The AMT data management experience

Claudia Castellani*, Gwenaëlle Moncoiffé and Juan Brown

British Oceanographic Data Centre, Joseph Proudman Building, 6 Brownlow Street, L3 5DA, Liverpool, Merseyside, U.K.

Abstract

As the UK's National marine data centre, a key responsibility of the British Oceanographic Data Centre (BODC) is to provide data management support for the scientific activities of complex multidisciplinary long term research programmes. Since the initial cruise in 1995, the NERC funded Atlantic Meridional Transect (AMT) project has undertaken 18 north-south transects of the Atlantic Ocean. As the project has evolved there has been a steady growth in the number of participants, the data volume, complexity and the demand for data. BODC became involved in AMT in 2002 at the beginning of phase II of this programme and since then has provided continuous support to both the AMT and wider scientific community through rescue, quality control, processing and access to the data. The data management comprises a team of specialist data managers using a sophisticated infrastructure of software and hardware to manage, integrate and serve the physical, biological and chemical data. Here, we discuss the approach adopted, techniques applied and some guiding principles for management of large multi-disciplinary programmes.

***Corresponding author: cxc@.bodc.ac.uk**

Keywords: Atlantic Meridional Transect, data processing, Inventories, Information centres, Instrument platforms, quality control.

Acknowledgement: Funding was provided by the UK Natural Environmental Research Council. We thank all the AMT participants for submitting their data to BODC and for improving the quality of data management through their feed-back. Particular thanks go to Carol Robinson for her continuous support and help with the acquisition of the data and to Dawn Ashby for her superb assistance. The ship personnel team is also gratefully acknowledged for their technical support and for supplying the CTD and underway data to BODC.

1. Introduction

In the enthusiasm that surrounds exciting and novel science it is easy to forget that a unified approach to data management aimed to ensure consistency in data quality, ease of dissemination between collaborators, secure archiving and future utilisation of the data is critical to the success of projects (Lowry *et al*., 2005). This is particularly true of major multi-disciplinary and multi-partner programmes. Collection of marine data involves expensive platform operations, such as running research vessels and training of qualified personnel. In addition, because of the inherent difficulties in their collection marine measurements are irreplaceable assets that are both expensive to collect and uniquely document the changing earth system.

Over the past two decades the scientific community has become increasingly aware of the importance of time series measurements as baselines for the quantification of changes in the marine environment (Beaugrand, 2002). It is essential that future generations of scientists and research users are able to access environmental data that has been properly managed and preserved, to known standards and formats, together with basic background information on how the data were collected and analysed (metadata). Scientists and funding bodies have also a legal and moral responsibility to make the data and results from publicly funded activities widely available. Moreover, insurance is required against accidental loss and technological redundancy through changes in storage media. Delivering all this requires a dedicated infrastructure of complex relational databases and staff that understand the data and their use as well as information technology.

The scarcity and high value of oceanographic data has led increasingly to the development of a culture where long-term data curation and data sharing, through specialised national data centres, has become the norm rather than the exception (Glover *et al* 2006, Seys *et al* 2006). Since its creation in 1979 the British Oceanographic Data Centre (BODC) has provided support to the UK marine science through its data management activities, by delivering fully integrated and quality controlled data for future utilisation by research scientists, the government, industry as well as the wider public. BODC deals with biological, chemical, physical, and geophysical data and its databases contain measurements of nearly 15,000 different variables. Its staff includes data scientists from a wide range of scientific disciplines who have direct experience of marine data collection and analysis working alongside information technology specialists developing and maintaining the databases and software infrastructure required to support data management and data distribution systems. One of the key responsibilities of National data centres such as BODC is to provide data management support for the scientific activities of complex multidisciplinary long term research programmes. In the present paper we discuss our experience in managing the data arising from the Atlantic Meridional Transect (AMT) project and provide some guiding principles for the data management of large multi-disciplinary programmes.

2. Overview of the AMT data set

Between 1995 and 2005 AMT has undertaken 18 cruises essentially twice yearly between the UK and the southern Atlantic Ocean which involved 46 international research groups (Table 1; Robinson, *et al.*, 2006). The programme was divided into two phases (i.e. phase I and phase II), separated by two years during which there was no fieldwork.

The scientific aims of phase I were to assess mesoscale to basin scale phytoplankton processes, the functional interpretation of bio-optical signatures and the seasonal, regional and latitudinal variations in mesozooplankton dynamics. During phase II, the program was broadened (Figure 1 and Robinson *et al.*, 2006) to address a suite of cross-disciplinary questions concerning ocean plankton ecology and biogeochemistry and their links to atmospheric processes. Broadly, the measurements comprised hydrographic and bio-optical properties, biogeochemistry, aerosol and rainwater composition, plankton community structure and plankton physiology, as summarised in Robinson *et al.* (2006) and Figure 1.

The objectives included the determination of how 1) the structure, functional properties and trophic status of the major planktonic ecosystems vary in space and time; 2) physical processes control the rates of nutrient supply, including dissolved organic matter, to the planktonic ecosystem and 3) atmosphere-ocean exchange and photodegradation influence the formation and fate of organic matter. Determinands include temperature, salinity, inorganic nutrients, carbon dioxide, oxygen, nitrous oxide, methane, dimethylsulphide, dissolved and particulate carbon, nitrogen and phosphorus, chlorophyll, phytoplankton pigments and taxonomy, photosynthesis, respiration, new production, nitrogen fixation, calcification, bacterial and microzooplankton abundance and activity, microbial molecular diversity, viral activity, mesozooplankton community structure and physiology (i.e. respiration, copepod egg production rates), atmospheric dust deposition and characterisation, bio-optical properties, and coloured dissolved organic matter (Table 2). An “individual dataset” or “AMT dataset” in the context of this paper is a single or multiple sets of related measurements originating from a single or a team of data originators during a given AMT cruise.

