Oceanography Committee **ICES CM 2002/C:11**

Report of the

Working Group on Marine Data Management

Helsinki, Finland
17–19 April 2002

International Council for the Exploration of the Sea

Conseil International pour l'Exploration de la Mer

Palægade 2–4 DK–1261 Copenhagen K Denmark

TABLE OF CONTENTS

1 OPENING OF THE MEETING

The meeting was opened by A. Isenor, Chair, at 9:00 am on 17 April, 2002, hosted by the Finnish Institute of Marine Research (FIMR), Helsinki, Finland. Participants were welcomed to the meeting by Professor Pentti Mälkki, President of ICES. R. Olsonen also welcomed participants and explained the local arrangements.

Members of the Working Group present were: P. Alenius (Finland), G. Dawson (UK), P. Ennet (Estonia), M. Fichaut (France), L. Fyrberg (Sweden), J. Gagnon (Canada), M. Garcia (Spain), R. Gelfeld (USA), A. Isenor (Chair, Canada), S. Jans (Belgium), F. Nast (Germany), R. Olsonen (Finland), L. Rickards (UK), S. Sagan (Poland), H. Sagen (Norway), G. Slesser (UK), J. Szaron (Sweden). The ICES Oceanographer, H. Dooley was also present. Various members of the IOC/IODE Group of Technical Experts on Data Exchange (GETADE) were also present including: G. Reed (IOC Consultant), D. Collins (USA) and E. Vanden Berghe (Belgium). Other participants included T. de Bruin (on behalf of MDM member N. Kaaijk, The Netherlands), T. Carval (France) and R. Hietala (Finland). Apologies for absence were received from S. Almeida (Portugal), S. Feistel (Germany), K. Medler (UK) and O. Ni Cheileachair (Ireland). A complete list of names, addresses and contact points of participants can be found in Annex 1.

2 ADOPTION OF THE AGENDA

The agenda (see Annex 2 for current Terms of Reference and last years Action Items) for the WGMDM (see Annex 3 for acronyms) meeting was adopted as a resolution of the Annual Science Meeting in Oslo, Norway (C.Res. 2001/2C12).

3 OCEANOGRAPHY COMMITTEE REVIEW

A. Isenor informed the WGMDM that J. Szaron had presented the Report of the Working Group on Marine Data Management to the Oceanography Committee at the Annual Science Conference in Oslo, Norway. The report was well received by the Committee. The external review was generally more positive than previous years.

The report from the Oceanography Committee (see also ICES Annual Report) and the full review comments are reproduced here in Annex 4. The MDM noted that the main comments from the review process could be summarised as dealing with:

1) Communication with other Study and Working Groups,
2) Operational Issues,
3) Promotional Issues.

Participants were asked to keep these three points in mind over the course of the meeting.

Some members of WGMDM expressed a concern that there is a lack of understanding by other ICES working groups towards the work carried out by the WGMDM. To address this problem it was suggested that the MDM sponsor a theme session at the ICES Annual Science Conference. The last time WGMDM made such a contribution to the conference was in 1994, to a Joint Session on Quality Assurance of Marine Measurements.

Some participants expressed an interest in relating the theme session to metadata. The importance of complete and coherent metadata cannot be over-stressed, as data is sometimes worthless without proper metadata. Furthermore, collectors often wish to conduct the collection exercise using their own methods and procedures. Proper documentation on these methods and procedures is sometimes difficult to obtain.

Although the participants expressed an interest in a theme session, no members were presently willing to commit sufficient time to such a task. Without direct member support, it would be difficult to sponsor a theme session.

4 PRESENTATIONS

Meeting participants described activities at their own data centre/laboratory over the past year and looked to developments in the future. Executive summaries of the presentations can be found in Annex 5.

5 DATA TYPE GUIDELINES (TOR 2)

A presentation by A. Isenor outlined the history of the initial and current guideline development.

The initial MDM guidelines were produced in the 1980s. This early development focused on the information and data required by the data centres. In the early 1990s further guidelines were developed, initially in collaboration with the WGOH and those guidelines previously developed were updated. These early guidelines were presented in many forms and often lacked inter-guideline consistency.

In 1999, the MDM re-examined the existing guidelines. A decision at the Ottawa (1999) meeting was made to rework the existing guidelines to a form compatible with a Data Centres perspective on the flow of data. The 1999 effort first established a framework for the guidelines, where each guideline would describe the data and metadata expected by the data centre, the processing within the data centre and the service delivered by the data centre to clients. Particular data types would be targeted and described with this framework. This data type targeting was considered more time consuming than general, all-encompassing guidelines. However, targeted data types were considered more client-friendly as someone collecting data would be able to obtain a guideline particular to their needs.

The 1999 meeting also resulted in an example application of the framework to the Seasoar/Batfish data type. During the 1999–2000 inter-sessional period, the MDM attempted to compile nine draft guidelines covering data types: CTD, water level, XBT, Seasoar/Batfish, shipboard ADCP, moored ADCP, moored current meter, surface underway, and a combined oxygen, nutrients, chlorophyll guideline.

At the Hamburg (2000) meeting, the MDM reviewed existing guidelines and suggested revisions. These revisions were primarily related to the inter-guideline consistency. These revisions were conducted during the 2000–2001 inter-sessional period. One guideline (oxygen, nutrients and chlorophyll) required considerable work and was sent back to the subgroup for revision. Two new guidelines were added this year (floats and drifting buoys).

At the UK (2001) meeting, the float guidelines (which evolved into the PALACE guidelines) and the drifting buoy guidelines were reviewed and revisions suggested. Again, these revisions were primarily related to the inter-guideline consistency. During a workshop-like session, the MDM revised the oxygen, nutrients and chlorophyll guidelines that were consequently renamed to the discrete water sample guideline.

Over the past year, the complete set of guidelines has undergone another review by members. The review indicated minor inconsistencies in the completed set of nine guidelines, with more substantial revisions in the remaining three guidelines. This review completed the remaining three guidelines on drifting buoy data, discrete water sample data and biological net tow data. These guidelines are presented in Annex 6.

The MDM then considered the future of the guidelines by focusing on five key questions:

- How will we advertise the guidelines?
- Would IOC/IODE be interested from the vantage point of training material?
- Log Sheets - Do we want to develop a set of deployment log sheets?
- New guidelines? - Mini logger
- Expanded role? - Detailed CTD Calibration (matching to water sample levels)

The MDM recognised the need to promote and advertise the guidelines. The Group decided that only one copy of the guidelines should be maintained that copy residing on the ICES web site.

