ICES Advisory Committee on Ecosystems
ICES CM 2004/ACE:05 Ref. ACFM, ACME

Report of the Study Group on Management of Integrated Data

31 March–2 April 2004
Hamburg, Germany

TABLE OF CONTENTS

Section **Page**

EXECUTIVE SUMMARY

SGMID reviewed existing central databases at ICES and distributed databases within the ICES community, as well as data policies and experiences with the setup of integrated databases. It became apparent that there is an urgent need for the development of additional databases within ICES, *inter alia* as a prerequisite to develop an integrated advice for ICES clients. SGMID made a number of recommendations for future actions. Amongst these the group recommends that central databases physically located at ICES should only be developed where necessary (i.e. a database for the handling of catch-at-age and other primary fish stock assessment input data), while existing databases (such as survey databases) should be kept distributed in order to reduce the development effort. ICES should, however, take the lead in defining standards for these distributed databases and should become the focal point for their integration. ICES should, wherever possible, use integration tools already available rather than developing new tools. SGMID will focus on these technical solutions during its next meeting. SGMID also favoured open accessibility of ICES data rather than restricted access.

1 INTRODUCTION

1.1 Terms of Reference (ToR)

The **Study Group on Management of Integrated Data** [SGMID] (Co-Chairs: P. Wiebe, USA and C. Zimmermann, Germany) met in Hamburg during 31March – 02 April 2004 to:

a) review the development within ICES towards integrated databases of oceanographic, environmental, and fisheries data;
b) identify data sources relevant to a), above, not yet integrated into the ICES databases;
c) review existing integrated data systems for fisheries/environmental data and review data integration work in existing projects inside and outside of ICES;
d) propose strategies and technical solutions for integrating available data including the possibility that data are not physically located in one site;
e) evaluate and recommend the level of integration and aggregation of data in connection with management issues from an ecosystem perspective including the use of GIS systems;
f) evaluate problems associated with the accessibility of data.

1.2 Participants

Larry Atkinson	USA
Doug Beare	UK/Scotland
Ole Folmer	Denmark
Kristin Kaschner	Canada
Henrik Kjems-Nielsen	ICES Secretariat
Jule Harries (part time)	UK/England
Todd O'Brien	USA
Lesley Rickards	UK/England
Yolanda Sarminaga	Spain
Reiner Schlitzer	Germany
Hein Rune Skjoldal (part time)	Norway
Manfred Stein	Germany
Peter Wiebe (co-chair)	USA
Sunhild Wilhelms (part time)	Germany
Werner Wosniok (part time)	Germany
Christopher Zimmermann (co-chair)	Germany

1.3 Importance of data integration for ICES

There is an on-going effort at ICES and elsewhere to implement an ecosystem approach to management of marine resources. Assessment is an important element of an ecosystem approach. Assessment covers a range of activities from

more narrow thematic assessments of specific conditions or features, e.g., the status of a fish population, to broader assessments of environmental or ecosystem status, e.g. quality status reports (QSRs). ICES has a large number of working groups (WGs) that regularly contribute to or carry out assessments.

The various types of assessments put demands on access to a wide range of data from multiple sources. These include data on physical oceanography, plankton, benthos, fish, etc., as well as information on a wide range of human activities or influences on the marine environments (e.g., inputs of contaminants and nutrients). There is also a demand to have access to data in real or recent time. In assessing the current situation, it became apparent that the current practise of data access and data integration within the ICES community would not be sufficient to deliver integrated advice.

ICES therefore decided to set up a Study Group to give advice to the secretariat and ultimately to the community on how to proceed with data integration. There are a number of different study and working groups dealing with data (e.g., WGMDM, SGXML, SHHAC), but it was felt that these groups were dealing more with technical aspects of data handling than with the general perspective. SGMID was set up at the Annual Science Conference in 2002, but could not convene in 2003 due to a lack of participants.

1.4 Structure of the Report

The report is structured along the terms of reference for the group. Following the Introduction, the second section (relevant to ToR a) gives an inventory of the data held in ICES databases. These are physically located at ICES headquarters. Recent developments for data integration at ICES are reviewed, namely a web-accessible database holding metadata. A list of data- and database deficiencies is given along with an explanation why these data are needed to develop an integrated advice (relevant to ToR b). A prioritisation for future developments can be found in Section 6. The third section (addressing ToR f) reviews different policies in place for handling integrated data, focuses on data accessibility, and lists possible problems. This should assist in formulating a revision of the ICES data policy in the near future. The fourth section (relevant to ToR c) reviews a number of approaches for data integration and the lessons learned from these attempts. This should help ICES to develop such a system with as little effort as possible. This section also indicates where further information can be obtained from websites and contact details. The fifth section (ToR d) compares the different approaches and discusses quality control issues and the user's perspective to a future integrated ICES database. The sixth section (relevant to ToR e) provides an outline for the ICES secretariat on how to continue with database development and data integration. It also gives a number of recommendations with priorities indicated. The annex lists contact details for SGMID participants.

2 CURRENT STATUS OF DATABASES WITH FOCUS ON DATA INTEGRATION WITHIN ICES

2.1 ICES database overview (ToR a)

The table below shows all databases currently in the ICES Secretariat, and whether it is included in the inventory on the web.

Disciplines	Database	Year range	Approx. no. stations	Included in Inventory db on the internet
Oceanography	CTD	1970-2004	150,000	**Yes**
	Hydrochemistry	1891-2004	900,000	**Yes**
	Surface	1891-2004	1,500,000	**Yes**
	Pump	1980-2004	400,000	**Yes**
Environment	Contaminants in sea water	1977-2004	20,000	**Yes**
	Contaminants in sediment	1977-2004	4,000	**Yes**
	Contaminants in biota	1978-2004	6,000	**Yes**
	Fish Disease	1981-2004	3,000	**Yes**
	Biological community	1979-2004	-	No
Fisheries	DATRAS	1965-2004	16,000	**Yes**
	Stock Assessment summary	2001-2004	400	No
	Statlant (official annual catch data)	1973-2004	7,500	No

2.2 ICES Integrated Inventory (a metadata-base on the web)

ICES receives data from three different disciplines: fisheries, environment and oceanography. These data are validated and stored at the ICES Secretariat. A researcher can receive data if requested. For historical reasons, the request for data procedures has been split in the 3 disciplines, each with its separate ways.

The purpose of the ICES Integrated Inventory is to:

- Provide easy access on the internet;
- Easily be able to see what data are stored at ICES;
- Enable data searches across disciplines. One site including data from all 3 disciplines;
- Make it easy to request data. An email sent from the ICES homepage with request limits.

The ICES has completed the whole system which includes:

- Design and implementation of the database;
- Importation of the metadata;
- Development of an internet program (ASP);
- Testing of the system;
- Completing the implementation on ICES homepage

2.3 Ongoing developments for data integration

2.3.1 DATRAS (DATabase on TRAwl Surveys)

The DATRAS database is a database that integrates three separate fisheries trawl surveys, where the trawl survey data originate from IBTS, BITS and BTS. National laboratories submit survey data to ICES on the web. ICES validates and uploads the data to a Microsoft SQL Server relational database. Indices for some fisheries assessment working groups are calculated in DATRAS. The project is in the final phase and close to being completed.

2.3.2 DOME (Database on Oceanography and Marine Ecosystems)

The DOME database is going to hold data for HELCOM, OSPAR, and ICES on contaminants in seawater, sediments, and biota, including biological effects and fish diseases. The database is a development from existing systems where these data are held in separate databases. The development of this database is concurrent with the development of the new reporting format (version 3.2). The design of the database is going on at the moment.

2.4 Database and data deficiencies (ToR b)

Following the review of databases currently maintained at ICES, and considering the data handled by ICES working groups, it became apparent that most data needed for the development of integrated and comprehensive advice are already available within the ICES community. Most of these data are, however, not easily accessible and/or distributed in a variety of formats, or the level of disaggregation is insufficient. Examples for this are aggregated data in working group reports or documents, raw data in files of individual stock/species/area coordinates, in databases of the national laboratories that collate data for specific surveys. The current inaccessibility of important primary data has serious implications not only for the future development of the ICES advice, but is regarded to be a key issue for quality control on ICES WG's work.

In the view of SGMID, there are some key areas where ICES needs to take an initiative to make data available. The development of a database to handle (to aggregate) catch and **catch-at-age data for fish stock assessment purposes**, and to store assessment input data is certainly a matter of high priority. As this database should be usable by all assessment working groups, a common structure is required. Therefore, a central database held at ICES HQ appears to be the most cost-efficient solution. The catch-at-age/assessment database would address quality control issues rather than data integration issues, but is also a prerequisite for the development of a fleet-based advice and 4m-advice ("multi-fleet, multi-species, multi-area, multi-annual") in the near future. As funding for the development of a central database is available, SGMID expects it to be developed within the next year.

The situation is more difficult when it comes to **survey data**. Currently, data collected on multidisciplinary cruises are split into different fields before it is submitted to ICES (e.g., hydrography data, environmental data, fisheries data) and it appears difficult to reunite these data. ICES needs to address this issue. In addition, the only comprehensive dataset accessible through ICES for any of the numerous surveys conducted within the ICES area is originating from the Bottom Trawl Surveys (IBTS in different quarters, in the North Sea, the Baltic, and West of the British Isles). Most of these surveys are already coordinated by ICES groups. Data are available in very different formats and on different levels of aggregation at national co-ordinator's home laboratories. As the different surveys serve different purposes and use different methods (e.g., multidisciplinary surveys, plankton surveys [among these egg- and larvae surveys], trawl surveys, acoustic surveys, aerial surveys), the definition of a standard format or database structure appears to be an impossible task. To make these data accessible as fast as possible, and to use the expertise present in expert group and national laboratories, SGMID suggests that ICES focuses on tools for the integration of these distributed databases rather than creating a central database. This approach would, however, require the obligation of national labs to make the data available e.g. on a web-based database. This approach would also address quality control issues, as the data would be consistency checked by experts for the specific survey rather than by database specialists with limited expertise in all surveys they handle. Also, the correction of errors would be facilitated if only a single copy of the data were physically stored in one place.