3. Challenges posed by the AMT dataset management: phase I and phase II

The two phases of the AMT project presented BODC with very different challenges. When the formal data management began at the beginning of phase II in 2002 our database system was already designed to deal with the complexity of biological and biogeochemical measurements. However, the number of individual sources of data, the size of the AMT community, each successive cruise often having new scientists and students and the available funding all presented significant challenges for the successful management of the data. The relative importance of these factors varied for each phase.

3.1 Phase I

As no provision had been made for data management during phase I, BODC and the AMT community were essentially involved in a data rescue effort. First additional funding was secured. As a large number of individuals (including temporary staff & PhD students) had been involved in collection, the data were in a varied state of repair. Consequently, the data had been kept in assorted formats on a range of media, there were no authoritative protocols for collection and often there was no definitive version of data sets. Thus, there was considerable potential for: data loss; confusion during analysis and subsequent interpretation; and wasted effort tracking down and re-processing data. The process of acquiring these data was helped considerably as the benefits of a central data management effort were realised during phase II and by the realisation of AMT participants that data from phase I were not readily available.

3.2 Phase II

Phase II presented the opportunity to undertake a more formal and structured approach to the data management. Initially, many scientists were unaccustomed to a culture of integrated data management. Whilst BODC had pioneered such an approach during the NERC North Sea programme in the early 1980's (Lowry *et al*., 2005) and later during other NERC and European projects, a period of organisational upheaval in UK marine science during the late 1990's meant that a number of the guiding principles had to be re-learned. Previously, it had often been the practice that BODC data scientists would accompany cruises to provide direct support on fundamental on-board data management activities (e.g. CTD and underway data processing). Regrettably, space constraints precluded BODC personnel participation during AMT cruises. This situation led to the development and implementation of a more formalised strategy for cruise preparation. This included the writing up of guideline documents of what was expected at the end of a cruise with regards to data and ancillary information associated with CTD and underway continuous measurements and also individual scientist data. Additionally, there was insufficient continuity in shipboard personnel to ensure consistent calibration and quality of the CTD and underway data. Consequently, it proved most efficient for BODC to undertake these tasks for which further resource had to be found. Figure 2 illustrates the disparity between the initial estimate of effort required for the data management and that actually expended.

4. How are the data managed by BODC?

4.1 BODC communication and networking culture

BODC's philosophy is that communication lays the foundations to good data management. Data management begins with the exchange of information between the different parties involved in the project (i.e. BODC staff/AMT data manager, scientists and ship personnel) at the planning stage and throughout the fieldwork. Whilst space constraints prevented data scientists from participating on AMT cruises, BODC participated in the AMT planning and science meetings, preparation of newsletter and contributed to reports. As mentioned previously, a number of what should have been standard working practices had to be re-learned to attain the processes described below.

4.2 Initial acquisition of cruise data in liaison with the ship technical personnel & the PI

Before the starts of each cruise BODC liaise closely with the ship's technical personnel based at the National Marine Facility and the Principal Investigator (PI). This enables data scientists to understand the scope of data collection activities and the nature of equipment deployed, but also enables BODC to provide a set of guidelines to assist in good data management practices and the recording of the metadata necessary for the accurate future description of the data. Following the cruise, the PI is required to compile a comprehensive list of the data collected during the cruise into the Cruise Summary Report (CSR, formerly the

"ROSCOP") conceived by the Intergovernmental Oceanographic Commission (IOC) in the late 1960s to provide a low level inventory for tracking oceanographic data collected on Research Vessels. It is expected that the PI submits the CSR, within a week of the cruise and a full "Cruise Report" collating the cruise narrative and a description of the methodology of the data sets collected by the scientists on board, within 6 month from the end of a cruise. These documents provide the key information outlining the nature of the data collected and the techniques employed. Alongside this, an electronic version of the CTD and underway navigation data is submitted with the appropriate documentation relating to the calibration and configuration of the scientific instrumentation employed.

4.3 Data tracking and banking

The initial source of information collated by the PI at the end of the AMT cruise is recorded by BODC into a series of inventory tables, one for each data set collected, together with the details of the scientist responsible for the data. The inventory tables are interfaced with the Oracle database to display the AMT data holdings dynamically on the BODC website (http://www.bodc.ac.uk/projects/uk/amt/data_inventories/). The records contained in the inventory tables are continuously updated following dialogue with data originators and used to keep the AMT participants informed about the availability and the processing status of the data.

4.3.1 Physical security

Clearly, physical security of data is one of BODC's primary concerns and the organisation's approach to this has two components: an 'accession system' and an 'archive system' for long-term preservation of the data. When the data arrive at BODC they are recorded on an electronic accession table and the data copied into the inventory via the Unix operating system. The physical integrity of the data is secured by preservation of the original media together with a copy placed in the BODC data archive and wherever possible, an additional version of the data supplied is saved into ASCII format. The archive system is supported by an accession system containing the metadata record which provide the data submission with a unique identifier and describe its contents and provenance.

4.3.2 Data reformatting and standardisation

Marine data sets cover an unusually wide range of physical, chemical, biological and geological data types. Even a single water sample may be analysed for several hundred parameters. Often, scientists use different words to describe the same type of data. All this means that there can be confusion when seeking data, which may cause errors in reports or misunderstandings between parties. Many organisations also want to be able to provide, search and manipulate data over the internet. To be able to integrate these data into a database and be able to use them reliably there must be no uncertainty surrounding the terms that are used to describe data. Therefore, standardisation of the file format and the parameter defining the data is essential for their professional management.

4.3.2.1 The BODC file format

CTD and underway data are reformatted into the QXF format which is a BODC defined subset of the binary format netCDF. Other data sets relative to discrete measurements are handled in ASCII format prior to loading into a relational database under the Oracle Relational Database Management system. Additionally, each of the data maintained by BODC is held in standard units and it is assigned a parameter code described by the Parameter Dictionary (http://www.bodc.ac.uk/projects/uk/enpardis/).

4.3.2.2 The Parameter Dictionary

To solve the problem of terminology standardization we have established a dictionary of terms which maps directly to dictionaries used by other leading international data management organizations. Containing more than 18,000 terms, the parameter dictionary developed by BODC (http://www.bodc.ac.uk/projects/uk/enpardis/) is a powerful data mark-up tool which uses a single 8-bytes parameter code to associate a data value to its parameter name and methodology through a semantic model. The names of biological entities in the parameter dictionary have been also standardised against the Integrated Taxonomic Information System (ITIS) that further enriches the metadata through access to a biological taxonomy.