In this context, the Group was reminded of a decision made some years ago to maintain a collection of “Information Sheets on Oceanographic Data Centres in the ICES area”. Originally, these sheets were maintained in hard copy and updated by the Service Hydrographique biennially. Several years ago the product became web-based and held centrally on the ICES web site. At the MDM meeting in Ottawa (1999) it was decided that maintaining such a collection centrally was somewhat clumsy and inhibited regular updating. Therefore, the decision was made to maintain the information at the individual data centres, with ICES linking to the particular data centre page. Unfortunately only MEDS (Canada) has responded to this decision. Consequently the members were asked to implement this decision in full so far as their data centre was concerned, as soon as possible. Ultimately this ICES web page should only contain the links to this collection. The catalogue is available from http://www.ices.dk/committe/occ/mdm/odcentre/

The Group noted that having web access to the guidelines automatically makes them available via on-line searches. For example, using the Google Search Engine and the strings “Marine Data”+”Guidelines”, results in the MDM data guidelines listed first on the results page. However, the MDM does recognise the need to promote the guidelines to other organisations concerned with marine science and data management.

The Group noted that communication had been established with the ICES/OSPAR Steering Group on Quality Assurance of Biological Measurements in the Northeast Atlantic (SGQAE). A summary of the discrete water sample guidelines has been provided to the SGQAE Chair, with the intention of providing the actual guideline when approved by the Oceanography Committee.

The following actions resulted from this discussion.

Action 1: ***After review by the Oceanography Committee, H. Dooley will ensure the Profiling Floats, Discrete Water Sampling and Net Tow guidelines are added to the Data Type Guidelines web page on the ICES web site.***

Action 2: ***All MDM members to ensure that the description of their data centres is up-to-date and presented in a structure similar to the Canadian page.***

Action 3: ***A. Isenor and G. Reed, GETADE Chair, will discuss links between IOC/IODE and ICES WGMDM guidelines.***

Action 4: ***All members will install links between their data centre web sites and the ICES data guidelines web page.***

Action 5: ***A. Isenor will notify JCOMM of the data type guidelines.***

Action 6: ***L. Rickards will inform the Argo data management team of the data type guidelines for float data.***

Action 7: ***M. Fichaut will inform the Co-Chairs of the IODE Sea Surface Salinity Data Pilot Project of the data type guidelines for Underway data.***

Action 8: ***A. Isenor will inform GE-BCDMEP (Group of Experts on Biological and Chemical Data Management and Exchange Practices) of the data type guidelines.***

Action 9: ***A. Isenor, G. Dawson and E. Vanden Berghe will discuss a possible guideline poster for the up-coming "Colour of Data" meeting to be held in Belgium.***

The MDM then considered application of the guidelines within the various data centres. An example was shown of the calibration data supplied for the current meter data on the VEINS CD-ROM. The guidelines do not specifically indicate how the data centre should supply the metadata in relation to the final datasets.

The MDM considered a member recommendation for a change in the wording of the current meter guidelines to address the issue of specific calibration documentation. However, the MDM noted that these documents were guidelines, and thus only represent suggestions. As such, the guidelines should not dictate the form of packaging of the metadata with the data. More strict rules on the content of packaging would move the “guidelines” towards “standards”.

Action 10: ***A. Isenor will inform K. Medler that the request for a change in the wording of the current meter guidelines will not take place after discussion amongst the group members.***

The next consideration was with regard to log sheets for the guidelines. Positive views were that they helped give the user a visual guide to the guidelines and that they helped summarise the guidelines. Negative views included indications that log sheets were unnecessary. It was concluded that for the moment no additional log sheets should be produced and that this subject could be addressed again at a later date.

The WGMDM discussed sending the guidelines to equipment manufactures. WGMDM decided not to pursue this at the moment.

The need for a temperature and pressure mini-logger guideline was addressed. It was thought that this particular guideline was unnecessary as it could be accommodated using the CTD or current meter guidelines. In addition, a possible guideline for sediment sampling was discussed. MDM noted that OSPAR already had in place a set of guidelines for sediment sampling. This began a discussion on related guidelines from other organisations or projects.

Action 11: ***F. Nast will assemble a list of similar data type guidelines as created by other organisations or projects. Once completed, the list will be forwarded to the Chair.***

Discussion took place regarding how the merging of water samples taken during a CTD deployment was matched with CTD data. This has been an ongoing problem for several years. Some organisations close the bottles as the CTD is being lowered, some as the CTD is being raised and others can take the CTD and water samples on separate casts, possibly even separate days. It was decided that a better understanding of procedures at individual centres or institutes is required before tackling this issue.

Action 12: ***Selected members (F. Nast, J. Szaron, A. Isenor, P. Alenius, T. de Bruin, M. Garcia, G. Slesser) will provide to the WGMDM Chair, a summary of the method used in merging the CTD data with the water bottle samples.***

6 REPORTS ON OTHER MEETINGS

This item is to inform the group of other related meetings and to investigate how links can be established or maintained between WGMDM and these other groups.

6.1 IODE Sea Surface Salinity Data Pilot Project (SSSDPP)

The SSSDPP steering group met in November 2001. The report of this meeting is on the IODE web site as a pdf document. The meeting noted three key issues related to creating a work plan for establishing an international underway data centre: data collection and transfer, processing, archiving and products. Three sub groups were established to address these issues. WGMDM members (L. Rickards, H. Dooley and A. Isenor) are members of SSSDPP sub groups. The sub groups were asked to work together to write sections for the project plan. The next meeting is September 2002.

The WGMDM discussed how they might contribute to the efforts of the SSSDPP. It was suggested that WGMDM members hold considerable underway data at their respective centres or institutes. Thus, the group may be able to contribute data holdings information to the SSSDPP.

Action 13: ***L. Rickards will compile a list of questions related to underway data collections. L. Rickards will then contact S. Jans, T. de Bruin, A. Isenor, P. Alenius, J. Szaron, F. Nast and M. Fichaut for answers to these questions, with the aim to creating an inventory of underway datasets at member centres.***

6.2 Argo Data Management Team

There was a brief report on the Argo data management implementation. It was noted that the netCDF data formats were in the process of being finalised for the profiling floats. The global data centres are now handling profile, track and metadata in netCDF.

It was noted that the Argo Data Management Handbook was nearing completion. The Handbook outlines the data management practices for the Argo members.

6.3 ICES Study Group of Integrated Data Management (SGIDM)

This Study Group was formed under the ICES Advisory Committee on Ecosystems (ACE). The SGIDM were initially scheduled to meet in January 2002. However, the initial meeting was delayed and it is now unclear when the group will conduct its first meeting. The Terms of Reference for the SGIDM were reviewed and some confusion resulted when comparing the TORs to some WGMDM member's impression that the SGIDM was to be a non-technical, user oriented group. It was decided that WGMDM would wait for the SGIDM meeting report to determine their focus. The WGMDM Chair will attempt to keep informed of SGIDM activities.

6.4 ICES Steering Group on GOOS (SGGOOS)

It was noted that the Steering Group will be meeting in Halifax next week and that the WGMDM Chair will attempt to attend. H. Dooley informed the group of NORSEPP (North Sea Ecosystem Pilot Project). It contained eight sub elements/packages. Two of the packages were Data Compilation and Data Management and Exchange. SGGOOS will consider NORSEPP at its next meeting.