There is an urgent need for access to **plankton** and higher taxonomic biota (**seabirds, marine mammals**) data, which are derived from the integrated survey (meta-)database. This is a prerequisite for the development of an ecosystem advice.

The sampling of **plankton** data (e.g., zooplankton biomass, composition, and abundance sampled by net and other gear) is illustrated for example in the WGZE Plankton Status Reports. It is suggested that ICES include these data in its current data processing efforts, starting with those data of the WGZE, WGPE, and WGHABD working groups and reports. Additional plankton data for the ICES regions are available from the US-NOAA/NMFS Global Plankton Database (contact: Todd.OBrien@noaa.gov).

Seabirds have been studied in most marine ecosystems of the world. While work has been targeting breeding seabirds in colonies over long time periods, work has also focused on seabirds at sea during the last three decades. Many databases have been established to collect data on breeding populations. However, annual numbers over many decades are only available for few places and/or species (e.g., southern North Sea). Much more often, seabird breeding censuses were conducted every 10–20 years as is the case e.g. in Britain and Ireland (around 1970, mid 1980s, around 2000). These data are stored in databases in connection with metadata such as methods employed, observers etc. From studies in seabird breeding colonies many data sets have been established that not only list population sizes, but also information on breeding biology, e.g., reproductive performance (number of eggs, percentage of eggs hatched, percentage of chicks fledged), adult survival, breeding phenology, diet, foraging trip duration. Usually, such data have been collected for one or few species at dedicated places. Since the 1970s, increasingly more data have been collected on the distribution of seabirds at sea, mostly as ship-based surveys, but also by aerial surveys. There are a number of methods employed that are comprehensively reviewed in the literature. In all these cases, databases have been set up to handle these data. Effort has been made to standardise counting methods and collection, which resulted for example in the founding of the European Seabirds at Sea Co-Ordinating Group, running the European Seabirds-at-Sea Database (Reid and Camphuysen, 1998). This database holds seabirds-at-sea data from nearly all groups working in NW Europe, with strongest emphasis on the North Sea and the waters to the north, west, and south of Britain and Ireland, but also including data from the Baltic Sea, the Bay of Biscay, and waters north up to Iceland. The database was first set up in 1990 and has been upgraded and updated on a regular basis. The current database format is a three-table structure with a table for bird-related information (e.g., species, number, behaviour), a table for positions and related observation conditions and a third table collecting general information (type of platform, observer, methodological details etc.). The third type of data collected originate from equipping individual seabirds with telemetry devices and/or data loggers. These instruments either transmit positional information (radio and satellite telemetry) or store data that can be downloaded (data loggers). The latter have usually been related to dive depths and behaviour of the birds, but comprise more and more positional information. Technological achievements have enabled seabird researchers now to collect synoptic very precise positional data by GPS data loggers along with very accurate temperature and water depth measurements. A review of these three different kinds of data collections would be very worthwile and would need the expertise of at least one person being familiar with seabirds in colonies and one person being allocated to seabirds at sea and data logger studies.

With respect to **marine mammals**, data should include information on and/or data required for the estimation of abundance, distribution, population dynamics, and mortality of marine mammals in ICES waters, as well as data that allows the assessment of impacts of other anthropogenic activities (e.g., pollution, seismic surveys, military exercises, etc.). This may require the incorporation of particular types of effort information not necessarily needed or provided in fisheries statistics (e.g., including gear types, deployment [soak] times, net lengths for by-catch estimation, etc.). Possible data sources include data obtained through observer program monitoring incidental catches of cetaceans in

fisheries (EU Council Regulation 2003/0163 [CNS]), dedicated sighting survey databases and the atlas of cetacean distribution in north-west European waters (Reid et al., 2003), ICES marine mammal by-catch records, and maybe national stranding network and incidental sightings databases.

References

Reid JB and Camphuysen C.J. (1998) The European Seabirds at Sea database. Biol. Cons. Fauna 102: 291. (http://www.jncc.gov.uk/marine/seabirds/seabird_sea.htm)
Reid JB, Evans PGH, and Northridge SP (2003) An atlas of cetacean distribution in north-west European waters http://www.jncc.gov.uk/Publications/cetaceanatlas/

2.5 Data archaeology – missing historic data

One important issue for quality control is the revision, updating, and completing of data already available in databases. Even in cases where ICES collates data, there appear to be gaps in the time series which should be filled. For example, there is a considerable amount of oceanographic data which is at present not available to national and international data centres. The reason for non-availability is manifold. These data are either "buried" in national institutes and labs, stored on manuscripts, ancient electronic storing devices, or "undetected" on computer hard disks. These data may still reside in national or international data centres and have not been treated in a timely manner due to e.g., shortage in personal. Every effort should be made to have these data available for the world scientific community. This will enhance data coverage and enable building of historic time series. Such attempts to complete existing data series should be coordinated within the IOC Global Oceanographic Data Archaeology and Rescue (GODAR) project.

3 DATA POLICIES AND DATA ACCESSIBILITY (TOR F)

In this section the Study Group reviews the data policies of ICES and other oceanographic and fisheries groups and makes recommendations. A recent article on data sharing stated:

> "Recent national and international multinational investments in networking and continued gains in information technological capability have given rise to a complex cyberinfrastructure that is rapidly increasing our ability to produce, manage, and use data. As research becomes increasingly global, data-intensive, and multi-faceted, it is imperative to address national and international data access and share issues systematically in a policy arena that transcends national jurisdictions. Open access to publicly funded data provides greater returns from the public investment in research, generates wealth through downstream commercialization of outputs, and provides decision-makers with facts needed to address complex, often transnational problems." [Arzberger *et al*., 2004, An international framework to promote access to data, Science, 303, 1777-1778.]

The same paper gave the following general operating principles for data access:

- Openness
- Transparency and active data dissemination
- Assignment and assumption of formal responsibility
- Technical and semantic interoperability of databases
- Quality control, data validation, authentication, and authorization
- Operational efficiency and flexibility
- Respect for intellectual property and other ethical and legal requirements
- Management accountability, including funding approaches

This article addressed the general issues of data access that are clearly applicable to the ICES situation.

3.1 Review of data ICES and other oceanographic groups data policies.

3.1.1 ICES Data Policies

[Note: the following material is taken from a three page internal ICES document]

ICES last reviewed its data policy in 1994 and prepared a base document for ICES data policy (ICES C.M. 1994/Del: 10). Since 1994 new policies have been added, but does not always provide information why access to data is regulated. The following information is taken directly from an internal document produced a few years ago.

3.1.1.1 Oceanographic data

3.1.1.1.1 The ROSCOP data

Provides details of research cruises conducted by ICES member countries. Additionally, it serves as a catalogue for the contents of the cm and hydro-chemistry databank.

No access restriction applies to this data set. All data, plus supporting software, are freely available by ftp and the Web.

3.1.1.1.2 The hydro-chemistry data

A collection of oceanographic profiles (CTD and Bottles), containing temperature / salinity / oxygen / nutrient / chlorophyll / pH / alkalinity / H_2S data, plus t/S surface data originating from the ICES ship route program. In addition to these data, time series stations from some coastal stations and from the North Atlantic Weather Ship stations are maintained. Most ICES oceanographic projects are also being maintained as a sub-set of these data. These include the Overflow experiments, Joint Skagerrak Expedition, the IBTS, BOSEX, JONSDAP, SKAGEX, NANSEN, and the Greenland Sea Project.

Very few data originators submit data to ICES without restrictions for subsequent re-distribution. As a matter of routine, ICES requests the approval of originators to release data collected within the past ten years. Most originators insist on clear acknowledgement of their data in any subsequent publication, including the opportunity to co-author, if deemed appropriate. This applies only in the case of raw data supplied for bona fide scientific use. Products are supplied, whether as a standard via the ICES web pages, or on demand, if for scientific use. In the case of a commercial requester, a standard sum of 2500 DKK is made, together with the above restrictions.

ICES reserves the right to refuse or hinder a request if the applicant, or his institute, is also a potential data supplier, but presently refrains from supplying or exchanging data with the ICES Secretariat.

3.1.1.2 Environment data

The following categories of data are available:

1. Contaminants in invertebrates, fish, marine birds, and mammals.
2. Contaminants in seawater.
3. Contaminants in sediments.
4. Data from certain types of measurements of the biological effects of contaminants.
5. Fish disease prevalence data.
6. Marine mammals by-catch data.
7. Quality assurance information.

ICES serves as the data centre for the Oslo and Paris Convention (OSPAR), under which a considerable proportion of the data on contaminants and biological effects are submitted. ICES is also the Thematic Data Centre for the marine component of the Arctic Monitoring and Assessment Programme (AMAP).

Data maintained for ICES or OSPAR are freely available to members of relevant ICES or OSPAR Working Groups. For scientists/institutions who are not members of these groups, there is a 2-year moratorium on the distribution of OSPAR data to allow time for them to be reviewed, assessed, and published before being made available to outside parties. The ICES Secretariat has over the years developed as practise that, when a request for raw data is received, the data originator is requested to give permission before any data are passed on to a third party. However, inventories of data holdings, available through the WWW or ftp, are freely available for everybody.