4.3.3 Data Quality control

After reformatting and attributing parameter codes, the quality control of many oceanographic data sets is operated through data visualisation. The two main approaches to data quality control are either via screening using bespoke soft-wares or simply via direct visualisation of the data by experienced data scientists:

4.3.3.1 EDSERPLO: the BODC way forward to CTD profile and time series data screening

Edserplo (Editing and SERies PLOtting) is the soft-ware developed by BODC for screening 1- and 2-dimensions continuous data series including continuous underway and CTD profile data. This software can be used to visualise multiple parameters and series and it has a quick editing tool which allows the quality control of data through the flagging spikes and suspect data points (Fig 3). Since 2006 EDSERPLO runs on PC computers compared to a previous version, operating on Silicon Graphic stations, allowing a more efficient and faster processing of the data.

4.3.3.2 Quality control of non continuous profiles or time series data

Discrete data measurements measured, for instance, from CTD cast bottles, net hauls or during experiments are also quality controlled through direct visualisation of data points on spreadsheets and graphical plots of the data. After the data are screened, quality controlled and reformatted they become integrated in the databases.

4.4 Integration of AMT data in the databases

Once reformatted and quality controlled, all the data collected during AMT cruises which originates from discrete water column sampling, continuous profilers, tows or benthic sampling, are managed within the BODC databases where they are fully integrated with concomitant oceanographic measurements and associated metadata. The data are stored into the BODC databases into a series of tables linked to each other to various degree of complexity to minimise duplication of information. Figure 4 shows a simplified representation of the BODC databases within the 3 main groups of tables;

- The sampling metadata tables consisting of the sampling activity or event description table with links to the fieldwork description table and to the sampling gear code table.

- The data tables consisting of a series of linkage and data storage tables for each main type of sampling or data collection techniques.

- The parameter dictionary tables defining the 8-byte parameter codes

All data storage tables have the same field structure consisting of a unique key linking each record to the linkage tables, a data parameter code, a data value, a data quality control flag, a data originator code and a loading date. The data linkage tables invariably control the one-to-many relationship between the events and the data. Their structure has been adapted differently for each main type of sampling or data collection technique in order to incorporate specific metadata information such as for example bottle depth and bottle type for water collection events, plankton net depth range, mesh size, mouth area for plankton net hauls. The structure of the database is such that it may be easily expanded to include new sampling gear, new methodologies and instrumentation, and new parameters.

5. AMT data policy and data dissemination

The distribution of the AMT data is regulated by the data policy (http://www.bodc.ac.uk/projects/uk/amt/data_policy/) drawn up by the AMT Scientific Steering Committee (SSC). The policy was developed to ensure an appropriate balance between the protection of data originators' intellectual property rights and the potential benefits that may arise via data use by the programme, the wider research community and other interested parties. According to this policy, when AMT data are transferred to BODC they become available to other investigators within the AMT programme on the condition that the originator is kept informed about how the data are being used and he/she is acknowledged in any exploitation of that data. The AMT data can also be made available to the wider scientific community, immediately upon permission being granted by the data originator and the signature of a licence or after 6 months from the end of the program in April 2006.

BODC supplies the AMT data to both internal and external users either directly via email, the ftp system or in the case of the CTD and underway data, via automatic download through the BODC website (http://www.bodc.ac.uk/data/online_delivery/amt/). Before any of the AMT data can be downloaded, however, users are required to be registered as BODC web users. The data are supplied in a fixed ASCII (CSV) format and they are associated with documentation providing information about sensors, quality control, calibrations and processing status. Each time modifications to data are made by BODC, these files are updated and changes are noted in the documentation.

6. AMT data requests

Requests for AMT data can be made directly to the BODC staff for all data or via the BODC website for CTD and underway data. Data requests and queries made directly to BODC are handled by either the AMT data manager or the BODC requests officer. Data requests and the processing of data sets required by PhD students are given priority over all other requests. When data are lodged at BODC, it is the aim that requests be serviced within 1 to 3 working days. Currently only the CTD and underway cruise data can be downloaded from the BODC website whereas discrete data sets can only be obtained contacting the BODC staff. However, BODC is developing a system for online delivery of its data holding and AMT data will become available alongside all other publicly available data held in its databases.

Statistical information on AMT data request is only available from 2002. Since then 570 requests for AMT data have been handled (Fig. 5). Figure 5 also shows that data requests have steadily increased from 16 in 2002 to190 in 2005. This increase in data requests reflects both the expansion of the scientific community to approximately 180 scientists from 11 countries and the provision of an on-line system for the automatic download of the CTD profiles and the underway data. On the other hand, the slight decrease in 2006 is probably the results of the conclusion of the phase II of the program and the graduation of many of the "AMT PhD students" (who accounted for a significant proportion of the data requests). Since 2003 (i.e. after the introduction of the BODC web data delivery system) the proportion of data requests from the web has represented between 47% and 50% of total requests. Interestingly, however, although the number of AMT requests received from external users has also increased, their proportion has remained overall stable over time at around 20 % of the total

AMT requests (Fig 5). The largest data request was received from USA scientists followed by that of the UK and French scientific community (Figure 6).

7. Discussion and conclusions

The planning and implementation of research and the efficient management of the resulting data often appear to be two widely separate worlds. This is because, whereas data managers consider data collection, management and dissemination as essential for the effective use of research funds, many researchers consider data management as technical and secondary to publications. As a consequence data management is often insufficiently planned or not planned at all (Seys et al 2006). Perhaps, the realisation by the scientific community, in the 1990s, of the importance of long term time series to predict the impact of future changes in the earth system on human activities together with the need to handle increasingly large and complex data sets, has resulted in a change in attitude towards the role of marine data management.