Action 14: A. Isenor will contact and inform the NORSEPP Planning Group of the data type guidelines.

6.5 Mediterranean Forecasting

M Fichaut informed the group of the Mediterranean Forecasting System Pilot Project (MFSPP). This project is conducting monitoring work using XBT, profiling and moored instrumentation for input to models. Members were not aware of this project.

6.6 WOCE Data Products Committee

WGMDM members were informed that the final WOCE data conference was held in Hobart and that a final data set would be produced for November 2002. Copies will be made available to WGMDM members upon request (contact L. Rickards).

7 INTEGRATED TAXONOMIC INFORMATION SYSTEM (TOR 4)

R. Gelfeld (US NODC) presented the current ITIS website. New developments included a new search facility, all hierarchy shown and direct access from the web. ITIS codes are now mirrored in Canada and Mexico. The ITIS codes are now part of Species 2000, which is a co-ordinating organisation of which ITIS is a partner.

At the 2001 MDM meeting it had been agreed to ask ITIS to provide their database on a CD. It was identified that the US-NODC would need to write access software and that the CD would become out of date in a month, as about 10,000 records are added monthly. Even considering these issues, the MDM recognised that many users do not have web access. In such cases a CD version is essential. It was agreed that a CD version issued annually would be a good idea.

Regarding updates to ITIS, R. Gelfeld informed the WGMDM that the ITIS consortium in the US would now give priority to ICES species updates passed through T. O'Brien. Formerly updates have been very slow. It was also noted that ITIS would welcome the idea of ICES becoming a local ITIS server.

Some conflicting procedures for using the Taxonomic Serial Number (TSN) were noted. At present, the US server provides the TSN while the Canadian and Mexico servers do not. Some participants felt that extra codes, such as TSN, are not needed as they just caused confusion and it was better to use an unambiguous name. Others thought the TSN was essential for communication between the databases.

Action 15: E. Vanden Berghe and T. O'Brien, with help from N. Kaaijk, F. Nast, Marc Costello and David Nicholls, to resolve the ICES MDM position on the use of TSNs by November 2002.

Action 16: E. Vanden Berghe to contact Todd O'Brien to work out the annual production of the ITIS CDs. E. Vanden Berghe software may be used.

Action 17: H. Dooley and T. O'Brien to investigate ICES mirroring of the ITIS site.

Action 18: E. Vanden Berghe will test the update speed of ITIS without use of the quick method via T. O'Brien.

The MDM noted a link between a previous issue and the current ITIS discussion. MDM felt that ITIS could be used as a bridging mechanism to introduce the MDM to other ICES Groups. However, any ITIS brief to the other Working Groups should occur after the ITIS mirroring was established at ICES. It was noted that ITIS is aware of MDM activities as T. O'Brien has provided ITIS with information on the WGMDM guidelines (the guideline information was very well received by the ITIS organisation). It was suggested that any brief should be an agenda item for the meeting of WG Chairs at the ICES Annual Science Conference.

MDM 2001 Action 12 had tasked BODC to match their directory to ITIS. L. Rickards reported that this was nearly completed.

Action 19: ***Todd O'Brien (with Chair's help) will co-ordinate and produce a brief on ITIS for other ICES WGs. This is to be circulated inter-sessionally to MDM for comment.***

Action 20: ***ICES MDM Chair will ensure that ITIS brief is on the agenda of the WG Chair's meeting.***

Action 21: ***ICES WGMDM Chair to encourage the completion of the matching of BODC 's directory to ITIS.***

8 ICES DATA SUBMISSION REPORT (TOR 1)

H. Dooley noted that his intention was to review the data submitted to the ICES Data Centre over the year. The full report in presented in Annex 7.

There has been a major shortfall in data submitted which was damaging the interests of ICES. ICES have attempted to collect nutrient data for OSPAR and provide trends for major estuaries. There was data for Elbe, Wesser, Rhine, Scheldt, Thames and Humber but there was not enough data in the western UK, Ireland, Spain and Portugal. They were looking for a response to OSPAR's actions to clean up the North Sea but because of the shortfall in data the study will have to be repeated.

H. Dooley noted that ICES had recently received, from Dr Rothburg of Hamburg University who was retiring, a nutrient atlas of the North Sea containing data from 7,000 stations in the North Sea. Only 16% of these data were already held in the ICES database.

Some of the members questioned tabulated values in the Data Centre Submission, in particular the figures shown for ROSCOP data provided by Canada. It was also noted that although ICES was the global collection point for ROSCOP, it appeared that only ICES countries were providing ROSCOPS to ICES.

The lack of nutrient data was discussed. H. Dooley noted that the data was not being supplied to ICES and some countries took a long time to process the data. There were other quality control problems noted. A particular example dealt with a SOOP vessel that provided some low salinity values in underway data. On investigation it appeared that the fire pumps had been turned on resulting in 3 days of bad data. Apparently there is a lack of proper quality control (QC) on underway thermo-salinograph data. It was apparent that the data was fed to GTS with only very coarse automatic QC checks applied.

Concerning silicate, nitrate and phosphate data in the NW Atlantic, systematic differences between nutrient values from different countries was noted. This was probably due to quality of manpower and there was a need for firm standards and procedures. This will be brought to the attention of the ICES Marine Chemistry Working Group.

The general question of data trace-ability was raised. Duplicating datasets at various Centres runs the risk of having different versions of the same dataset. In this situation, the problem is where the master copy of a dataset was held and the fact that ownership may not be traceable. This problem may be addressed in a distributed system using XML or something similar. However, the MDM recognised that the distributed model is a long-term solution. History had shown the need for a deep, long-term archive as this avoided the need to go to several places for data. It was also noted that not all Data Centres or agencies had the resources to make their data available via a website and in turn this indicates a continuing need for central or regional archives.

It was agreed that all Centres need to track the data carefully to ensure we always have the latest and best available. checking for duplicates is a huge problem for the NODCs – considerable effort goes into this task.

Action 22: ***H. Dooley to check the number of ROSCOP submissions provided by Canada as indicated in Annex 7, Table 1.***

Action 23: ***P. Ennet will investigate and report to the Chair, the situation regarding the lack of Estonian data submissions to ICES.***

Action 24: ***H. Dooley to draw IODE SSSDPP attention to serious quality control problems with underway data.***

Action 25: ***J. Gagnon will contact Germany, Russia and Iceland to resolve the inconsistencies in the chemical data in the NW Atlantic as presented by H. Dooley (see Annex 7, Figure 3).***

9 PARAMETER DICTIONARIES (TOR 3)

L. Rickards presented the group with an update on parameter dictionary activities. At present the UK and France are using the BODC dictionary, while Canada is using a modified version of GF3. There is also planned an SGXML activity that provides the basis for mapping the various dictionaries. It was noted that mapping exercises could be difficult. As an example, L. Rickards presented temperature, which currently has 26 different codes in the BODC dictionary. It is a difficult task to map even this variable to the SISMER or MEDS dictionary.