Raw data are made available as data extractions in a standard format. If more sophisticated products are requested, the Data Bank might charge the requester the costs of preparing such products.

3.1.1.3 Fisheries data

Concerning costs charged for data requests, ICES users and governments departments or contracting parties are given the data free of charge. All other users have to pay the costs of producing the data extraction.

3.1.1.3.1 Statlant

Official catch statistics for STATLANT area 27 provided by national statistical reporting offices. No restrictions apply to the release of these data.

3.1.1.3.2 IBTS data

Data from the International Bottom Trawl Survey carried out in the North Sea and Division IIIa. The data stored consist of the raw haul-by-haul data together with various levels of aggregation.

Without restrictions the data are available to all usage in connection with ICES working groups or research projects within the ICES work programme. For all other users there is an important distinction between raw data and aggregated data.

For raw haul data the following is a summary of the procedure: Applicants have to fill out a form indicating the data requested, their level of aggregation or disaggregation, the reasons why the request is made, the title and description of the project for which the data are to be used, for whom the project is conducted and particularly whether the project is done under contract. Once the form is filled in and signed by the applicant and the undertakings on the form effectively agreed upon by the applicant, it will be sent to the national contact person of the countries responsible for supplying the data. In order for matters to move smoothly and efficiently, deadlines for responses will be given. Objections or specific requirements, when arising, will be handled by referring the applicant to the country, which had objected.

For aggregated data down to the level of statistical rectangle (but without identification of the country or haul) the IBTS working group has suggested that the data should be in the public domain but that all requests should go through the national contact persons to secure proper use of the data and guidance of the user. Until this has been accepted, all requests are referred by the Secretariat to the national contact persons.

3.1.1.3.3 IFAP

Aggregated catch data by stock and some effort data by fleet required by the Assessment Working Groups have been stored at ICES in the IFAP system. This system was discontinued in 2001. Data were used without restrictions by assessment Working Groups and Working Group members working intersessionally. Most IFAP data are publicly available in tables and figures in the Working Group reports, and they are freely exchanged between all members of ICES Working Groups.

Beginning in 2004, working group reports are published immediately after a WG meeting (on the ICES website), and aggregated data are therefore unrestrictedly available to the public.

3.1.1.4 Summary

The various rules for data access are summarised in the following table.

	Raw data	**Data products**
ROSCOP	F	F
Hydro-chem <= 10 years	O/I	F/I
Hydro-chem> 10 years	F/I	F/I
Environment data	O	Inventories: F Other products: 0
Statlant	F	F
IBTS	ICES WGs + projects: F Others: O	ICES WGs + projects: F Others: O
IFAP	Assessment WGs + members: F Others: X	F
F: freely exchangeable **O**: approval from originator **I**: ICES decides **X**: no access		

*Note: IBTS is now included in the Datras database.

3.1.2 Other Data Policies

There are now many published data access policies for the myriad of oceanographic and fisheries programs. Some representative policy statements are presented here for comparison. The group recognizes that many of data access issues ICES faces are considerably more complicated than those programmes presented here.

3.1.2.1 Intergovernmental Oceanographic Commission (IOC)

The IOC Oceanographic Data Exchange Policy is given in the Annex to Draft Resolution IOC-XXII/DR.3. It is as follows [from IODE Website]:

> "Preamble
>
> The timely, free and unrestricted international exchange of oceanographic data is essential for the efficient acquisition, integration and use of ocean observations gathered by the countries of the world for a wide variety of purposes including the prediction of weather and climate, the operational forecasting of the marine environment, the preservation of life, the mitigation of human-induced changes in the marine and coastal environment, as well as for the advancement of scientific understanding that makes this possible.
>
> Recognising the vital importance of these purposes to all humankind and the role of IOC and its programmes in this regard, the Member States of the Intergovernmental Oceanographic Commission agree that the following clauses shall frame the IOC policy for the international exchange of oceanographic data and its associated metadata.
>
> Clause 1 - Member States shall provide timely, free and unrestricted access to all data, associated metadata and products generated under the auspices of IOC programmes.
>
> Clause 2 - Member States are encouraged to provide timely, free and unrestricted access to relevant data and associated metadata from non-IOC programmes that are essential for application to the preservation of life, beneficial public use and protection of the ocean environment, the forecasting of weather, the operational forecasting of the marine environment, the monitoring and modelling of climate and sustainable development in the marine environment.
>
> Clause 3 - Member States are encouraged to provide timely, free and unrestricted access to oceanographic data and associated metadata, as referred to in Clauses 1 and 2 above, for non-commercial use by the research and education communities, provided that any products or results of such use shall be published in the open literature without delay or restriction.
>
> Clause 4 - With the objective of encouraging the participation of governmental and non-governmental marine data gathering bodies in international oceanographic data exchange and maximizing the contribution of oceanographic data from all sources, this Policy acknowledges the right of Member States and data originators to determine the terms of such exchange, in a manner consistent with international conventions, where applicable.
>
> Clause 5- Member States shall, to the best practicable degree, use data centres linked to IODE's NODC and WDC network as long-term repositories for oceanographic data and associated metadata. IOC programmes will co-operate with data contributors to ensure that data can be accepted into the appropriate systems and can meet quality requirements.
>
> Clause 6 - Member States shall enhance the capacity in developing countries to obtain and manage oceanographic data and information and assist them to benefit fully from the exchange of oceanographic data, associated metadata and products. This shall be achieved through the non-discriminatory transfer of technology and knowledge using appropriate means, including IOC's Training Education and Mutual Assistance (TEMA) programme and through other relevant IOC programs."

3.1.2.2 GLOBal ocean ECosystems dynamics (GLOBEC)

The basic principles of GLOBEC data policy is as follows [From GLOBEC website http://www.pml.ac.uk/globec/main.htm] :

- GLOBEC will use a decentralised data management and distribution system. The 'centralised' component will be a comprehensive inventory of GLOBEC and 'GLOBEC-like' metadata with 'pointers' to the data location and key contact persons.

- Data sharing and exchange are an essential component of GLOBEC interdisciplinary studies. GLOBEC encourages the sharing of models as well as of data. Investigators should be prepared to share their data, models and data products and should recognise the 'proprietorship' of such data ('rights to first publication/ authorship') acquired from other investigators. GLOBEC protocols for international data exchange should be followed. GLOBEC recognises that sensitive fishery stock assessment data may need special consideration as to its availability.

- GLOBEC investigators retain the primary responsibility for data quality control and assurance. Once metadata is submitted to the IPO, GLOBEC investigators should inform the IPO of updates, corrections and dataset 'versions' to ensure that the GLOBEC Metadata Inventory is accurate.

- Within-project (during the funded lifetime of GLOBEC programmes) data exchange, archiving, and cataloguing are an important component of each GLOBEC programme and should be addressed (in detail) within the implementation plan of each GLOBEC programme. Metadata should be sent to the IPO for inclusion in the GLOBEC Data Inventory within 6 months of data being collected or created. Each national and regional programme and contributors to these programmes will retain the data collected under its auspices and will be responsible for ensuring that the data are properly archived. The IPO will ensure that the wider GLOBEC community is aware of the data collected within the programme. Regional and National programmes should assist in programme co-ordination and in the provision of central services e.g. data management.

- Each GLOBEC programme should address the long term archival of observational, laboratory, and model data and data products in order to ensure a lasting contribution to marine science.

- This data policy is in addition to the requirements of the individual Data Policies of the National and Regional Programmes.

3.1.2.3 Joint Global Ocean Flux Study (JGOFS)

The parts of the JGOFS data policy related to access is as follows [from JGOFS web site http://usjgofs.whoi.edu/jgofs-data-policy.html]:

PROPRIETARY RIGHTS

While recognizing the legitimate rights of data originators to the first use of the data they collect, U.S. JGOFS also encourages the oceanographic community to use data collected by the program, and in particular, believes that data availability should be restricted only in exceptional cases. Data normally becomes publicly available for use without restriction two years after origination. It is expected that all users will properly acknowledge the source of the data.

Guideline: Basic Core Measurements will be available to U.S. JGOFS investigators without restriction following submission to the DMO. However, it is expected that published works will cite the originator and will be sent to him or her for prior review.

Guideline: All non-Core Measurements will be available to U.S. JGOFS investigators with restrictions requested by the PI, such as: can be viewed by all JGOFS PI's, view and use only with consent of originator, use only in collaboration with originator, etc.

Guideline: Data will be released to the public domain two years after the date of collection. Before then, all requests from investigators outside the JGOFS program will be referred to the PI unless the DM has been given explicit permission to distribute the data freely.

3.1.2.4 Global Ocean Observing System (GOOS)

The following is taken from Section 5.2 (“Data Sharing, Access and Release”) of the Global Ocean Observing System report on Data and Information Management Strategy and Plan June 2001 [http://ioc.unesco.org/goos/GOOSdm_final.pdf].