A long-term multi-disciplinary international programs such as AMT provide numerous challenges for the managements of the data they produce in terms of their large volume, diversity of measurements, changes in sampling instrumentation and increased sophistication of analytical methods. Increase in diversity and number of variables measured also means increasing demand and needs from the end users for their rational integration. Measurements also need to be supported by a rich metadata if they are to have value for the future. Clearly, when it comes to satisfying the requirements of modern science on the delivery of fully integrated and quality controlled complex data sets, such challenge cannot longer simply be met by individuals or single scientific organisations without large investment of time and resources. BODC through its support to multidisciplinary programs has specialised in handling data sets which are small in volume, but extremely high in complexity covering a wide range of physical, chemical, biological, atmospheric and geological parameters. Progresses in the acquisition and management of AMT data were facilitated by a culture of communication and collaboration between the data scientists and IT specialists with the AMT scientists and ship technical staff. Thanks to this collaboration, BODC has rescued > 90 % of the data collected during the phase I of the project by scientists who have moved on in their careers often loosing contact with the AMT community.

Within modern data management, the role of a National Data Centres such is also evolving from one of data repository to one of active data delivery and redistribution between scientific partner organisations. Such evolution is being fostered and made possible thank to concomitant technological advances in IT soft-wares and the advent of the internet. Although considerable effort has been spent by BODC to ensure the physical security of all data accessed, its primary mission is to guarantee that data may be reused with confidence and without any need for recourse to the data originator no matter how much time has elapsed since the data were submitted. Ensuring this forms the bulk of the work of BODC's data scientists. Thus, another important role played by BODC for the AMT community, and for other large multidisciplinary programs, has been in its ability to manage and disseminate data of increasing complexity through the development of bespoke software and web-portals. The development of the BODC Parameter Dictionary and of screening soft-wares like EDSERPLO, has greatly contributed to the effective data management and dissemination of data. BODC data scientists also use sophisticated software, to extract quickly and efficiently integrated data sets from the databases to satisfy data requests. In addition, the on-line data distribution system has introduced by BODC since 2003 via its website has also allowed the expansion and efficient data delivery within and beyond the AMT community.

Finally, it is important to note that the funding initially allocated for the data management of AMT represented considerably less than the minimum required to manage the data of a large multidisciplinary program as estimated by Glover *et al*., (2006). In their review of the data management of the US JGOFS project, they recommended that good data management requires devoted resources to be between 5 - 10 % of the total cost of the programme. Overall, the AMT program received funding by NERC of the order of £2.38 millions of which £62k (2.6%) was initially allocated for data management. In the final analysis an additional £140k has been sourced, meaning the total cost of the AMT data management will exceed 8.5% of the total funding.

8. Lessons learnt and recommendations for managing future long term programs

The following recommendations can be drawn from the experience acquired by BODC through the data management of a large multidisciplinary program such as AMT:

- **At the beginning:** The data center should be involved at the planning stage of any scientific program to ensure that data management can be properly planned and resourced.

- **A sound data policy:** A comprehensive data policy must be drawn at the start to enable the steady and harmonious integration of the data into the database and their subsequent dissemination in relation the protection of the intellectual property rights of the data originator.

- **Early contact with PIs:** An early data management liaison with the project PIs improves data quality and the identification and logging of the information in a data inventory.

- **A good cruise report:** The preparation of a detailed and comprehensive cruise report by the cruise Principal Scientist in collaboration with the technical ship personnel is key to the subsequent management of the data collected during the cruise

- **Proactive sea-going support:** The activity undertaken by the ship technical personnel in monitoring the functioning of the underway ship sensors and during CTD operation also represents a crucial step towards gathering high quality data and support for successful data management.

- **Early data acquisition:** To prevent/minimise the risk of their loss, data and metadata should be acquired by the data centre as early as possible, from the time of their collection. This is particularly important when data processing and delivery is handled by staff on short term contract and students.

- **Data management culture:** Close collaboration with project participants including attendance at planning and science meetings, and where possible data scientists participating on cruises helps foster a culture of rigorous Data management.

- **Data centralization:** Centralisation of the data sets collected during the scientific program through their submission to the data centre also prevents confusion that may arise through the dissemination of multiple versions of the same data sets by different parties.

References

Atlantic Meridional Transect, NERC Consortium programme (2005) 2004/05 Review of the NERC Funded Marine Sector Consortium Programmes. NERC, Swindon, 14 April 2005, 90 pp.

Beaugrand G, PC Reid, F Ibanez, JA Lindley, M Edwards (2002) Reorganization of North Atlantic Marine Copepod Biodiversity and Climate, Science, 296: 1692- 1694

Glover DM, Chandler CL, Doney SC, Buesseler KO, Heimerdinger G, Bishop JKB, Flierl GR (2006) The US JGOFS data management experience. Deep-Sea Research II, 53:793-802

Lowry R, Rickards L and Brown J (2005) Adding Value to Oceanographic Data at the British Oceanographic Data Centre. Proceedings of the Ensuring Long-term Preservation and Adding Value to Scientific and Technical data (PV 2005) Conference, Edinburgh, November 2005.

Robinson Carol, Alex J. Poulton, Patrick M. Holligan, Alex R. Baker, Grant Forster, Niki Gist, Tim D. Jickells, Gill Malin, Rob Upstill-Goddard, Richard G. Williams, E. Malcolm S. Woodward, Mikhail V. Zubkov (2006) The Atlantic Meridional Transect (AMT) Programme: A contextual view 1995–2005. Deep-Sea Research II 53:1485–1515

Seys J, Mees J, Vanden Berghe W and Pissierssens P (2006) Marine Data management: we can do more, but can we do better? (www.iode.org)

List of Figures

List of Tables

Figure 1:

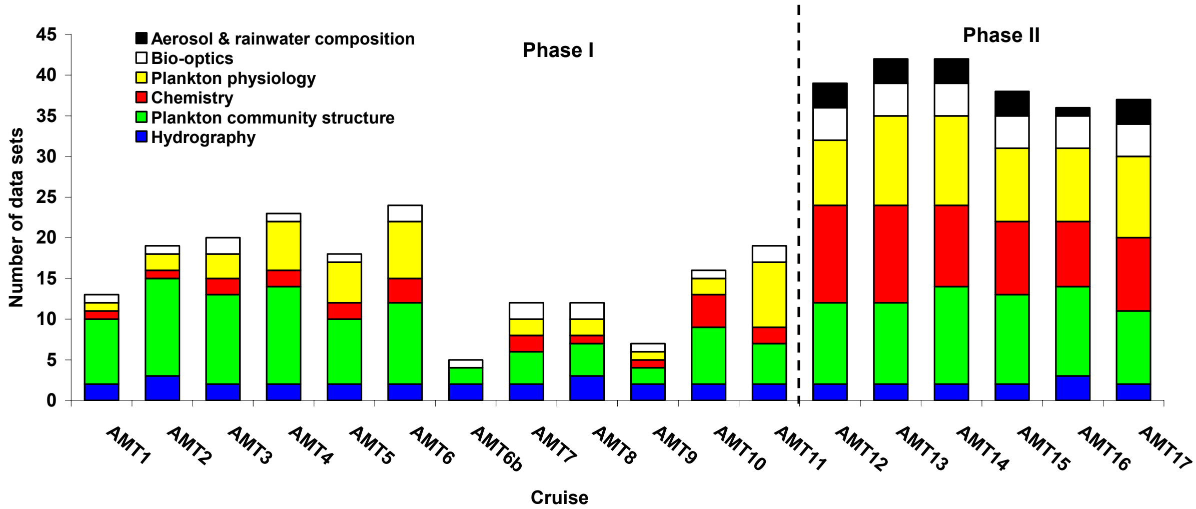

Figure 2:

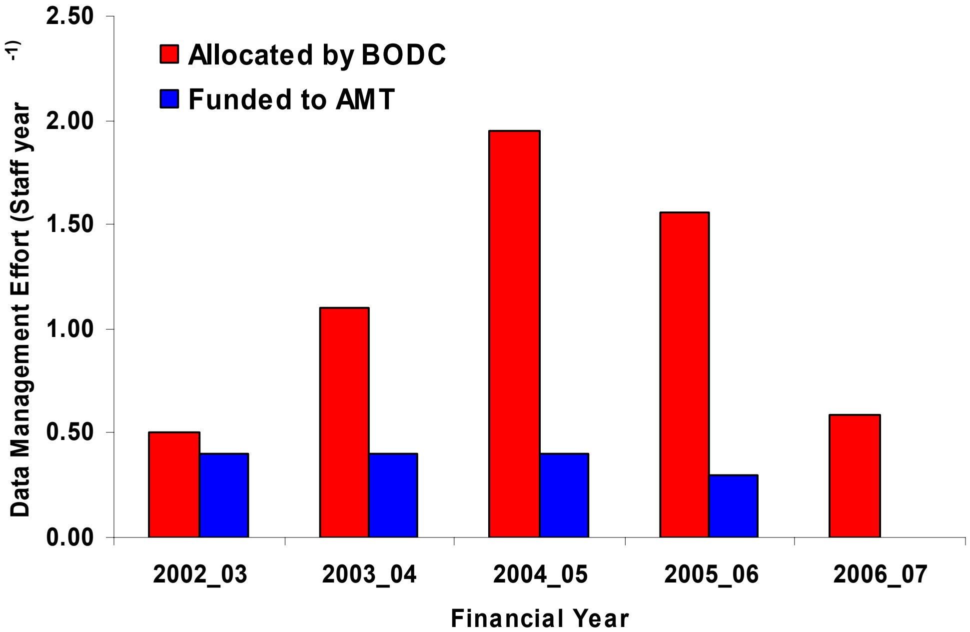

Figure 3:

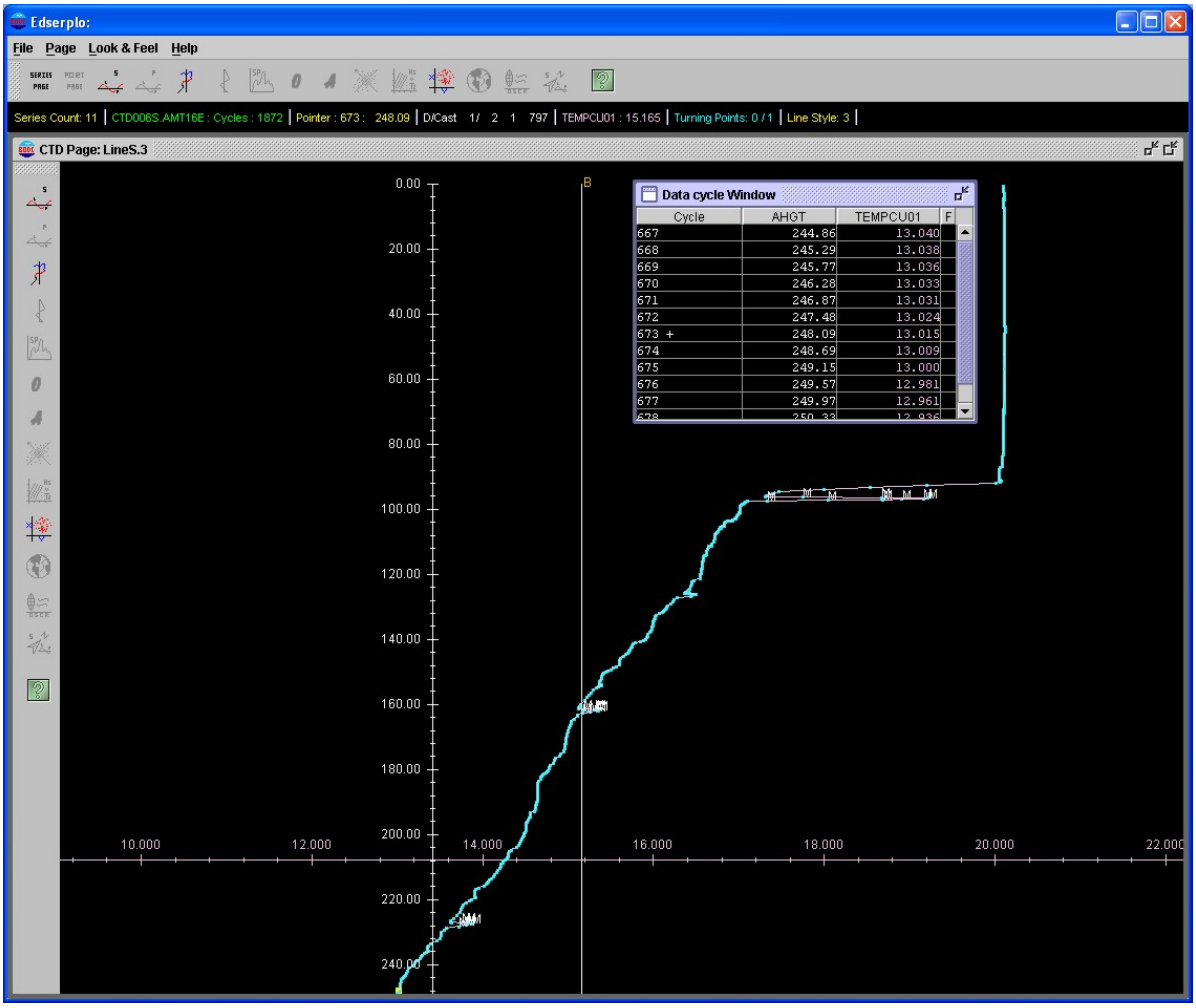
Edserplo:
File Page Look & Feel Help
Series Count: 11 | CTD006S.AMT16E : Cycles : 1872 | Pointer : 673 : 248.09 | D/Cast 1/ 2 1 797 | TEMPCU01 : 15.165 | Turning Points: 0 / 1 | Line Style: 3 |
CTD Page: LineS.3
Data cycle Window
Cycle AHGT TEMPCU01 F
667 244.86 13.040
668 245.29 13.038
669 245.77 13.036
670 246.28 13.033
671 246.87 13.031
672 247.48 13.024
673 + 248.09 13.015
674 248.69 13.009
675 249.15 13.000
676 249.57 12.981
677 249.97 12.961
0.00
20.00
40.00
60.00
80.00
100.00
120.00
140.00
160.00
180.00
200.00
220.00
240.00
10.000
12.000
14.000
16.000
18.000
20.000
22.000