In terms of metadata, internationally we have ROSCOP/CSR, MEDI/GCMD and EDMED. There was a recognised need for a defined hierarchy in any defined metadata structure. The structure needs to provide links between the individual presentations provided by MEDI/GCMD, EDMED and ROSCOP. It was noted that metadata in the ROSCOP and EDMED systems have a reasonably good mapping. More information has been added with the EDMED system.

The importance of improving the BODC parameter dictionary categories was also noted. These categories represent large groups of codes, and do not form part of the actual eight character code. The categories have been constructed out of convenience and all need to be rationalised.

The WGMDM then began a general discussion on several questions:

1) Do we want to define a hierarchy? Do we want to start all over again or use older ones?
2) Do we want to wait for the SGXML activities and the mapping dictionaries into a common XML structure?
3) Do we want to improve BODC parameter code categories?

The MDM noted that the members must strive to use existing codes rather than creating another dictionary. However, waiting for SGXML may delay the process. A connection with acoustic groups was also suggested. It was generally felt BODC recognises the need to improve the categories in the BODC dictionary. However, it is a low priority.

Action 26: ***Information from SGXML should be made available on MDM Yahoo website. L. Rickards will evaluate the results of the SGXML mapping exercise and entrain other WGMDM members to help with the evaluation.***

10 10. REPORT ON XML (TOR 5)

L. Rickards initiated a general discussion on the activities at the ICES/IOC SGXML meeting earlier in the week, by summarising the main points of the meeting. These points included the difference and importance of syntax vs. semantics. General metadata developments were noted and in particular the common DTD of GMCD and MEDI. The MDM were informed of existing standards for metadata including FGDC, ISO19115/TC211, and defined MARC DTDs. It is expected that all such geo-spatial standards will converge around ISO metadata standard.

One point from SGXML meeting was that the Group should be using what is available in GML with reference to date/time, latitude, longitude, point, line and polygons. Open GML has good spatial descriptions and SGXML will attempt to capitalise on these developments. The Russian MedBlack DODS example was noted and the importance of a common, well-formed data model and dictionary was essential.

The Keeley bricks were considered by SGXML and in particular relation to the question of markup vs. content. A complete list of elements was needed. In addition, the SGXML activity to place GF3 into an XML structure was noted.

The parameter dictionary mapping to a developed XML structure during the SGXML meeting was noted. There are at present 11 groups participating in this mapping. The general conclusion of SGXML to concentrate on metadata and point data was noted.

The MDM considered the relevance of the SGXML activity to MDM activities. The Group considered the relevance to be within the broader context of data exchange. In particular, the distributed data model may provide the framework for implementing the SGXML results to the international data stream.

The distributed model (as originally described by N. Mikhailov during SGXML) is technically possible today. Consider a user at a browser window where they specify the spatial-temporal space of interest. The user may specify the parameters being sought. This request is then sent to a "navigator", which identifies the various networked databases that contain data in the spatial-temporal-parameter space being requested. The "navigator" then distributes the request to each site that satisfies the request. At the local sites, the "integrator" accepts the request. The "integrator" is simply a piece of JAVA code using the Document Object Model (DOM) for XML. The "integrator" accepts the request and parses it into SQL and sends the request to the local database. The database returns a record set to the "integrator". The "integrator" then converts the record set to XML and sends it back to the "navigator". The "navigator" combines all the returning XML documents and sends the data to the client. The Russian NODC system is hoping to have some of this model operational sometime in 2002.

This type of virtual data centre approach has many internal issues. For example, a search on temperature could result in numerous different parameter codes from the many online databases. This would tend to confuse clients if no mapping to a single parameter space was conducted. There are also issues with mirroring of sources to prevent off-line occurrences. There may also be granularity issues related to information in the databases and the XML data stream.

In the distributed model, a data state indicator is an important concept. This has the potential to identify the best quality dataset. However, the system would need a mechanism to identify that a higher quality version existed.

The MDM noted that the 2001 action items under this topic have all been completed. The WGMDM considered if this TOR should be continued. It was decided not to include an XML specific TOR in the current report for next year. The ICES/IOC SGXML can deal with the XML issues.

However, the Group did recognise MDMs role in using the results from SGXML. It was suggested that MDM develop a data management strategy using the tools developed by SGXML. MDM may also want to improve the virtual data model concept while removing the XML TOR. A new TOR for the current year was suggested.

11 OTHER BUSINESS

11.1 Remaining Action Items from Last Year

The Chair began this discussion by reviewing all of last years action items that were not formally addressed during the previous three days (see Annex 2 for a complete list of last years action items). Of the 22 action items from last year, 10 items were reviewed with the following outcomes:

Action Item 1:	Completed
Action Item 2:	Completed
Action Item 3:	Not Completed.
Action Item 4:	Completed.
Action Item 5:	Not Completed.
Action Item 6:	Not Completed.
Action Item 9:	Completed
Action Item 20:	Not Completed.
Action Item 21:	Completed
Action Item 22:	Completed

It was noted that Action Item 4 resulted in a list of recommended changes to be implemented, but no changes have yet been made.

Action 27: A. Isenor and H. Dooley will review the list of recommendations from MDM 2001 Action #4 and incorporate into the ICES MDM site as required.

11.2 Review of Current Years Action Items

The action items from the meeting were presented to the MDM. It was noted that in all cases except one, names have been attached to the items rather than the generic "MDM" as a group. It was noted that every member had his or her name appearing in the Action Item list.

The MDM also referred back to the initial three items discussed under Section 3 of this report, pertaining to the external reviews. It was felt that these action items make substantial progress on comments 1 (communication) and 3 (promotion). It was noted that comment 2 lacks sufficient action. However, next years TOR5 (see Annex 8) should address this in the longer term.

11.3 Next Years TORs

Suggested TORs for the coming year were presented to the MDM. The group made several edits to the TORs. The resulting TORS are given in Annex 8.

11.4 GETADE Meeting

G. Reed proceeded to invite all WGMDM members to the upcoming GETADE meeting to take place on Saturday April 20 and Monday April 22. Several WGMDM members will be attending this meeting.

11.5 Next Meeting and Concluding Remarks

On behalf of Sweden, J. Szaron and L. Fyrberg volunteered to host the next meeting in Gothenborg, Norköping, immediately following Easter. SGXML and WGMDM will again be juxtaposed within a week period. The exact dates of the meeting will be determined at a later date. The dates have to be co-ordinated with an EDIOS meeting in Norway.

The Chair closed the meeting by thanking the participants. On behalf of the WGMDM, the Chair also thanked the Finnish Institute for their hospitality and arrangements and in particular acknowledged the efforts of Riitta Olsonen and Pekka Alenius. The interest of Riikka Hietula to endure the entire week was acknowledged. The meeting closed at 14:40 on April 19, 2002.

ANNEX 1: NAMES AND ADDRESSES

Names, addresses and contact points of participants. Note that WGMDM members are denoted by (WG), GETADE members are denoted (G) and other participants by (O).