> One of the primary GOOS principles (D7) deals with managing, processing, and distribution of data. Specifically,
>
> > "In concert with the policies of IODE, IGOSS (now JCOMM) and GCOS, and following the data management plan for the World Weather Watch of the WMO, commitment is required by GOOS participants to establishing, maintaining, validating, making accessible, and distributing high quality, long term data meeting internationally agreed standards. Preservation of GOOS data is required in suitable archives following appropriate procedures and criteria for data acquisition and retention, and should include information about data holdings. Data should be processed to a level that is generally suitable for the generation of operational products and for research, and described in internationally accessible on-line computerized directories that can also be made available by other means. GOOS contributors are responsible for full, open and timely sharing and exchange of GOOS-relevant data and products for non-commercial activities. Exchange implies that donation by individual nations gains access to data from others as well as to products derived using all available data, such that the benefit of cooperation exceeds the cost."
>
> GOOS has also established guiding principles for data sharing.
>
> > a) The data obtained by GOOS will be the most useful to the most people if there are no periods of proprietary holding, nor any restrictions on to whom or when the data are disseminated.
> >
> > b) The second principle is that the quality must be assured, which may mean some delays for some of the data.
> >
> > c) The third principle is that national and personal interests and needs are part of the overall equation, so it may not be possible to apply the first principle totally in all cases.
>
> It may not always be possible to hold to the first principle. Some research data sets will not be part of GOOS, if the investigators involved are not able to release the data in a timely way for any reason. GOOS does not plan to "publish scientific papers" based on the GOOS data, but rather to see that end users such as coastal zone managers and ship routers can get the data and products they need, when they need them. GOOS needs for data, and a scientist's need for the data, may be quite compatible. GOOS may be able to obtain access to scientific data for compilation into products provided the data are not released in original form for a specified time.
>
> Both scientific and operational agencies are becoming more aware of the need to make data available quickly even if it must be released in a preliminary form. One of the objectives of the CLIVAR data management system is that CLIVAR data should be assembled and distributed swiftly and that all data, model output, data assimilations, and reanalyses should be made accessible to the entire climate community (free and open exchange).
>
> CEOS has adopted principles on satellite data exchange in support of operational environmental use for the public benefit, and for support of global change research. These principles recognize that environmental data should be made available to users who need them within time scales compatible with the requirements.
>
> Operational environmental use for the public benefit is defined as:
>
> - Use of data to provide a regular environmental service for the public benefit;
> - Carried out by public national or international Earth observation agencies, or other entities designated by governments or public authorities, to support public benefit mandate;
> - Examples include use of data to carry out a mandate of environmental observation and prediction or missions relating to environmental management or regulation.

The last point is of particular interest because it applies the term operational to activities with much longer time scales than are usual. Environmental management and regulation is a process that takes place over months to years rather than hours to months.

3.1.2.5 World Ocean Circulation Experiment (WOCE)

The following is directly from the WOCE data policy web page (http://www.nodc.noaa.gov/woce_V2/disk02/datashar.htm).

There is a fundamental trade-off in WOCE - on the one hand, the protection of the intellectual effort and time of originating investigators (those who plan an experiment, collect, calibrate, and process a data set to answer some questions about the ocean), and on the other hand the need to compare various data sets and data types to check their consistency, to better understand the ocean processes involved, and to see how well the numerical models describe the real ocean. The policy adopted by WOCE is a trade-off between these conflicting needs.

Any data collected as part of WOCE should be made publicly available no later than 2 years (the publication rights period) from collection, unless differing proprietary rights are specifically granted by the WOCE Scientific Steering Group (SSG) and funding agencies. However, in the case of hydrographic data submitted to the WHPO-DAC the data will not be made publicly available until the DQE process has been completed and a final data set submitted by the chief scientist to the WHPO. Prior to that time the data will be considered preliminary and can be obtained, if necessary, only from the chief scientist.

Collection is interpreted as the completion of the determination of the value of the particular parameter. Thus, for example, tritium/helium collection may not be complete for over a year after return to shore/laboratory. Individual WOCE programs (hydrography, surface velocity, etc.) require all participating investigators to submit (usually within a few months after collection) data collected as part of WOCE to the WHPO or other DAC for the purposes of quality control and data synthesis during the public rights period. In that case the recipient DAC may not redistribute such data, or a derivative containing most of the information unless specifically approved by the originating PI, and should use the data for the stated purpose only.

Originating investigators are strongly encouraged to share their data before the end of the publication rights period. The receiving investigator should not publish any paper during that period based predominantly on the received data, should co-author results with the originating investigator, and should not redistribute the data.

3.2 Summary

A review of the current policies regarding data within international organizations shows that openness is the common denominator. Organizations make every attempt to provide as much data as possible to anyone who needs it. Organizations have realized that there must occasionally be some restrictions and they have accommodated them. However, it seems clear that organizations must try to be as open as possible.

4 REVIEW OF INTEGRATED DATA SYSTEMS ALREADY AVAILABLE (TOR C)

A series of presentations were given by members of the Group describing various integrated systems, both centralised and distributed, already developed or under development. Summaries of these are given below together with contact details for further information.

4.1 The British Oceanographic Data Centre (BODC) integrated databases

BODC acts as the UK's national oceanographic data centre, with part of its remit to provide data management support for UK marine science projects. These projects are primarily multi-disciplinary in nature, and over the past 15 years this has led to the development of integrated databases to handle the many different data types collected (e.g. navigation, CTD/rosette stations, net hauls, thermosalinograph, air samples, nutrients, sediment traps, ADCP, current meters, towed undulators, cores, drifting buoys, bottom landers, plankton recorders, production experiments and underwater cameras). BODC assembles the data from these large, multidisciplinary oceanographic research projects into a single high quality coherent data set, maintaining spatial and temporal relationships and ensures full documentation of data sets and the final banking and publication of the project data set (on CD-ROM). Work is underway to develop a web-based user interface to the data stored in these databases.

The concept of 'data event' is central to BODC's database design and covers everything collected on each cruise. An 'event' may be as simple as opening a tap to collect a sample of pumped sea surface water or more substantial activities such as: drifting buoy, bottom lander, SeaSoar tow segment, CTD cast, instrument mooring, or net haul. Each data event given a line entry in the main Oracle database table: date/time, geographic position, gear code, water depth, and various other identifiers.

The BODC parameter dictionary is also central to the integrated database. Each parameter is described with an 8-byte code; the first 4 bytes provide a coarse resolution parameter description, and defines the units, absent data values and valid data range; the remaining 4 bytes providing more detailed information.

BODC is currently undertaking a project (Enabling Parameter Discovery (EnParDis) to enhance dictionary. This will involve cleaning and tidying the Dictionary, expanding and clarifying entry descriptions, extending the range of parameters covered and enhancing parameter access through improved dictionary classification. BODC is also collaborating with other organisations to map their dictionaries to the BODC dictionary. At present some of the BODC parameter codes are used with some of the more exotic parameters included in the oceanographic database at ICES, but it would be useful to extend this and carry out a mapping or comparison of the parameters in the environmental databases.

For further information contact: Lesley Rickards (E-mail: ljr@bodc.ac.uk) or Roy Lowry (rkl@bodc.ac.uk), and see the BODC web-site (www.bodc.ac.uk)

Jones, M.T. 1996. All at sea with the British Oceanographic Data Centre. NERC NEWS Winter 1996.

4.2 The EU SeaSearch-II Project

Oceanographic and marine data, information and knowledge are important resources to a wide range of users in government, scientific community, and industry. The SeaSearch-II project seeks to bring together national oceanographic data centres in Europe to improve the data management infrastructure, and builds on the achievements of the earlier SeaSearch-I project.

The aims of SeaSearch I included providing users with a central overview and better access to ocean and marine data and information in Europe, to exchanging experience, and to cooperating in development, promotion and implementation of data and information management practices and methods. It also fosters dissemination of these practices to other institutes and networks. Initial work during SeaSearch I concentrated on building up metadata catalogues, primarily the European Directory of Marine Environmental Data (EDMED) – a discovery metadata directory – and the Cruise Summary Report (CSR) database. EDMED is a similar directory to the NASA Global Change Master Directory (GCMD), and at the end of SeaSearch-I comprised over 2800 data set descriptions from over 570 organisations in 12 countries and 2 international organisations (including ICES).

SeaSearch-II began in late 2002, is building on the experience of SeaSearch-I, and includes partners from over 30 countries. The new partners are building up their EDMED and CSR entries. Taking into account the fast growth of the internet, a change in attitude to communication and making data available, it is apparent that users are demanding more transparent insight in existing data resources scattered over various locations and more efficient ways of accessing and manipulating the data or of requesting data in case of conditional access. Thus SeaSearch-II is adopting new technologies (e.g., XML, high bandwidth communication) and standards (e.g. ISO19115, OpenGIS). One development is the Common Data Index (CDI), which will link the data held at (initially a subset of) the centres involved in the project. The CDI will comprise an index of data held at partner data centres, which will be searchable via a web-enabled GIS interface. The master index will be updated automatically at regular intervals. Agreed common codes for parameters, instruments, and platforms are being used. Searches of the CDI will point to online data, further information held on-line, or e-mail a request for data. The CDI should be available during the second part of 2004.

As a follow on from SeaSearch-II, a proposal has been submitted, SeaDataNet, to the EU. This proposed project will be a design study investigating on-line integrated data access to distributed heterogeneous systems, and will undertake a pilot project amongst selected project participants. The main components of the project are technical development, standards and communication protocols, quality control/quality assurance, and a distributed data system. There are some parallels to US IOOS DMACs and if the project is funded, appropriate links will be made.

For further information contact: Lesley Rickards (E-mail: ljr@bodc.ac.uk) or visit the SeaSearch web-site www.sea-search.net.

4.3 The WOCE Experience

WOCE completed the task that it set for itself of the management and delivery of high quality data to the oceanographic research community. The task has evolved from establishing standards for data (for example CTD measurements, analysis of water samples, tracers, nutrients), the development of better processing methods (e.g., chemistry and XBTs), and the quality control of the data by scientists with expertise to a role more focused on the delivery of uniformly formatted and described data to the web, CD ROMS and DVD's. These products have involved greater integration of the different data streams by increasing the standardization and consistency of naming conventions across data sets, through the use of self-describing data formats, development of tools capable of searching over spacio-temporal information and variables, the delivery of data directly to applications such as Matlab, Ferret, and standard programming languages and products for quickly viewing WOCE data such as eWOCE (Schlitzer, 2002).

