in Lambourn Downs Area

Archeology of East and West Coasts of Scotland

Analysis of sustainable energy sources by estimating solar irradiance from satellite data

Modelling of Scatness archeological site using imagery to create land-use data

Role of recent glacier fluctuations for Icelandic ice-marginal landscapes and sediment budgets

Flood effects on Rio Granado, near El Churro, Spain

*Urban land use, London* 

Carbon sequestration over urban centres (Edinburgh, Oxford, Sheffield, Leicester, Glasgow)

Development and validation of a lowland permeable catchment coupled hydrological and fine sediment dynamics model (Ref. NER/S/S/2003/11927)

Coral Reef Ecology and Monitoring

Risk taking in Semibalanus balanoides (Firth of Clyde)

*Use of satellite imagery to track jaguar habitats in jungles in Belize* (Ref. NER/S/A/2004/12432)

Spatial variation of sediments in the Avon Estuary, Southhams, Devon, and quality of sediment with regard to nutrients (nitrate and phosphate) (Ref. NER/S/M/2004/12811)

Modelling British plants & animals: using NEXTMap data to generate climate surfaces and habitat suitability maps at a fine resolution (Ref. R8-H12-01)

Characterisation of groundwater-surface water interactions in a number of groundwater fed river catchments (Ref. NER/T/S/2001/00948)

Developing models of morphological change in coastal case study areas, in response to climate change and sea level rise.

Checking georectification for ARSF imagery

Modelling and predicting pollution from diffuse sources of phosphorus (Ref. NER/S/J/2003/13105)

Capturing geomorphological change for the coastal simulator

Landscape Evolution of South-Central England

Sediments and nutrient problems in the Avon Estuary, South Devon (Ref. NER/S/M/2004/12811)

Accelerated upland erosion following recent wildfire (Ref: NER/A/S/2003/00790)

Geocorrection and removal of topographic effects from CASI-2 data (Ref. ARSF 04/26)

# 4.2 Appendix 2 NEODC position paper 'Data Management for the EO Community'

#### Introduction

NERC's Data Policy requires the curation of data generated by NERC-funded research, which is implemented through a set of designated data centres. Additionally, NERC encourages, and often explicitly funds the efficient exploitation of data via its data centres, rather than supporting duplication of effort in the community. In the case of the Earth Observation community, the NEODC provides curation and facilitation support (in addition to the services offered by the other data centres which are also available to the EO community).

NERC has recently funded a 6-month scoping study aimed initially at understanding the data curation and facilitation requirements of the Centres of Excellence, but with the eventual aim of identifying the requirements of the wider community. While significant information has already been collected, and a meeting held in Reading on 23 March with representatives from ESSC, DARC, CLASSIC, CTCD, BADC and NEODC, it is clear that only a sample of the community has been contacted and/or visited.

Accordingly, this document is a general position paper, in order to garner input from the wider community via the NERC EO mailing list. The main thrust of the paper is to establish a "UK EO data<sup>3</sup> commons" where the sum of the entities is greater than the sum of the parts, but this requires the community to do work to support each other.

#### 1. Data Management

The objectives of data management in general are to ensure that:

- NERC preserves the investment in gathering data for future scientific exploitation (it is expected that individuals and their employing institutions retain the rights for commercial exploitation if this is possible).
- NERC funds are most efficiently used and there is minimal unnecessary duplication of effort in handling datasets in the community.

These twin objectives are and will be implemented by ensuring that:

- Datasets of use to the wider scientific community are archived and distributed in a suitable manner
- Potentially scientifically valuable data are kept for the long-term.
- High quality documented data archives are created.
- Appropriate metadata standards are used to facilitate the identification, re-use and integration of scientific data and provide information that future scientists can use to evaluate the data.
- Appropriate data support is provided to scientists

#### Additionally, ideally:

• Scientists should be recognised and given credit for the scientific contribution of the datasets that they produce as well as for the analysis of those data (i.e. datasets become citeable publications)

<sup>&</sup>lt;sup>3</sup> It is understood that EO data means any data required in the EO community, and that NEODC and the other NERC data centres will endeavour to support the EO community as an integrated activity, not as individual data centres.

#### 2. Activities

In this section a series of requirements are presented, which as outlined above, came from a subgroup of EO centres of excellence. Some of these requirements are already satisfied, or should be satisfied, by existing core funding of the NEODC. Existing data management activities within the centres of excellence are also not covered here.

#### 2.1 Central Archive Functions

#### (Activities that should be covered by current funding)

As a matter of course, NEODC should be seen as the default data centre for new datasets produced by the EO community.