Alenius, Pekka, (WG)
Finnish Institute of Marine Research,
P.O. Box 33, (Lyypekinkuja 3),
00931 Helsinki,
Finland
Tel (operator): +358 9 613 941
Tel (direct): +358 9 613 94439
Fax: +358 0 61394494
E-mail: pekka.alenius@fim.fi
Web page: http://www2.fimr.fi/ or www.fimr.fi

Carval, Thierry, (O)
Institut Francais pour le Recherche et
l'Exploitation de la Mer (IFREMER),
Centre de Brest,
Departement IDM,
BP 70,
29280 Plouzane
France
Tel: 33–2-98–22–4597
E-Mail: theirry.carval@ifremer.fr

Collins, Donald W., (G)
U.S. National Oceanographic Data Center
1315 East West Highway, 4th Floor,
Silver Spring MD, 20910,
USA.
Tel: +1 301 713 3275 extn 179
Fax: +1 301 713 3302
E-mail: donald.collins@noaa.gov

de Bruin, Taco (O)
Secretary National Oceanographic Data Committee (NODC) of the Netherlands,
Royal Netherlands Institute for Sea Research (NIOZ)
E-mail: bruin@nioz.nl
Web page: http://www.nioz.nl or
www.nodc.nl

Dawson, Garry, (WG)
Maritime Environment Information Centre,
UK Hydrographic Office,
Admiralty Way, Taunton,
Somerset TA1 2DN,
UK
Tel: +44 1823 337900 extn 3225
Fax: +44 1823 284077
E-mail: garry.dawson@ukho.gov.uk
Web page: http://www.hydro.gov.uk/

Ennet, Peeter, (WG)
Marine Systems Modelling Section, Estonian Marine Institute,
University of Tartu,
Paldisni Str. 1
Estonia
E-mail: peeter@sea.ee

Dooley, Harry, (WG)
ICES Oceanographer,
International Council for the Exploration of the Sea (ICES),
Palaegade 2-4,
1261 Copenhagen K,
Denmark
Tel (operator): +45 33 154225
Tel (direct): +45 33 152677 (tone) 210
Fax: +45 33 934215
E-mail: harry@ices.dk
Web page: http://www.ices.dk

Fichaut, Michele, (WG)
Institut Francais pour le Recherche et
l'Exploitation de la Mer (IFREMER),
Centre de Brest,
Departement IDM,
BP 70,
29280 Plouzane
France
Tel: 33–2-98–22–6663
E-Mail: michele.fichaut@ifremer.fr

Fyrberg, Lotta, (WG)
Swedish Meteorological and Hydrological Institute,
Oceanographic Services,
Nya Varvet 31,
SE- 426 71 Vastra Frolunda,
Sweden
Tel: + 46 (0)31 751 8978
Fax: +46 (0)31 751 8980
Web page: http://www.smhi.se
E-mail: lotta.fyrberg@smhi.se

Gagnon, Jean, (WG)
Marine Environmental Data Service (MEDS),
Department of Fisheries and Oceans,
200 Kent Street, 12th Floor,
Ottawa, Ontario K1A OE6,
Canada
Tel: +1 613 990-0260
Fax: +1 613 993-4658
E-mail: GagnonJ@dfo-mpo.gc.ca
Web page: http://www.meds-sdmm.dfo-mpo.gc.ca/

Garcia, Maria Jesus (WG)
Instituto Espanol de Oceanografia
Corazon de Maria 8
28002 Madrid
Spain
Tel: +34 1 3473612
Fax: +34 1 4135597
E-mail: mjesus.garcia@md.ieo.es
Web page: www.ieo.es

Gelfeld, Robert D., (WG)
U.S. National Oceanographic Data Center/
World Data Center - A Oceanography,
1315 East West Highway, 4th Floor,
Silver Spring MD, 20910-3282,
USA
Tel: +1 301 713 3295 extn 179
Fax: +1 301 713 3303
E-mail: rgelfeld@nodc.noaa.gov
Web page: http://www.nodc.noaa.gov

Hietula, Riikka, (O)
Finnish Institute of Marine Research
P.O. Box 33, Fin-00931,
Helsinki,
Finland
E-Mail: riikka.hietula@fimr.fi

Isenor, Anthony, (WG)
Bedford Institute of Oceanography,
P.O. Box 1006,
Dartmouth,
Nova Scotia B2Y 4A2,
Canada
Tel: 902 426 4960
Fax: 902 426 7827
E-mail: isenora@mar.dfo-mpo.gc.ca
Web page:
http://www.mar.dfo-mpo.gc.ca/science/ocean/welcome.html

Jans, Siegrid (WG)
Management Unit of the North Sea
Mathematical Models (MUMM),
Gulledelle 100,
B-1200 Brussel,
Belgium
E-mail: s.jans@mumm.ac.be
Web page: http://www.mumm.ac.be

Nast, Friedrich, (WG)
Deutsches Ozeanographisches Datenzentrum (DOD),
Bundesamt für Seeschiffahrt und Hydrographie
Bernhard-Nocht-Str. 78
D-20359 Hamburg,
Germany
Tel: +49- (0) 40 - 3190–3530
Fax: +49- (0) 40 - 3190–5000
E-mail: friedrich.nast@bsh.de
Web page:
http://http://www.bsh.de/Oceanography/DOD/htm

Olsonen, Riitta, (WG)
Finnish Institute of Marine Research,
P.O. Box 33, (Lyypekinkuja 3),
00931 Helsinki,
Finland
E-mail: riitta.olsonen@fimr.fi
Web page: http://www2.fimr.fi/ or www.fimr.fi

Reed, Greg, (G)
Intergovernmental Oceanographic Commission (IOC),
1 Rue Miollis
75732 Paris Cedex 15,
France
Tel: 01 45 68 3960
E-Mail: g.reed@unesco.org

Rickards, Lesley, (WG)
British Oceanographic Data Centre,
Proudman Oceanographic Laboratory,
Bidston Observatory, Prenton,
Merseyside, CH43 7RA,
United Kingdom
Tel: +44 151 653 1514
Fax: +44 151 652 3950
E-mail: ljr@bodc.ac.uk
Web page: http://www.bodc.ac.uk;
http://www.oceannet.org

Sagan, Slawomir, (WG)
Institute of Oceanology PAS,
Powstancow Warszawy 55,
81–712 SOPOT, PL
Poland
Tel: +(48 58) 5517 283 x211,
Fax: +(48 58) 5512 130
E-mail: sagan@iopan.gda.pl
Web page: http://www.iopan.gda.pl

Sagen, Helge, (WG)
Institute of Marine Research
Norwegian Marine Data Centre
Nordnesgt 50,
5817, Bergen
Norway
Tel: 47 55 23 8500
E-Mail: helge.sagen@imr.no