These outcomes would have not been possible if the WOCE planners had not made the strategic decisions that encouraged resources to be allocated towards data management including tracking of the observations, and without the willing and active participation of science users as well as scientists at the WOCE Data Assembly Centres.

The WOCE Data Resource consists of 18 different elements distributed across the globe managing and quality controlling the 12 different observational data streams. The Data Information Unit tracked the progress of the WOCE field programmes and gathered the necessary field information to ensure that the Data Assembly Centres (DACs) were receiving data and were aware of the data sets were being collected by WOCE investigators. The DACs were mainly divided by instrument type and always had a close association with scientists using these data streams as active research users. The 12 data streams consisted of the hydrographic programme, surface drifters, upper ocean thermal data, sea-level, subsurface floats, moored measurements, surface meteorology/air-sea flux, surface salinity, satellite altimetry and sea-surface temperature (which was later expanded to include satellite surface winds), bathymetry, and shipboard acoustic doppler current profiler (Lindstrom and Legler, 2001).

The WOCE Data Resource is not perfect. There are overheads created by being distributed, which with the world wide web and emerging new technologies have been or will be largely overcome. The distributed data system with each data assembly centre representing a single data type can also have particular problems. For example, if one wanted to combine ADCP data, hydrographic data, and surface meteorology data from a single WOCE hydrographic section then the data would have had to be retrieved from three different DACs and merged by the individual researcher. For some data types, aggregating the WOCE data streams to enable joint analysis presents special problems as the relationship between the different variables can be obscured and even lost through the distributed system. WOCE allowed for this by the use of unambiguous experiment codes that were assigned to all the data collected on a single voyage. On balance, for WOCE, the distributed data system based on data type has been advantageous to research using only a single data type (for example hydrography), but does present disadvantages to those trying to integrate the data types by individual researchers.

For further information about the WOCE data etc, go to: http://woce.nodc.noaa.gov/wdiu/

References

Bindoff, N. and Legler, D. The WOCE global data resource: Lessons for CLIVAR data and data requirements. CLIVAR. Exchanges 2003 (6 pages). [http://www.clivar.org/publications/exchanges/ex26/pdf/s26_bind.pdf.]

Lindstrom, E.J., and D. M. Legler, 2001: Developing the WOCE Global Data System. In: Siedler G., J. A. Church, and J. Gould, (Eds.): *Ocean Circulation and Climate*. Academic Press, International Geophysics Series, Volume 77, 715pp.

Schlitzer, R., 2002: eWOCE, Electronic Atlas of WOCE data, www.awi-bremerhaven.de/GEO/eWOCE

4.4 The World Ocean Database (WOD) Project

The *World Ocean Database* (WOD) project is a global database of historical and current oceanographic data maintained by the U.S. National Oceanographic Data Center (NODC) in cooperation with the World Data Center (WDC) and IOC Global Oceanographic Data Archaeology & Rescue (GODAR) project. Its current content of over seven million stations contains samples from the early 1800's to the present (with the majority of the data covering from the 1950's to present). The WOD is an integrated database, containing co-sampled and co-stored physical, chemical (nutrients), and plankton data (chlorophyll, plankton composition, abundance, and biomass data).

One of the lessons learned from the WOD effort is the risk of duplicate data submissions through different pathways (for example, data first received from the original investigator, also received on a BODC project CD-ROM, and then

later received from the ICES Data Centre). These duplicates are not obvious, as corrections applied to some versions of the data (and not others) and minor formatting differences may cause the replicate data to be slightly different and having position differences up to 0.5 degrees or data differences up to one day. ICES would benefit from discussing duplicate detection and visualization techniques used by the WOD project.

The WOD project also has 8+ years of experience in working with plankton data (including metadata categories, quality control, and product generation). Todd O'Brien, developer of the plankton component of the WOD database, is currently working with ICES database programmers to review their metadata and database structure. Todd also suggested that the best test of the new ICES database plankton capability would be to start loading plankton data into the system, perhaps starting with the relevant WOD data from the ICES regions.

Contact Info: Todd O'Brien (Todd.OBrien@noaa.gov)

References

O'Brien, T. D., M. E. Conkright, T. P. Boyer, C. Stephens, J. I. Antonov, R. A. Locarnini, H. E. Garcia, 2002: World Ocean Atlas 2001, Volume 5: Plankton., NOAA Atlas NESDIS 53, U.S. Government Printing Office, Wash., D.C., 95 pp., CD-ROMs.

O'Brien, T.D., M. E. Conkright, T. P. Boyer, J. I. Antonov, O. K. Baranova, H. E. Garcia, R. Gelfeld, D. Johnson, R. A. Locarnini, P. P. Murphy, I. Smolyar, C. Stephens, 2002: NOAA Atlas NESDIS 48 WORLD OCEAN DATABASE 2001 Volume 7: Temporal Distribution of Chlorophyll and Plankton Data., U.S. Gov. Printing Office, Wash., D.C., 219pp.

4.5 The JGOFS/GLOBEC Data Management Open Source Software

The JGOFS/GLOBEC software is a distributed, object-based data management system that was developed for multidisciplinary, multi-institutional programs. It enables the exchange and synthesis of extremely diverse and widely spread data sets and it provides the capability for scientists to work with the data without regard for the storage format or for the actual location where the data resides. The JGOFS/GLOBEC software uses an object oriented approach by tying a method (software program) and the actual data together. While similar data types can share the same (or similar) software, the entity that the system deals with is the data object, which includes the necessary software to access and manipulate the data. Furthermore, since the system uses standard http protocol to communicate between multiple servers and multiple clients, a user is able to take advantage of the implicit and explicit tools and techniques available for displaying information via a Web browser. For example, when a user requests to see the available data, the table is generated in real time, based on whatever data are currently available. There is no static version of a data directory; rather the Web's ability to run a common gateway interface (CGI) script is used to generate the directory list on demand. The directory listing is created with the built in links so that the user can click on a data object (to look at the data) or click to another data category to see other available data objects.

Once a data object is chosen, by clicking on the object name, the actual data available are displayed, typically organized in a hierarchical fashion, with the slowest varying variables listed first. The data are formatted for display on the browser each time a request is made so that even if the data are changed, the display is always up to date.

The system is flexible and extensible, able to handle ASCII tables, binary data, and even images and movies. This is due to the inherent object based approach and the reliance on standard Web protocols. The data's method (computer program) is specific to the data. These programs are typically written in C, but any computer language is supported. These programs accommodate how the data are formatted (binary, ASCII, etc.). The programming task is simplified by requiring only the read routines be specific to the data. These read routines are used by standard, pre-existing data manipulation and reformatting code, which takes care of converting the data into a form that the rest of the JGOFS software can use.

Image and movie data are accommodated by creating "data files" of URL links to these data sets and serving these data files as if they were "regular" ASCII data. This allows objects composed of images and movies to be treated as all the other JGOFS data objects. The only requirement is that the user's browser support the image format (e.g. gif, tiff, jpeg, etc.) or movie format (e.g., mov, quicktime, fli/flc, etc.) that is used in the object. Further power and flexibility is gained because of the ability of the remote data servers to reformat data as necessary. For example, satellite images are served as gif images to the browser. However, these images are stored in an alternate format and may also be compressed. When a user requests to see one of these images, the original image is uncompressed (if necessary) and then reformatted as a gif image before being sent to the user's browser. This happens automatically, without the user aware of these steps. By serving a contributor's original data, from their own system, the necessity of duplicating data are avoided and, more importantly, the data viewed by others is always the most current version.

In addition to serving data, the JGOFS/GLOBEC software can also use the Web to distribute information including copies of meeting reports, cruise reports, the protocol documentation, and other information that is related to the data. Many of these pages are more or less static, such as the thesaurus documentation, but other pages are generated at the time of the request by reading the inventory data (also stored in the data management system as an object) and generating the appropriate Web formatted page.

For more information, the contact person is:

Robert C. Groman
Swift House MS #38
Woods Hole Oceanographic Institution
Woods Hole, MA 02543-1127
email: rgroman@whoi.edu
voice: (508) 289-2409
FAX: (508) 457-2169

JGOFS/GLOBEC software: The US GLOBEC version of the JGOFS software is available here:

http://globec.whoi.edu/globec-dir/software_tools.html

Download the last available tar file from: ftp://globec.whoi.edu/pub/software/JGOFS_GLOBEC

References

Groman, R.C. and P.H. Wiebe.1997. Management of biological, physical, and chemical data within the GLOBEC project. In " Proceeding of the International Workshop on Oceanographic Biological and Chemical Data Management", NOAA Technical Report NESDIS 87: 79-88.

Groman, R.C., and P.H. Wiebe. 1997. The U.S. GLOBEC Georges Bank study data and information management systems. 1997 ICES ASC Handbook (ICES CM 1997/T:10). Pg 154.

Groman, R.C., and P.H. Wiebe. 1998. Data management in the U.S. GLOBEC Georges Bank Program. 1998. Ocean Community Conference‘98 Proceedings, Marine Technology Society, Baltimore Md. pp. 807-812.

4.6 The Ocean Biogeographic Information System (OBIS)

The Ocean Biogeographic Information System (OBIS), a part of the Census of Marine Life (CoML), provides global geo-referenced information on marine species. It is a web-based portal for a number of species and habitat level databases that provides a variety of spatial query tools for visualizing relationships among species and their environment. The proximate goal of OBIS is to assess and integrate biological, physical, and chemical oceanographic data from multiple sources. The ultimate goal is to “... assess and explain the diversity, distribution, and abundance of life in the oceans - past, present, and future.” The OBIS Portal provides open access to data, information infrastructure, and informatics tools - maps, visualizations, and models.