- NEODC needs to clearly define what archival activities are currently covered by existing funding arrangements.
- Where appropriate, the datasets may be held elsewhere, but NEODC should still be advised of all datasets produced in the EO community. All recipients of NERC EO funding need to be regularly made aware of their obligations and how to meet them.
- Proposals for NERC funding of EO-related science should include some provision for
  ongoing data management of the data inputs to, and outputs from, the proposed activities.
  Such proposals may or may not include centralised resources but will be audited against the
  NERC data policy. Where these data fall outside the scope of the existing archival activities
  and/or funding arrangements, the NEODC should be able to provide a simple cost model to
  quote in such proposals.

#### (Activities that would require additional funding)

- The community should be regularly polled to identify key datasets being used, or desired by the community, and such datasets should be acquired (if necessary purchased) and made available centrally.
  - The prioritisation of such datasets is a community activity, and will need to be done
    by an appropriate group of community representatives. See below for a suggested
    committee structure to support this.

#### 2.2 Search Function

#### (Activities that should be covered by current funding)

• NEODC operates and continually enhances its web site in order to provide an effective guide to information sources for aspects of Earth Observation and links to other NERC, UK and worldwide sources of Earth Observation data. The web site includes both links and a catalogue.

#### (Additional Search Function)

- The NEODC catalogue interface should provide information about data holdings in other organisations, and the information and links provided should be responsive to community requirements.
  - As a consequence, there needs to be a mechanism for community requirements to be fed into the NEODC on a regular basis, and an infrastructure to support catalogue interoperability.

#### 2.3 Clearing House

(Existing Activities)

• NEODC provides an on demand acquisition service for satellite- and airborne-derived imagery required in support of NERC science together with the provision of an effective archive and access service for these data (e.g. Landsat, SPOT, Ikonos).

## (Additional 'Clearing House' Activities)

- A more general "clearing house" function should exist for any EO data purchased with NERC funding, where users can check what others have already acquired, and assistance can be provided to purchase data if required.
  - o There needs to be a mechanism by which such datasets can be registered, and/or lodged securely so they can be shared in ways that do not violate associated licenses.
  - o The community needs to be educated to do this.

### 2.4 Internal Community Interface.

#### (Existing Activities)

- The NEODC already provides services to deliver EO data products it holds and associated
  information in response to requests from accredited users, it provides data curation for
  remotely sensed imagery acquired by the NERC Airborne Research and Survey Facility and
  other NERC remotely sensed data (e.g. DSRS AVHRR secondary archive, NERC/BNSC
  SHAC data etc.)
- (Additionally, as a special activity this year, the NEODC has been funded to provide advice and assistance to develop data management policy and data protocol documents for EO centres of excellence)

## (Additional Internal Community Interface Activities)

- Not all data held in the UK EO community is held at the NEODC (or any other NERC data centre). It is desirable that significant data holdings in the community are discoverable and obtainable by other members of the community.
  - O Data holders should either make such datasets available from the NEODC or publish catalogues (in an appropriate format) and make the data available from their own sites. In either case, the EO researchers themselves have to take responsibility for populating this catalogue; this is an *education activity*.

#### 2.5 External Community Interface

## (New activities)

- Interface with ESA: A dedicated portal for UK users of ESA datasets is required. Ideally, this would involve delegated registration of users for ESA datasets and act as a go-between to smooth out access and delivery problems.
- Interface with other agencies: Where information is required regarding e.g. access to other agencies' EO data, there should be a contact point who can either provide information or refer to others who can.

#### 3. Implementation Requirements

In summary, the following tasks have been identified which are not already covered by existing NEODC core activities:

- The prioritisation of key datasets used or desired by the community, by an appropriate group of community representatives
- Operate a mechanism for community requirements for data from other agencies to be fed into the NEODC, and an infrastructure to support catalogue interoperability.

- Operate a "clearing house" system, including a data registration/sharing mechanism, and educate community to use it
- Educate EO researchers to make datasets available from the NEODC or publish catalogues and make the data available from their own sites
- Interface with ESA and other agencies

#### 3.1 Financial Implications

If the new activities proposed are to occur, additional funding will be required to implement the functions described above. One outcome of feedback to this position paper will be a proposal to NERC for such funding.

#### 3.2 Datasets of interest

As identified above, a regular appraisal of dataset priorities should be carried out. Initially, NEODC has asked the EO Centres of Excellence to indicate the highest priority datasets to which NEODC should facilitate access. Obviously, non-CoE researchers need to be involved in the polling process.

In the longer term, it is suggested that a dataset priority form could be accessed via a website and EO researchers are regularly prompted to fill in such a form. Where feasible, the highest priority data should then be acquired for distribution and availability publicised.

o When a priority dataset/activity list is produced, a budget will be required, and a mechanism to make the decisions of what/how much should be done.

#### 3.3 Steering committee

To make decisions on the direction and activities of data support, an advisory body representing the EO community should be available for consultation, and meet regularly (perhaps annually). Any proposal based on this position paper will include outline terms of reference for such a body. The membership would normally consist of representatives both of the CoEs and the wider EO community.