Slesser, George, (WG)
Marine Laboratory,
Fisheries Research Services (FRS),
P.O. Box 101,
Victoria Road,
Aberdeen, AB9 8DB,
Scotland
Tel: +44 1224 876544
Fax: +44 1224 295511
E-mail: slesser@marlab.ac.uk
Web page: http://www.marlab.ac.uk

Szaron, Jan, (WG)
Swedish Meteorological and Hydrological Institute,
Oceanographic Services,
Nya Varvet 31,
SE - 426 71 Vastra Frolunda,
Sweden
Tel: +46 (0)31 751 8971
Fax: +46 (0)31 751 8980
E-mail: jan.szaron@smhi.se
Web page: http://www.smhi.se

Tikka, Kimmo (O)
Finnish Institute of Marine Research,
P.O. Box 33, (Lyypekinkuja 3),
00931 Helsinki,
Finland
E-mail: kimmo.tikka@fimr.fi
Web page: http://www2.fimr.fi/ or www.fimr.fi

Vanden Berghe, Edward, (G)
Manager, Flanders Marine Data and Information Centre
Flanders Marine Institute
Vismijn, Pakhuizen 45–52,
B-8400 Ostend
Belgium
Tel: +32 59 342130
Fax: +32 59 342131
E-Mail wardvdb@vliz.be
Web Page: http://www.vliz.be

ANNEX 2: 2001/2002 TERMS OF REFERENCE AND ACTION ITEMS

TOR 1) Quantitatively assess the last 5 years data (1997–2001) sent to the ICES Oceanographic Data Centre by each Member Country, identify problems and suggest solutions.

The amount of oceanographic data received by ICES continues to decrease. This item will provide the impetus to encourage an increased data flow to the ICES Oceanographic Data Centre from Member Countries. This item will provide the impetus to encourage an increased data flow to the ICES Service Hydrographique from Member Countries;

*Action 1) The **Chair** will draft a note on behalf of the MDM to the Oceanography Committee raising our concerns regarding the recent changes within the ICES Secretariat and the consequent jeopardy of its valued archives and services. The present procedures of maintaining off-site back-ups of the ICES archives at the WDC-A will be emphasised.*

*Action 2) The **Chair** will draft a note on behalf of the MDM requesting that the WGOH evaluate the product.*

*Action 3) The **Chair** will re-organise the present MDM e-group web site into appropriate folders, create a folder for co-ordination of the final report, and add a file of this year's attendees with their current mailing addresses, e-mail, and URL's.*

*Action 4) **J. Gagnon** agreed to affect a review of the contents of provisional MDM web site contents in order to migrate relevant information to the ICES web site.*

*Action 5) The **Chair** will create a folder on the MDM e-group web where members can input their list of data CDs. Identification, originator/contact and a short description of each CD's contents were some of the information deemed relevant to record.*

*Action 6) The **Chair** is to include a comment regarding the inconsistent data policy of international project CD-ROMS and extend an offer to review project CDs upon request, as part of his summary to the ICES Oceanography Committee.*

TOR 2) Continue to critically evaluate the guidelines for data management and exchange developed inter-sessionally for the following data types: moored current meter data, shipboard and moored ADCP, CTD, XBT/XCTD, sea level, surface underway measurements, nutrients, oxygen and chlorophyll.

There is a need for simple guidelines for those processing, quality assuring and managing data. The existence of written guidelines has distinct advantages. It shows laboratories reporting data to the ICES Oceanographic Data Centre how important it is to apply quality control procedures on the data, and it will provide ICES with data sets which are easier to handle and have a properly documented quality control history behind them. This leads to an improved data set being available to the ICES community

*Action 7) **L. Rickard's** will present the guidelines to the ARGO Committee Meeting in September for comment.*

*Action 8) **A. Isenor** will conduct more inter-guideline consistency checks together with guideline maintenance. Consistency checks will include the detailed wording in the various sections, possible inclusion of deck sheets similar to the ADCP guidelines,* etc.

*Action 9) **R. Gelfeld** will send a copy of the revised guidelines with cover letter to the Chairs of the Working Groups under the Oceanography Committee. An accompanying letter should explain the framework for the guidelines, and volunteer to produce others at the suggestions of the Working Groups.*

*Action 10) **A. Isenor** will examine the algae bloom working group guidelines.*

*Action 11) **T. O'Brien** will discuss our existing guidelines and the guideline framework with the Chair of Zooplankton Working Group.*

Suggestion 1) The guidelines could be combined with example submission files in an ICES CDROM product.

Suggestion 2) *The guidelines could be extended into standards. (It is important to note the difference between guidelines and standards. Guidelines simply provide an outline of the information required. A standard implies that we would like submissions to meet a particular level of process or service.)*

TOR 3) Report on the parameter dictionaries use in ICES Member Countries and evaluate present ROSCOP system to see how these new metadata procedures can change and improve it.

A number of parameter dictionaries for oceanographic data have been developed by the oceanographic community. The current ROSCOP system will be evaluated to see how these new metadata procedures can change and improve it with a view to recommending use with ICES, if appropriate;

Action 12) T. O'Brien/R. Lowry to set up link table between BODC parameter dictionary and ITIS

TOR 4) Report on common taxonomic coding systems use in ICES Member Countries. Formulate a model of how the Integrated Taxonomic Information System (ITIS) might be expanded internationally.

A number of taxonomic coding systems for oceanographic data have been developed by the oceanographic community. Many institutes are maintaining their own systems, and there is concern over the lack of co-ordination. The working group recommended increasing coordination among WGMDM members and also between WGMDM and ITIS. This could involve pooling taxonomic specialists (resources) to speed the review of local biota for both the benefit of ITIS and the local institute. It also recommended ITIS being used as a taxonomic authority, and encouraged its members to follow the examples of (Germany and Netherlands) which provide ITIS codes in addition to their own codes.

TOR 5) Report on XML code use in ICES Member Countries.

A specialized Study Group with a view to recommending use within ICES will investigate these and make appropriate recommendations.

Action 13) The ***MDM*** *will form an XML Study Group (XSG) to examine XML in a marine context. The XSG will be Chaired by R. Gelfeld and will consist of the participants noted in the above table. The XSG will proceed with the three XML projects as outlined:*

Point data	*– profile, underway, water sample*
Metadata	*– metadata cruise information*
Biology	*– integrated tows*

Action 14) The ***XML Study Group*** *(XSG) will prepare an action plan to initiate and implement a marine XML based on the outlined Projects.*

Action 15) ***MDM*** *will inform the Marine Consortium that ICES MDM (or someone to represent it) will be a partner. Details regarding Consortium fees will be determined.*

Action 16) ***MDM*** *will encourage the use of tags that remove data content from the tag structure.*

Action 17) This TOR should be placed on next year's agenda as a stand-alone Term of Reference.