The OBIS portal currently provides links to 18 databases divided by topic areas:

<u>Taxon</u>

Biogeoinformatics of Hexacorals; Biotic Database of Indo-Pacific Marine Mollusks; CephBase; FishBase; FishNet; History of Marine Animal Populations; Integrated Taxonomic Information System; OBIS-SEAMAP; Species2000; ZooGene;

<u>Habitat and Taxon</u>

SeamountsOnline; The Huntsman Marine Science Centre; ChEss - Biogeography of Chemosynthetic Ecosystems; SOC Atlantic Bathypelagic Biota

<u>Regional</u>

Gulf of Maine Biogeographic Information System

Time-Series

Zooplankton and Micronekton Species Time-Series in the Subtropical Atlantic

National DataBases

National Oceanic Data Center; Maritimes Region Homepage (Canada)

OBIS is becoming integrated into the larger data system for worldwide biological data (GBIF) and intends to become the major component for ocean biogeography and systematics. GBIF is under the sponsorship of the Organization for Economic Co-operation and Development (OECD).

A system of quality control existing in at least one of the OBIS components, FishBase, may be noteworthy in this context. Data entered into FishBase is largely provided through experts that voluntarily participate in this initiative and are rewarded by the acknowledgement of their work only. All data entered into the database is subsequently checked by another expert/collaborator and credit is given to all collaborators that enter, provide or check the data. Quality control in FishBase is therefore in part ensured by the user who is actively encouraged to provide feedback when accessing the data by a.) using user-friendly interfaces for comments and corrections and b.) providing an incentive for collaboration through clearly defined system of acknowledgements (see also: http://www.fishbase.org/manual/fishbasehow_to_become_a_fishbase_collabo.htm and http://www.fishbase.org/manual/fishbasebugs_blanks_and_errors.htm).

Contacts

The Portal is under the direction of Dr. Frederick Grassle and Dr. Yunqing Zhang (system development) at Rutgers, the State University of New Jersey [phoebe@imcs.rutgers.edu].

References

Grassle, J. F. 2000.The Ocean Biogeographic Information System (OBIS): an on-line, worldwide atlas for accessing, modeling and mapping marine biological data in a multidimensional geographic context. Oceanography. 13 (3): pp. 5-7.

Tsontos, V. M., Kiefer, D. A. 2003. The Gulf of Maine biogeographical information system project: developing a spatial data management framework in support of OBIS. Oceanologica Acta. 25 (5): pp. 199-206.

Zhang, Y., Grassle, J. F. 2003. A portal for the Ocean Biogeographic Information System. Oceanologica Acta. 25 (5): pp. 193-197.

4.7 The National Virtual Ocean Data System – NVODS

The National Virtual Ocean Data System (NVODS) is a framework for the distribution and analysis of oceanographic (and other) data from researchers and other data providers to intermediate and end users, including other researchers, government and commercial managers and planners, educators, students at all levels, and the general public. The goal is to provide an Integrated Regional, National and International Data System for Oceanography that allows end users, whoever they may be, to access immediately whatever data they may require in a form they can use, using applications they already possess and are familiar with.

The NVODS was started in 2000 with support from the U.S. National Ocean Partnership Program. NVODS uses the OPeNDAP (Open-source Project for a Network Data Access Protocol) Data Access Protocol as the primary mechanism for moving data from provider to user. This protocol provides a discipline-neutral means of requesting and providing data across the World Wide Web and will be a fundamental component of systems that provide machine-to-machine interoperability with semantic meaning in a highly distributed environment of heterogeneous datasets.

A large number of data access routes to NVODS data have already been created. These are listed at: http://www.po.gso.uri.edu/tracking/vodhub/vodhubhome.html. More than 48 sites provide OPeNDAP/DODS Datasets.

Contacts

The NVODS site is maintained by Paul Hemenway and Eben Oldmixon at the University of Rhode Island, Graduate School of Oceanography.

Contact at: contact@nvods.org

Technical Support: Technical questions concerning the software are handled by the DODS support staff at: technicalsupport@nvods.org

References

Hemenway, P., P. Cornillon, and T. Sgouros. 2003. Open Source Project for a Network Data Access Protocol Technical Working Conference 2003 Boulder, Colorado. 19-21 March 2003.

http://www.po.gso.uri.edu/tracking/meetings/tech03/report-html/report.html

4.8 Integration of oceanographic, fisheries, and biological data at the Fundación AZTI

The "Itxasgis system" is a cooperative system to gather all marine data originated at the AZTI Foundation, conceived to give access to the multidisciplinary database to all PIs and to facilitate integrated analysis.

The system is structured as a central data storage disc, in which are stored in an ordered manner (hierarchical directories system) all the data files produced. These files have to meet some minimum format requirements regarding the file formats (shape files, ascii txt, jpg, and tiff) and the format for the reference fields (Lat, Long, Depth, Date, and Time) before being stored.

Stored data belong to 6 defined categories: "Land Cartography" (which gathers coastal land cartographic series at different resolutions and dates, "Bathymetric data" (which gathers different bathymetric data from different sources and resolutions), "Biological and Fisheries data" (which involves data from benthos sampling, fisheries landings, icthyoplankton surveys, zooplankton distribution data, etc), "Hydrography and oceanography data" (CTDs, ADCP, remote sensing products like SST, etc), "meteorological data" and "Uses and management" (ICES subdivisions map, marine reserves, submarine pipes, etc). Integration of most of these data into the system is still under progress.

Access to data is done through an application running under ARCVIEW GIS desktop. This application allows the user to query the database and get the data meeting his requirements (Geographical extent, date ranges and depth ranges). The application then searches into the database, and provides a copy of the records within the user specifications.

This system has advantages regarding its simplicity to use either for data administrators or for users, its flexibility (new data and formats can be easily fed into the system) and finally, output data are provided in very standard formats so that users can easily exploit them with nearly any common software tools.

However, its main disadvantage is that it relies on a commercial software package running with licenses which limits the number of users, it is not platform independent, and problems have arisen with the version updates of the software. These disadvantages could be overcome with an INTERNET platform independent interface applications.

For further information contact: Yolanda Sagarminago, AZTI, Spain (see Annex 1 for contact details).

4.9 Management of Environmental Data in the German Marine Environmental Data Base

The German Oceanographic Data Centre (DOD), founded in 1967, started handling marine environmental data early in the late 1970s with the implementations of the Conventions of Helsinki (HELCOM – Baltic Sea) and Oslo-Paris (OSPARCOM – North Sea and Northeast Atlantic) which were both signed by Germany. Within their Monitoring Programmes various types of data had to be obtained: besides the "classical" physical and hydrographical data. Hydrochemical and biological information became very important at that time. The Conventions asked for regular assessments of the state of the seas for which data evaluations are a prerequisite. Already at that stage, DOD was aware of the need of a tool for an integrated management of those different types of information. After a period of brainstorming and negotiations, the "Marine Environmental Data Base (MUDAB)" of Germany was born. It was baptized in 1987 by virtue of a contract between the Federal Maritime and Hyrdrographic Agency (BSH) and the Federal Environmental Agency (UBA) which defined "MUDAB" as a joint project.

MUDAB is an in-house product of BSH and started as a hierarchical system. The "heart" of the database was built by the so-called "ROSCOPs" (Reports of Observations/Samples collected by Oceanographic Programmes – nowadays called "CSRs" – Cruise Summary Reports). They provide meta information on monitoring and research cruises such as

vessel/platform used, scientists involved, geographical area, types of investigations made on board, methods used etc. The forms have to be submitted to DOD no later than two weeks after a cruise – they form the basis for linking the relevant data with the meta information. Some years later MUDAB was changed to a relational system, those days migrated from the Data Base Management System "OPEN INGRES" to "ORACLE".

To date MUDAB stores information on 6388 cruises of 283 vessels/platforms with data on 700 different variables. Information is grouped either related to disciplines: physics, chemistry, biology and geography (with sub-groupings, see Table 4.1) or related to the medium in which the data were sampled: atmosphere, biosphere, hydrosphere, lithosphere, pore water, suspended solids.

Table 4.1. Groups of variables in "MUDAB", related to disciplines (alphabetical order)

Biochemical	Oil
Biometric	Organochlorines
Chlorobenzenes	PAH
DDT	PCB
Dioxines	Pesticides
Furans	Phenols/Chlorophenols
General chemical	Physical
Grain size	Radionuclides
HCH	Solvents
Metals/Elements	Taxonomic
Meteorological	TBT
Nitro compounds	Toxaphenes
Nutrients	

All data are stored together with their relevant information on methods and quality assurance. Before loaded to the database, detailed checks are performed on the original data - intensive communications with data originators form an essential part of the work. National taxonomic species lists were developed for proper management of biological data.

MUDAB is subject to continuous development in order to meet current requirements for the up-to-date management of marine data. Besides the integration of new types of information, the optimisation of functionality of the database and retrieval and export tools plays an important role in data management (e.g., GIS-interface). In parallel the amount of data is growing steadily. The data base stores actual as well as historical data. (The oldest data date back to 1873.)

Retrieval of data is performed on a network PC - either guided by a menu or by using SQL. As starting point for a data request, one can choose the database or the GIS-interface. There are several levels of retrieval:

1. Meta information via internet: http://www.bsh.de/ - Marine data - Observations - DOD data centre: (http://www.bsh.de/en/Marine%20data/Observations/DOD%20Data%20Centre/index.jsp)
2. Meta information in MUDAB: Cruise inventories, data submissions, core information (descriptions e.g. for platforms, stations, institutions, methods etc.)
3. Data/values

A variety of export forms is offered:

Alphanumeric: Tables, basic statistics, text information, formatted data (e.g. for submission to ICES)

Graphical: Cruise plots, charts, graphs

Users of MUDAB recruit from different groups: National and foreign scientists, German Ministries and Federal Agencies, national and international bodies as well as environmentally interested citizens and groups.