Figure 4:

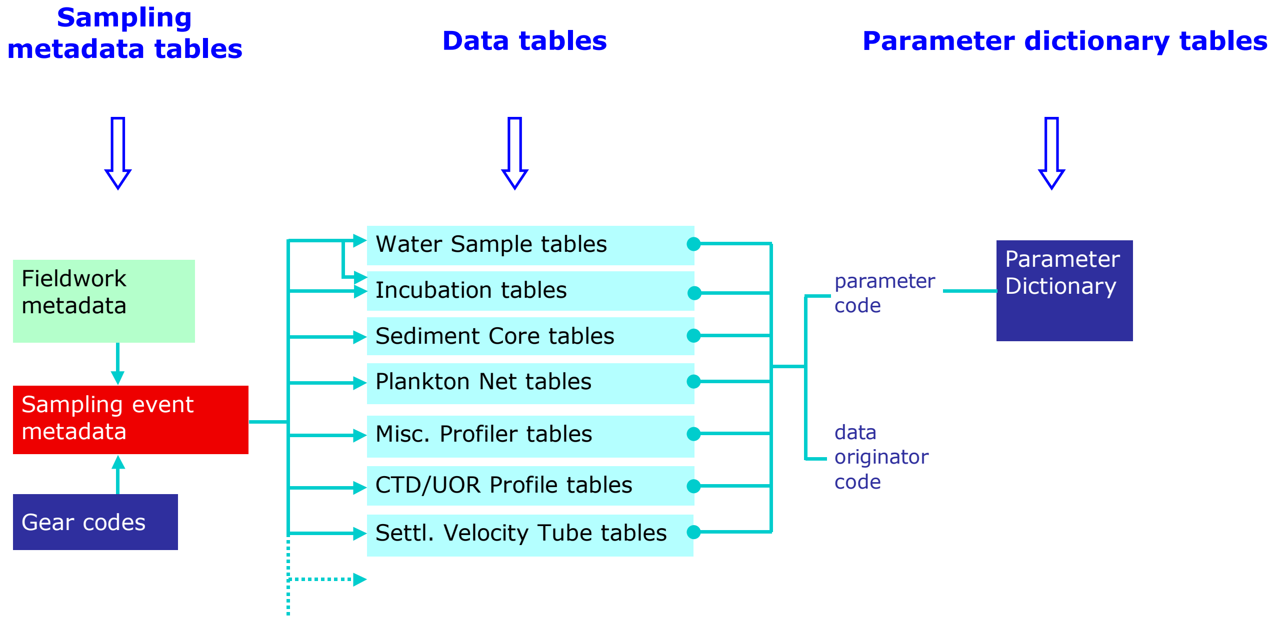
Sampling metadata tables
Data tables
Parameter dictionary tables
Fieldwork metadata
Sampling event metadata
Gear codes
Water Sample tables
Incubation tables
Sediment Core tables
Plankton Net tables
Misc. Profiler tables
CTD/UOR Profile tables
Settl. Velocity Tube tables
parameter code
data originator code
Parameter Dictionary

Figure 5:

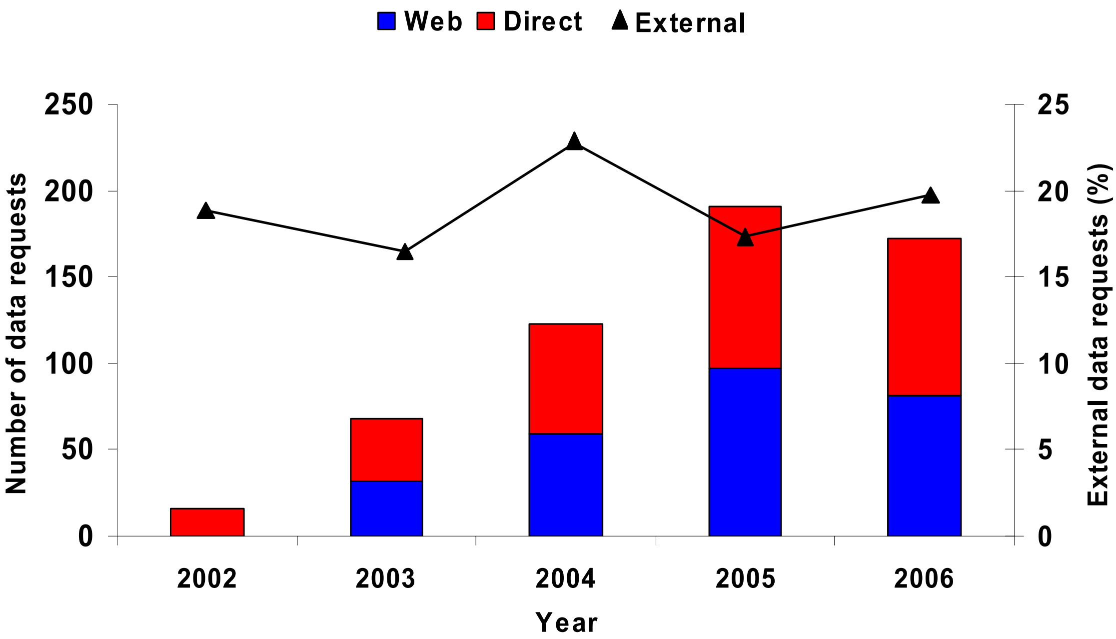

Figure 6:

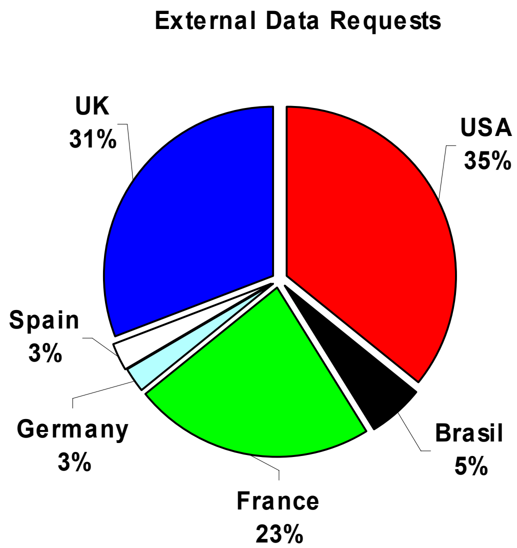

Table 1:

Cruise	Departure/Arrival	Date	PSO/Institute
AMT1 (JCR XX1)	UK – Falkland Islands	21/09/95 – 24/10/95	Dave Robins/ PML
AMT2 (JCR 13)	Falkland Islands – UK	22/04/96 – 28/05/96	Dave Robins/ PML
AMT3 (JCR 15)	UK – Falkland Islands	16/09/96 – 25/10/96	Tony Bale/ PML
AMT4 (JCR 21)	Falkland Islands – UK	21/04/97 – 27/05/97	Tony Bale/ PML
AMT5 (JCR 23)	UK – Falkland Islands	14/09/97 – 17/10/97	Jim Aiken/ PML
AMT6b (JCR 31)	Falkland Islands – UK	05/04/98 – 04/05/98	Gerald Moore/ PML
AMT6 (JCR 32)	South Africa – UK	14/05/98 – 15/06/98	Jim Aiken, PML
AMT7 (JCR 34)	UK – Falkland Islands	14/09/98 – 25/10/98	Jim Aiken/ PML
AMT8 (JCR 41)	Falkland Islands – UK	25/04/99 – 07/06/99	Nigel Rees/ PML
AMT9 (JCR 45)	UK – Uruguay	15/09/99 – 13/10/99	Nigel Rees/ PML
AMT10 (JCR 49)	Uruguay – UK	12/04/00 – 07/05/00	Chris Gallienne/ PML
AMT11 (JCR 52)	UK – Uruguay	11/09/00 – 13/10/00	Malcolm Woodward/ PML
AMT12 (JCR 90)	Falkland Islands – UK	12/05/03 – 17/06/03	Tim Jickells/ UEA
AMT13 (JCR 91)	UK – Falkland Islands	10/09/03 – 14/10/03	Carol Robinson/ PML
AMT14 (JCR 101)	Falkland Islands – UK	26/04/04 – 02/06/04	Patrick Holligan/ NOC
AMT15 (D 284)	UK – South Africa	17/09/04 – 29/10/04	Andrew Rees/ PML
AMT16 (D 294)	South Africa – UK	19/05/05 – 29/06/05	Tony Bale/ PML
AMT17 (D 299)	UK – South Africa	17/10/05 – 28/11/05	Patrick Holligan/ NOC

Table 2:

Measurement Category	Measurement	Protocol	Cruises
Aerosol & rainwater composition	Organic compounds	High volume air sampler	12, 13, 14,15, 17
	Ammonia	Low volume vacuum pump	12, 13, 14, 15, 17
	Major ions and trace metals	High volume air sampler	12, 13, 14,15, 16, 17
Bio-optics	Coloured Dissolved Organic Matter absorption	Waveguide+spectrometer	6b, 11, 12, 13, 14,15, 16, 17
	Inherent Optical Properties (absorption, scattering and attenuation)	One or more of ac9, particle absorption by filter papers, VSF	1, 2, 3, 6, 12, 13, 14, 15, 16, 17
	Phytoplankton fluorescence	Fast repetition rate fluorometer (FRRF)	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	Apparent Optical Properties (upwelling/downwelling light and calculation of diffuse attenuation coefficient)	Freefall and/or cable lowered light sensors	3, 4, 5, 6, 7, 8, 12, 13, 14, 15, 16, 17
Biogeochemistry	Hydrogen peroxide	Flow injection analytical system	16
	Urea	Autoanalyser	11
	Halocarbons	Gas chromatography Mass Spectrometry	8
	Carbon monoxide	Gas chromatography	10
	Nitrous oxide	Gas chromatography	12,13
	Methane	Gas chromatography	12,13
	Alkalinity	Titration	12, 13, 14, 15, 17
	DMS, DMSP concentrations	Gas chromatography	5, 9, 12, 13, 14
	pCO2	Infrared gas analyser	12, 13, 14, 15, 16, 17
	Dissolved organic carbon	High temperature catalytic oxidation	12, 13, 14, 15, 16, 17
	Dissolved oxygen	Winkler titrations	6, 12, 13, 14, 15, 16, 17
	Dissolved organic nitrogen	High temperature catalytic oxidation	10, 12, 13, 14, 15, 16, 17
	Dissolved organic phosphorus	Colorimetric analyses	10, 12, 13, 14, 15, 16, 17
	Dissolved iron	Flow Injection Chemiluminscence	3, 6, 12, 13, 15,16, 17
	Dissolved inorganic carbon	Coulometric titration	6, 7, 12, 13, 14,15, 16, 17
	Nitrate, nitrite, phosphate, silicate,	Autoanalyser, waveguide, gas diffusion and fluorescence (NH4+)	1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15, 16, 17
Hydrography	High vertical & horizontal resolution hydrography	Towed CTD (undulating oceanographic recorder,	2, 8, 16

		UOR; moving vessel profiler, MVP)	
	Attenuation, temperature, salinity, fluorescence, oxygen, transmission	CTD, SBE oxygen sensor	1, 2, 3, 4, 5, 6, 6b, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
	Underway navigation, bathymetry,fluorescence, temperature, salinity, meteorology, transmission	Thermosalinograph; Autosal salinometer, SurfMet	1, 2, 3, 4, 5, 6, 6b, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
Plankton community activity	Phytoplankton excretion	Incorporation of 14C into dissolved organic carbon	6
	Silica uptake	Uptake of 32Si, Liquid Scintillation Counter	14
	Mesozooplankton respiration and ammonia production	Winkler oxygen titration, nutrient analysis	11
	Net community production	O2/Ar ratios by Membrane Inlet Mass Spectrometry (MIMS)	16, 17
	Nitrification and ammonium regeneration	14C with nitrification inhibitors, 15N isotope dilution	6, 13
	Calcification	Incorporation of 14C into particulate inorganic carbon Liquid Scintillation Counter	14, 16, 17
	Export production	Thorium disequilibria	12, 13, 14
	Microzooplankton herbivory	Modified dilution assay and flow cytometry	13, 15, 16, 17
	N2 fixation	Acetylene reduction & 15N uptake	12, 13, 14, 15, 17
	Nitrate, ammonium uptake	15N uptake, mass spectrometry	5, 6, 11, 12, 13, 14,
	Microzooplankton bacterivory	Modified dilution assay and flow cytometry	3, 4, 14, 15, 16,17
	Mesozooplankton grazing	Flow cytometry and epifluorescence microscopy	4, 5, 11, 12, 13, 15
	Bacterial production	Thymidine & Leucine incorporation Flow cytometry; Microautoradiography Fluorescence In Situ Hybridization (MARFISH)	2, 3, 4, 11, 13, 14, 15, 16, 17
	Gross and net community production Dissolved oxygen flux,	Winkler oxygen titration	4, 5, 6, 11, 12, 13, 14, 15, 16, 17
	Community respiration Dissolved oxygen flux	Winkler oxygen titration	4, 5, 6, 11, 12, 13, 14, 15, 16, 17
	Photosynthetic efficiency	Fast Repetition Rate Fluorometry (FRRF)	6, 7, 8, 9, 10, 11, 12, 13, 14, 15,16, 17
	Primary production	Incorporation of 14C into particulate organic carbon	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17

Plankton community structure	Aerobic anoxygenic phototroph production	Kinetic Infra-Red Fluorometry	16
	Viral abundance	Filtration then Polymerase Chain Reaction (PCR)	2, 15, 16
	Particulate inorganic carbon & nitrogen	Inductively Coupled Plasma Argon Emission Spectrometry (ICPAES)	14, 15, 16, 17
	Biogenic silica	Spectrophotometer	14, 15, 16, 17
	Particle size	Coulter counter	2, 3, 4, 6b
	Mesozooplankton distribution	Optical plankton counter	1, 2, 3, 4, 5, 6, 10
	Heterotrophic nanoplankton	Microscopy, enzymatic assay, flow cytometry	3, 4, 12, 13, 14,15, 16, 17
	Cyanobacterial abundance	Flow cytometry; Fluorescence In situ Hybridisation (FISH)	2, 4, 6, 12, 13, 14, 15, 16, 17
	Mesozooplankton size-fractionated particulate organic carbon & nitrogen	CHN analysis	1, 2, 3, 4, 5, 6, 10, 12, 13, 14
	Heterotrophic bacterial abundance	Flow cytometry; Fluorescence In situ Hybridisation (FISH)	2, 3, 4, 6, 11, 12, 13, 14, 15, 16, 17
	Particulate organic carbon & nitrogen	CHN analyser	1, 2, 3, 4, 5, 6, 12, 13, 14, 15, 16, 17
	Mesozooplankton taxonomy and abundance	Microscopy	1, 2, 3, 4, 5, 6, 9, 10, 12, 13, 14, 15
	Coccolithophore composition and abundance	Scanning Electron Microscopy; optical microscopy; Calcareous Optical Detection, Fluorescent In Situ Hybridisation (CODFISH)	1, 2, 3, 4, 5, 6, 7, 8, 10,11, 12, 14, 15
	Phytoplankton composition and abundance	Microscopy, flow cytometry	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12,13, 14, 15, 16,17
	Microzooplankton composition and abundance	Microscopy, FlowCam	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12,13, 14, 15, 16, 17
	Chlorophyll pigments	Fluorometry & high performance liquid chromatography	1, 2, 3, 4, 5, 6, 6b, 7, 8,9, 10, 11, 12, 13, 14, 15, 16, 17