## 4.3 Appendix 3 NEODC Website access statistics for FY 2004-5

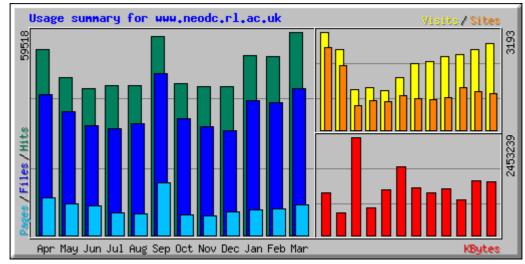


Figure 1 Graphical summary of access statistics for www.neodc.rl.ac.uk

Summary by Month										
	7			Sun	iiidi j	by William				
Month	Daily Avg				Monthly Totals					
Month	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Mar 2005	1919	1388	285	90	1182	1331648	2816	8851	43029	59518
Feb 2005	1868	1385	277	94	1262	1359183	2632	7762	38786	52304
Jan 2005	1875	1411	267	88	1383	900059	2475	7477	39527	52508
Dec 2004	1553	1093	250	85	1065	1162673	2382	7020	30618	43508
Nov 2004	1446	1062	193	74	979	1059265	2220	5813	31861	43402
Oct 2004	1431	1100	198	69	1031	1201361	2167	6144	34111	44368
<u>Sep 2004</u>	1938	1574	513	57	1126	1707709	1725	15396	47246	58165
Aug 2004	1410	1054	200	41	923	1145979	1299	6218	32693	43720
<u>Jul 2004</u>	1409	1008	214	44	953	700054	1390	6651	31274	43691
<u>Jun 2004</u>	1428	1073	284	43	800	2453239	1312	8541	32192	42865
May 2004	1489	1171	300	85	2104	559716	2639	9312	36319	46179
<u>Apr 2004</u>	1810	1375	366	106	2704	1051808	3193	11006	41277	54315
							I	II.	I	
Totals						14632695	26250	100191	438933	584543

Table 1 Tabular summary of access statistics for <a href="www.neodc.rl.ac.uk">www.neodc.rl.ac.uk</a> (see following notes for explanation of terms: e.g "visits" vs "hits")

#### **Explanatory notes for web access statistics**

Figure 1 and Table 1 show access statistics for the NEODC website at <a href="http://www.neodc.rl.ac.uk/">http://www.neodc.rl.ac.uk/</a>. These statistics are produced by the *Webalizer* analysis program (available from <a href="http://www.mrunix.net/webalizer/">http://www.mrunix.net/webalizer/</a>).

The input data to these statistics consist simply of the Apache web server access log file, once the relevant log files covering the time period of interest had been concatenated. Filtering was applied to exclude from the statistics any visits to the site from computers operated by NEODC staff (by means of specifying their IP addresses), which would have skewed the statistics.

Several measures are reported, defined as follows:

- *Hits*. These are individual requests to the web server for any item, be they HTML documents, images, data files, requests to run a server-side script or other items. When a remote host (i.e. a computer elsewhere on the internet) requests a web address (say, a web page) from the web server, the loading of that page in the client's browser usually results in the requesting of many more files (e.g. images, stylesheets) which are part of that page, but count as separate "hits". Note that requests to non-existent pages or web addresses also result in "hits".
- *Files.* These are items successfully returned from the web server to the remote host. A request does not always result in a "hit", for a variety of reasons (e.g. mis-typing an address, files that have moved location, incorrect access permissions held by the client, etc.).
- *Pages*. These are text documents written in HTML, or generated on-the-fly by a request to a server-side script, that result in HTML text being transferred to the client's web browser. Images and other "files" are excluded (on the basis of their filename extension).
- Sites. These are remote computers, recognised as distinct IP addresses when recorded in the log file, which have submitted requests for items on the web server. Depending on how the client's access to the internet is configured, it is possible that multiple computers connected to the same network and sharing a web proxy (or cache) can result in only 1 "site" being reported for these multiple computers (Example: a classroom full of students using the internet, all submitting requests to the web server, may all count as 1 "site", [and, according to next definition, as one "visit"]). "Site" statistics are therefore likely to be conservative.
- *Visits*. Requests from the same "site" that are either the first request from that "site", or separated from the last request from that "site" by a period of 30 minutes. This is probably the most useful measure reported in the statistics, and is probably conservative. For example, someone browsing through the web site will initially access the front page, at which time the clock starts ticking and all subsequent requests to the webserver from their "site" still count as one visit, so long as no more than 30 minutes inactivity is recorded. The first request after a period of 30 minutes' inactivity will result in the visit count being incremented.
- *KBytes.* Kilobytes of data transmitted by the web server in response to successful HTTP requests from remote computers.