For the future on-line access to released data is planned. By implementation of improved technologies a 50% reduction of personnel will be achieved.

For further information contact: Sunhild Wilhelms, Federal Maritime and Hydrographic Agency, Germany.

4.10 Sea Around Us web-based database of global fisheries catches and species distributions (www.seaaroundus.org)

The Sea Around Us Project is based at the Fisheries Centre, University of British Columbia, Vancouver, Canada and co-funded through Philadelphia's Pew Charitable Trusts (PI: Daniel Pauly). The aims of the project are to provide an integrated analysis of the impacts of fisheries on marine ecosystems on a global scale using spatial modelling and mapping techniques, and to devise policies that can mitigate and reverse harmful trends whilst ensuring the social and economic benefits of sustainable fisheries. One of the main outlets for products from this project is a web-based database combined with a GIS mapserver and other tools to generate a multitude of maps, tables and graphs summarizing/visualizing database contents and modelling products on the fly or to provide pre-canned analysis products. The database contains, among other things, time series maps of global catches, generated through rule-based spatial and taxonomic disaggregation of available fisheries statistics (FAO data, supplemented by national databases and ancillary data) (Watson *et al.* 2004 (in press)). In addition, predictions of patterns of species distribution, including more than a 1500 commercial fish species and higher taxonomic groups as well as crustacean and marine mammal species (with other groups forthcoming) can be viewed online. Distributions were produced using spatial habitat suitability modeling approaches (e.g. (Kaschner *et al.* in prep)). All results are presented in a global raster with 0.5 degree longitude by 0.5 degree latitude cell dimensions and modeling input parameters are also posted online and feedback from species experts regarding input settings and outputs is strongly encourage to continue to improve maps. One of the main strength of the web interface lies in the many ways catch and distribution information can be queried and viewed as well as the user-friendly navigational interface. Besides the available global maps, information can be obtained by country EEZ's, FAO statistical areas and Large Marine Ecosystems. Query outputs may include lists or graphical representations of all species occurring or caught in specific areas or countries with fishing access.

For further information contact: Kristin Kaschner, University of British Columbia, Vancouver, Canada.

References

Kaschner K, Watson R, Trites AW, Pauly D (in prep) Mapping world-wide distributions of marine mammals using a Relative Environmental Suitability (RES) model.

Watson R, Kitchingman A, Gelchu A, Pauly D (2004 (in press)) Mapping global fisheries: sharpening our focus. Fish and Fisheries.

5 STRATEGIES AND TECHNICAL SOLUTIONS FOR DATA INTEGRATION IN THE ICES ENVIRONMENT (TOR D)

5.1 Distributed vs. centralised databases (including a review of existing tools and formats/protocols useful for setting up integrated databases)

The Study Group discussed the benefits of "centralized" vs. "distributed" database models. A *centralized* database is where all data are physically located at and served from a single location, as is the current ICES database. A *distributed* database is where the data can be located at various geographically distributed nodes (e.g., multiple institutions and/or local data centres) and may be accessible from a single interface. The management issues and attributes of each model are summarized in Table 5.1.

The study group recognized that each model had its own advantages and disadvantages, and did not indicate either as the ideal choice for ICES. The group did come to consensus that some applications favoured the centralized model (e.g., large, data-heavy hydrographic databases) whereas others favoured the distributed model (e.g., complex biological data types that are better kept closer to the expertise that originally collected it). The two models do not have to be exclusive. For example, ICES could ideally have a single centralized database for certain data types and distributed sub-databases which serve the special needs or data of local institutions. The SGMID thought that the distributed systems might be developed faster and for less expense than centralised systems.

Table 5.1: Issues and Attributes of Centralized and Distributed Database Systems.

Management Issues / Attributes	**Centralized Model**	**Distributed Model**
Computer & Technical Issues	Centralized hardware and computer staff *(better support and hardware for less or same money, +)* Easier to keep entire (single) data system online and to provide fast internet access **(+)**	Each node requires capable hardware and staff (which might be present already) *(quality of both may differ between nodes, -)* Rural nodes may suffer poor internet access and/or power losses (-) *Note: A centralized data archive is still required for long term data archiving.*
Data Access & Re-Formating[1]	Reformatting and indexing effort done at initial load *(quicker access and response, +)*	Data are reformatted and/or indexed on the fly for each request, plus travel time from the nodes *(slower access and response, -)*
Quality Control	Data managers not as familiar with data and separated from its original collector (-) Corrections to original data require reloading into central database, slowing update (-) Easier to perform multi-data cross-variable statistical and quality control checks **(+)**	Data closest to the original investigator expertise **(+)** Corrections to local data are immediately available **(+)** Data may only be checked against themselves (-) Less risk of data duplication (+)
Apparent Data Ownership and/or Credit	Focus on centralized center may over-shadow (or omit) collecting institutes and parties (-) Investigator participation and submission interest may be lower (-)	Data stays local, offering better possibility of credit and ownership recognition **(+)** Possible better investigator participation & submission **(+)**

[1] Assuming that effort of reformatting to centralized database is approximately equivalent to programming an OPeNDAP protocol and/or minor reformatting required for the distributed database.

5.2 Quality control

There was agreement among participants that the scientific value of integrated datasets critically depends on the quality and precision of the data. Assuring the highest possible data quality during all stages of data collection, submission, integration, and dissemination is considered essential. As first step, data originators should make sure that data are carefully calibrated and should apply automated and visual data screening procedures in order to detect and identify outliers and systematic data errors. Systematic errors should be corrected immediately, outliers should be flagged rather than deleted from the dataset. At data centres, the integrity of individual datasets should be checked by trained data experts before incorporating the new data into the database. The integrity check should include visual and/or automated comparisons with historical data. In case of individual datasets served on a distributed data network, specifically assigned data quality experts should compare data from different laboratories and projects to identify possible calibration offsets and assure inter-laboratory consistency.

Consistency problems and calibration offsets detected by quality experts should be reported back to the data originator and corrections should be provided. Individual data values considered questionable or bad should be flagged instead of being removed from the dataset. This requires that both centralized and distributed data systems should support data

quality flags for individual data values. To detect possible data corruption occurring during data re-formatting and conversion, newly integrated datasets should be extracted and sent back to the data originator for cross-checking.

The study group realized that achieving and maintaining high data quality standards is not easy and involves tedious and labour intensive procedures. Good coordination between data originators, data centres, and data users is essential as are long-term resource commitments at all stages of the data flow. An easy-to-use feedback system to report errors is needed to encourage users to participate in data correction.

5.3 How to get data out of an integrated system

Centralized and distributed data servers must provide continuous, reliable access to data, facilitate the discovery of datasets, and provide the data in useful formats. The study group agreed that for the purpose of maximum scientific usefulness, the data should be made publicly available in their raw, unaggregated form on a timely manner. The revised ICES data policy should aim for a maximum of openness and for short proprietary periods. Timely access to data is a fundamental requirement for early detecting of environmental changes.

Integrated data systems serving large heterogeneous datasets need to have easy-to-navigate data catalogues and inventories to facilitate discovery and extraction of data by the user. These systems need to be web-based and available on a 24/7 basis. Information indexing technologies currently applied by general purpose search engines should be applied in the scientific context as well. Aggregation capabilities and tools for producing data holding summaries on-the-fly are required for distributed data systems. Present aggregation facilities are inadequate and there is a strong need for further development.

Data extraction from integrated data systems need to be automatic and response times should be as short as possible. Ideally, web-based, interactive graphical user interfaces should be used to allow easy specification of geographic domains, time periods and data types of interest. The Live Access Server (http://ferret.pmel.noaa.gov/Ferret/LAS/ferret_LAS.html) is an example of a GUI data retrieval system. Another interactive graphic user interface is MapServer, an open source development environment for building spatially enabled Internet applications [http://mapserver.gis.umn.edu/].

Users should be given a choice of output formats. The choice of output formats should be based on user demand, and formats of popular analysis and visualization software tools should be supported. This includes Ocean Data View (http://www.awi-bremerhaven.de(GEO/ODV), Matlab, commercial and open-source Geographic Information Systems (GIS) such as Open GIS Consortium, Inc. (OGC) [http://www.opengis.org/], and others. Further development of GUI analysis and visualization tools should be encouraged.

5.4 Open access/open source software vs commercial software, platform independency

The working group noted that there are several relevant GNU software packages available that might be useful in the construction of integrated databases. These include MySQL and PostgreSQL which can both be downloaded from the internet for free. They are available for Windows, Unix, and GNU Linux, and have a range of powerful features. So much so that they are both used in the commercial world, e.g., by banks and supermarkets. These databases can easily be linked to web browsers, and they can also be set up so they can be interrogated by other useful GNU software packages. 'R', 'Octave' and 'GRASS, for example, are free statistical and data-visualisation packages which are capable of querying MySQL and PostGreSQL databases directly from their command lines. Furthermore, open source programming languages such as PERL, Python, or PHP can be used as 'glue' languages allowing communication between database and visualisation softwares. The group noted, however, that such systems, although flexible, require considerable degrees of computing expertise, and more time is needed to learn how to use them than is probably the case for commercial solutions. ICES personnel who are ultimately responsible for building and maintaining the integrated database have opted to use Microsoft SQL server and it is probable that any integrated database would exploit that particular package. Nevertheless, this need not preclude the use of GNU software packages, which could be set up to 'talk' to the MS SQL server.