Taxonomic Coding Systems, CSR, and Other Business

Action 18) ***T. O'Brien*** *volunteered to suggest that ITIS produce a CD-ROM of their database to allow use without web access.*

Recommendation 1) The working group recommended increasing coordination among ***WGMDM members*** *and also between WGMDM and ITIS. This could involve pooling taxonomic specialists (resources) to speed the review of local biota for both the benefit of ITIS and the local institute.*

Recommendation 2) The working group recommended ITIS being used as a taxonomic authority, and encouraged its ***members*** *to follow the examples of Germany and The Netherlands which provide ITIS codes in addition to their own codes.*

Recommendation 3) The working group recommended to make the development of common taxonomic systems a separate Term of Reference.

Action 19) That the new study sub-group on XML inventories includes a review of ROSCOP contents and exchange process within its objectives and that members continue to complete and submit ROSCOP forms with their data submissions as in the past.

Action 20) ***L. Rickards*** *to provide a copy of the Microsoft Access 97 version of the EDMED database and supporting information to P. Geerders*

Action 21) ***P. Hadziabdic*** *to provide P. Geerders with contact details*

Action 22) All ***MDM*** *members to check their archives for data from the Caribbean.*

ANNEX 3: LIST OF ACRONYMS AND TERMS

Acronym or Term	Description
ADCP	Acoustic Doppler Current Profiler
Argo	The Array for Real-time Geostrophic Oceanography
BIO	Bedford Institute of Oceanography
BODC	British Oceanographic Data Centre
BOOS	Baltic Operational Oceanographic System
CSR	Cruise Summary Report
CTD	Conductivity-Temperature-Depth
DAC	Data Assembly Centre
DOM	Document Object Model
DPC	Data Products Committee
EDIOS	European Directory of the Initial Observing System
EDMED	European Directory of Marine Environmental Data
GCMD	Global Change Master Directory
GETADE	IOC's Group of Experts on the Technical Aspects of Data Exchange
GODAR	Global Oceanographic Data Archaeology and Rescue
GOOS	Global Ocean Observing System
GTSPP	Global Temperature-Salinity Profile Programme
HELCOM	Helsinki Commission
ICES	International Council for the Exploration of the Sea
IOC	Intergovernmental Oceanographic Commission
IODE	International Oceanographic Data and Information Exchange
ITIS	Integrated Taxonomic Information System
JGOFS	Joint Global Ocean Flux Study
MEDAR	Mediterranean Data Archaeology Rescue Project
MEDI	IOC Marine Environmental Data Information Referral Catalogue system
MEDS	Marine Environmental Data Services - Canada
MFSPP	Mediterranean Forecasting System Pilot Project
NODC	U.S. National Oceanographic Data Center
OCL	Ocean Climate Laboratory/U.S. NODC
OSPAR	Oslo-Paris Commission
PI	Principle Investigator
QC	Quality Control
ROSCOP	Report of Observations/Samples Collected by Oceanographic Programmes (now CSR)
SGQAE	ICES/OSPAR Steering Group on Quality Assurance of Biological Measurements in the Northeast Atlantic
SGXML	ICES/IOC Study Group on the Development of Marine Data Exchange Systems using XML
SISMER	French National Oceanographic Data Centre
SOOP	Ship of Opportunity Programme
SSSDPP	Sea Surface Salinity Data Pilot Project
SQL	Structured Query Language
TOR	Term of Reference
VEINS	Variability of Exchanges in the Northern Seas
WDCA	World Data Center for Oceanography/Silver Spring
WGMDM	Working Group on Marine Data Management
WGOH	Working Group on Oceanic Hydrography
WOA	World Ocean Atlas
WOAF	World Ocean Atlas Figures
WOCE	World Ocean Circulation Experiment
WOD	World Ocean Database
XBT	Expendable Bathythermograph
XML	Extensible Markup Language
XSL	Extensible Stylesheet Language

ANNEX 4: WGMDM REVIEWS

Oceanography Committee Comments
(Reproduced from the ICES Annual Report)

The report (Doc. C:09) was presented by Jan Szaron (Sweden). He explained that this Group is anxious to increase interaction with all Working Groups of the Committee, particularly as data management is central to the resolution of many of the issues considered by them. In this context he reminded the Committee that the Working Group had asked it in 2000 to participate in the development of data management guidelines, but so far it had not received any feedback. He also explained that it is a goal of this Group to introduce ICES to, and to implement within the ICES framework, new technologies for data management. For this reason he asked the

Committee to endorse the initiative to set up a Study Group on XML, which is being seen as an important building block for the development of distributed internet-based data systems.

Harald Loeng (Norway) presented the review of this report. He remarked that the Group had reacted positively to the rather critical review it had received last year. He also noted that the Group must more aggressively address operational issues if it is to remain relevant with regard to upcoming global operational initiatives (e.g., ARGO). It was also noted that the data management guidelines must be promoted more aggressively and also that the Group should work with other groups who are developing their own guidelines.

Review of the WGMDM Report

As requested by the Oceanography Committee, here are our comments on the ICES MDM WG report.

Overall, we found this year's report easy to read and well organised. We particularly liked the way it is organised with respect to action items, giving responsibility for each item to specific individuals. Furthermore, reading through the report, we had a sense that MDM as a group is moving into areas that needed attention for some time, such as XML and ITIS, and is trying to interact better with other data management initiatives. It is important that this trend is continued for MDM to be relevant within ICES and elsewhere.

The WG was given four Terms of Reference this year. The report indicates that they all have been met satisfactorily.

The reviewers noted that the WG discussed seriously the last year's review, and took the comments into account. The report indicates that MDM was concerned about the limited attendance at the OCC meeting. It is to be recognised that for MDM and other WGs to be more effective, better exchange among OCC members and better guidance from them will be necessary. The report could have brought out this point more clearly, and it would have been more informative to state the exact number of participants at OCC, which is available in the ICES annual report for 2000.

There were several points raised at the ASC last year regarding the MDM's mandate and performance. Taking on tasks that are of relevance to other ICES WGs and establishing effective communication with other international and national bodies were identified as areas where MDM could improve on. This report clearly indicates that MDM has made a serious attempt to achieve this, particularly with the IODE and its subsidiary Group of experts. However, we did not get a clear picture of how MDM proposes to interact with other ICES WGs. For example, we believe that MDM could play a role in the ICES/IOC steering group on GOOS.

The report also indicates that even though ICES member countries are increasingly moving towards operational oceanography, MDM continues to solely focus on 'delayed mode operations'. The report does not address the real-time aspect of data management. My suggestion is that MDM also establish links with the JCOMM Programme Area on data management and its subsidiary groups to meet the real-time requirements of ICES. This may be achieved by MDM being one of the observers in the Co-ordination Group of this Programme Area, or by nominating members of MDM by their countries to the various JCOMM Expert Teams.