6 THE ROLE OF ICES IN PROVIDING INTEGRATED DATA – RECOMMENDATIONS (TOR E)

The Study Group on Management of Integrated Data (SGMID) has the following recommendations:

Data Policy: ICES should revisit its data policy (Review and Update), emphasizing that the ICES database provide users with maximum accessibility and openness.

Data Integration within ICES: ICES should continue work on data integration and explore the possibilities of using distributed database systems (eg. for survey data). It should define minimum standards and facilitate quality control. ICES should provide a comprehensive list of data (and linkages between data) available in the ICES environment (even if currently not handled by the secretariat) and promote discussion about possible interdisciplinary future data use and how to meet necessary data demands

Expanded Database Information: There is an urgent need for development of fisheries databases, integration of plankton, survey, and marine mammal/seabird data as a prerequisite for the development of an integrated advice. ICES should make data easily available that are already in the ICES system, but hidden in reports of working groups or elsewhere, and retrieve (and quality check) historic data. This requires people power – so ICES should emphasize the need for support to do this task from the member countries.

Improved access to data: ICES should encourage open access to data wherever possible. Data should be directly downloadable via the internet to reduce work for secretariat staff and there should be increased accessibility of ICES products like WG reports.

Data Integration - Expanded Role In Community: ICES should take the lead in exchange format definitions (e.g., species lists, nets, fleets….) to ease data integration, make codes available if they are already agreed. Also specified formats needed to smooth data interchange without losing flexibility. Lead institutions with large data holdings should participate in creating the flexible expanded formats. This requires increased coordination and communication between ICES working groups and the secretariat to facilitate development of standards. ICES should be the focal point for data integration through maintenance of a catalogue (metadata and inventories of existing data and information with appropriate roadmap for how different information connected – improve search capability to be able to find documents and data (for example, by providing a site map button on home page). ICES should explore possible data retrieval and visualisation tools to facilitate data and site navigation (GIS interface/ map server).

ICES should use as much as possible existing formats, software, etc, rather than re-inventing/redeveloping database structures or formats. ICES should encourage national laboratories to contribute to software/database development or to take the lead for development/maintenance of specific distributed databases. ICES should utilize as much as possible open source software that can be used on all computer platforms.

ICES needs to participate in different ongoing initiatives for data integration within the ICES community (for example the work of the steering group for quality assurance of biological measurements in the Northeast Atlantic (SGQAE) and Baltic Sea (SGQAB) and the steering group for quality assurance of chemical measurements (SGQAC) needs to be better communicated within the ICES community), in Europe (SeaSearch-II etc.), and elsewhere (GLOBEC, WOCE, OBIS etc.). ICES should coordinate with the IOC GE-BICH (Group of Experts on Biological and Chemical Data Management and Exchange Practises (GEBCDMEP) –contact person Edward Vanden Berghe).

Theme Session on Marine Integrated Data: There is a recognized need for enhanced integration of diverse kinds of fisheries, oceanographic, and other marine environmental data and the development of tools to enable fisheries and environmental assessments. We propose that a theme session take place at the Annual Meeting in 2005 (Aberdeen, Scotland) with the aim to encourage the presentation of papers/posters that address these issues. This theme session should be jointly sponsored by the WGMDM and SGMID.

7 PROPOSED TERMS OF REFERENCE FOR 2005 MEETING OF SGMID

The Study Group on Management of Integrated Data (co-chairs: Peter Wiebe, USA, and Christopher Zimmermann, Germany), should meet at a venue to be defined in April 2005 to:

a) Review the development within ICES towards integrated databases of oceanographic, environmental, and fisheries data;
b) Identify data sources relevant to a), above, not yet integrated into the ICES databases;
c) Review existing integrated data systems for fisheries/environmental data and review data integration work in existing projects inside and outside of ICES;
d) Propose strategies and technical solutions for integrating available data including the possibility that data are not physically located in one site;
e) Evaluate and recommend the level of integration and aggregation of data in connection with management issues from an ecosystem perspective including the use of GIS systems;
f) Evaluate problems associated with the accessibility of data.
g) Review the progress of integrating the ICES database system into networks like SeaSearch II.

Acknowledgements

The group expresses its gratitude to the staff of the host institute for their help with the network and support during coffee breaks. The latter were sponsored by the German “Bundesmarktverband der Fischwirtschaft”.

ANNEX 1: LIST OF SGMID PARTICIPANTS

(as per 02 Apr 2004; + = attending the 2004 Hamburg meeting)

+Dr. Larry Atkinson
Commonwealth Cent. for Coastal
Physical Oceanography
Dept. of Ocean, Earth & Atmospheric Sciences
Crittenton Hall - 768 W. 52nd St.
Old Domion University, Norfolk VA 23529
USA
atkinson@ccpo.odu.edu

+Dr. Doug Beare
Fisheries Research Services
Marine Laboratory
P.O. Box 101
Victoria Road
Aberdeen AB11 9DB
United Kingdom
d.beare@marlab.ac.uk

+Ole Folmer
Danmarks Fiskeriundersøgelser
Charlottenlund Slot
DK-2920 Charlottenlund
Denmark
ofo@dfu.min.dk

Stefan Garthe
Christian Albrechts-Universitat Zu Kiel
Inst.Fur Meereskunde
Dusternbrooker Weg 20, D-24105 Kiel
Germany
sgarthe@ifm.uni-kiel.de

+Jule Harries
CEFAS, Burnham Laboratory
Remembrance Avenue,
Burnham-on-Crouch, Essex CM0 8HA
United Kingdom
Tel: 01621 787234
Fax: 01621 784989
j.harries@cefas.co.uk

+Kristin Kaschner
Marine Mammal Research Unit
Fisheries Centre
The University of British Columbia
Vancouver, B.C. V6T 1Z4
Telephone: (604) 822-8181
kaschner@zoology.ubc.caR

Dr. Ciaran Kelly
Fisheries Science Services, Marine Institute
Galway Technology Park, Galway
Ireland
Tel: +353 91 730454
Mobile: +353 86 8058013
Fax: +353 91 730470
Ciaran.kelly@marine.ie

+Henrik Kjems-Nielsen
ICES Secretariat
Palægade 2
1261 Copenhagen
Denmark
henrikkn@ices.dk

Dr. Thomas Lang
Inst. f. Fischereiökologie
BFA f. Fischerei
Deichstrasse 12
27472 Cuxhaven
Germany
thomas.lang@ifo.bfa-fisch.de

Nancy McHugh
Northeast Fisheries Science Center
NMFS/NOAA
116 Water Street
Woods Hole, MA 02543-1026
USA
Nancy.McHugh@noaa.gov

Dr Georgiy Moiseenko
Information Systems Laboratory
Russian Federal Res. Inst. of Fisheries & Oceanography (VNIRO)
17, Verkhne Krasnoselskaya
Moscow 107140
Russia
georgem@vniro.ru

Dr. Savithri Narayanan
Marine Environmental Data
Services Branch
200 Kent Street, 12th floor
Ottawa, Ontario K1A 0E6
Canada
NarayananS@DFO-MPO.GC.CA

+Todd O'Brian
NOAA-NMFS-Science and Technology
Marine Ecosystems Division
SSMC III F/ST7
1315 East-West-Hwy
Silver Spring, Maryland, USA
Tel +1 301 713-2363
Fax +1 301 713-174
todd.obrian@noaa.gov

Brian Rackham
CEFAS, Lowestoft Laboratory
Lowestoft
Suffolk NR33 0HT
United Kingdom

+Dr. Lesley Rickards
Proudman Oceanograhic Laboratory
Bidston Observatory
Birkenhead,
Merseyside L43 7RA
United Kingdom
ljr@bodc.ac.uk

+Ms Yolanda Sagarminaga
AZTI
Herrera kaia, Portualde z/g
20110 Pasaia (Gipuzkoa)
Spain
ysagarminaga@pas.azti.es

+Dr. Reiner Schlitzer
Alfred-Wegener-Institut for Marine and Polar Research
Columbus Center
D-27568 Bremerhaven
Germany
rschlitzer@awi-bremerhaven.de

Dr. Ursula Siebert
Forschungs- und Technologiezentrum Westküste
Hafentörn
25761 Büsum
Germany
usiebert@ftz-west.uni-kiel.de

+Prof. Hein Rune Skjoldal
Institute of Marine Research
P.O. Box 1870 Nordnes
N-5817 Bergen
Norway
hein.rune.skjoldal@imr.no

+Manfred Stein
Bundesforschungsanstalt f. Fischerei
Institut für Seefischerei
Palmaille 9
D-22767 Hamburg
Germany
manfred.stein@ish.bfa-fisch.de

Aleksander Toompuu
Marine Systems Institute
Tallinn Technical University
Akadeemia tee 21
12618 Tallinn
Estonia
alex@phys.sea.ee

+Dr. Peter Wiebe
Woods Hole Oceanographic Institution
Woods Hole, MA 02543
USA
pwiebe@whoi.edu

+Sunild Wilhelms
Bundesamt f. Seeschifffahrt und Hydrographie
Bernhard-Nocht-Str. 78
D-20359 Hamburg
Germany
Sunild.wilhelms@bsh.de

+Dr. Werner Wosniok (part time)
Inst. für Statistik der Universität Bremen
Fachbereich 3
Postfach 330 440
D-28334 Bremen
Germany
wwosniok@math.uni-bremen.de

+Dr. Christopher Zimmermann
Bundesforschungsanstalt f. Fischerei
Institut für Seefischerei
Palmaille 9
D-22767 Hamburg
Germany
czimmermann@ish.bfa-fisch.de