The MDM has been successful in producing guidelines for data management and exchange for a variety of data types. However, a guideline is useful only if the community accepts it and implements it in their programmes. Perhaps an action item here could be promoting these guidelines within ICES and within member countries'. Furthermore, there are other groups developing similar guidelines, and we are sure that both MDM and these groups will benefit from closer collaboration. This relates back to interagency collaborations. My suggestion here is for MDM to inform these groups (IODE, COOP, JCOMM, PICES, etc.) of the existence of guidelines and offer to provide advice on data management. We know individually a few countries are doing this, but unless MDM makes a real effort collectively, it would be very difficult to establish these guidelines.

ANNEX 5: PRESENTATIONS

A4.1 Day 1 Presentations:

J. Gagnon (Canada)

Web Access to MEDS Surface Gravity Wave Archives

Canada's Marine Environmental Data Service (MEDS) maintains a national database of surface gravity wave data collected in the Canadian Area of Interest (35˚N to 90˚N, 40˚W to180˚W). MEDS acquires, processes, quality controls, maintains archives and makes these data available to the general public. Spectral wave data, reported operationally on the GOES and ARGOS networks for the above area, are processed daily. MEDS' surface wave data archives include observations from over 400 locations around Canada. The archives date from the mid-1970s and contain over six million individual wave spectra.

All of MEDS observed wave data archives, including some specialized products and client PC display software, are now available for direct download from our prototype web page

http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Databases/WAVE/WAVE_e.htm

Some key components in the development of this capability include:

- Browse and view inventory information using Scalable Vector Graphic (SVG) mapping feature JAVA/XML plug-ins to standard PC web browsers.
- Direct download of user-selectable datasets, of recorded parameters from the archives into CSV ASCII coded files or in MEDS documented format files, to client's PC.
- Graphic displays of standard products, such as monthly time series plots of identified subsets in the archives.
- Options to download PC-based Visual Basic software developed at MEDS to further visualize the downloaded wave and ancillary meteorological datasets.
- Access to meta-data information by gauge location, such as status reports on active operational buoys.
- Access to published literature and symposia documents (in PDF format) specific to the field of wave research.
- Access to a Wind and Wave Climate Atlas for the Canadian Area of Interest.
- Links to other Canadian wave operational programmes.

S. Sagan (Poland)

Institute of Oceanology Polish Academy of Sciences (IOPAS)
Sławomir Sagan

Institute of Oceanology Polish Academy of Sciences is the leading marine research organisation in Poland. The data holdings at the database are mostly collected by Institute itself or acquired in the course of joint research activity in frame of international co-operation.

IOPAS has no obligation for collecting and maintaining marine data from other institutions in Poland.

Most of the data are gathered from Institute's research vessel "Oceania", which operates mainly on the Southern Baltic and European Arctic, from 1986. There are 12–14 cruises annually, with average time 220 days on sea. Type of data collected:

- Hydrology
- marine optics
- acoustic
- hydrochemistry
- marine biology

Data are stored according to cruise/field campaign identifiers. Each group of data are processed and quality checked within the data originator Labs, and are archived by the Data Centre. Data Centre maintains the meta-information base and assists with retrieval of requested data.

Meta data are available via the intranet web server, which points to the data originator.

Planned future activities:

Most of activities concerning development of better data management structure are planned within FP5-funded grant under the call *The integration of 'newly associated states' (NAS) in the European research area"*, due to start at the end of 2002. One of the Work Packages is

"To provide access to the Marine Data Bank to European Community"

IOPAS therefore will seek for an expertise and assistance in following areas

- implementation of guidelines for data management and exchange
- quality control expertise
- XML implementation

M. Fichaut (France)

ADDED VALUES TO OCEANOGRAPHIC MULTI-PARAMETERS DATA SETS, EXAMPLE OF AN EUROPEAN CONCERTED ACTION: MEDAR/MEDATLAS 2

M. FICHAUT [1] for MEDAR GROUP

[1]IFREMER, Centre de Brest, BP 70, 29280 Plouzané, France
(Michele.Fichaut@ifremer.fr, ☎33 (0)2.9822.4643, Fax 33 (0)2.9822.4644)

Basic oceanographic parameters like temperature, salinity and nutrients are needed for various scientific and technical studies. However most of the time, they remain dispersed among all the different organizations, which carry out oceanographic cruises. The overall objective of the EU concerted action MEDAR/MEDATLAS (MAS3-CT98–0174 & ERBIC20-CT98–0103) was to make available a comprehensive data product of such multi-disciplinary *in situ* data and information in the Mediterranean and Black Sea, through a wide co-operation of the Mediterranean and Black Sea countries.

The specific project objectives were: 1. to compile and safeguard historical data; 2. to make available comparable and compatible data sets of: temperature, salinity, oxygen, nitrate, nitrite, ammonia, total nitrogen, phosphate, total phosphorus, silicate, H2S, pH, alkalinity, chlorophyll-a profiles by using a common protocol for formatting and quality checking; 3. to prepare and disseminate qualified value added products by using efficient gridding and mapping methodology; 4. to enhance communication between data managers and scientists to improve the data circulation.

Each participant, who represents the National Co-ordinator for International Oceanographic Data and Information Exchange (IODE) at the Intergovernmental Oceanographic Commission (IOC) of UNESCO, have compiled and safeguarded copies of the data sets dispersed in the scientific laboratories of his country, and reformat them at the common MEDATLAS format. These data sets has been checked for quality (QC) according to the common protocol based on the international IOC, ICES and EC/MAST recommendations, with automatic (objective) and visual (subjective) checks.

The data management structure was distributed between four Regional Data Centres (RDC) and one co-ordinating and Global Assembling Centre (GAC). Each National Oceanographic Data Centre (NODC) or Designated National Agencies (DNA) of the participating countries sent his data set to the corresponding RDC for regional expertise. The data have been gathered and checked for quality in the RDC, and then sent to the GAC which finalise the last quality and duplicate checks. Finally a selection of all the "Good" data interpolated to pre-defined standard levels has been sent to the Analysis Centre (AC) for climatologies computation.

Thanks to this strong international cooperation the volume of available data represents now 286426 stations (vertical profiles) from about 150 sources laboratories of 33 countries. The data released which double the volume of the previously available data, consist of the following number of profiles:

DATA TYPE	NB of STATIONS
BOTTLE	88453
CTD	36054
MBT	81465
XBT	80425
Thermistor string	29
Total	**286426**

PARAMETER NAME	NB OF PROFILES
SEA TEMPERATURE	284946
PRACTICAL SALINITY	118509
DISSOLVED OXYGEN	44989
NITRATE (NO3-N)	10588
NITRITE (NO2-N)	10561
AMMONIUM	5301
SILICATE	15936
PHOSPHATE	20808
ALKALINITY	2548
PH	14548
CHLOROPHYLL-A TOTAL	4716
HYDROGEN SULPHIDE (H2S)	1843
TOTAL NITROGEN	153
TOTAL PHOSPORUS	2381

The *in situ* and climatological data, the climatological maps and all the documentation about the project are in publication on a set of 4 CD-Roms. The *in situ* data are accessible through software designed as a user-friendly interface so that users can extract data from the whole data set, following several criteria.

This software, available for PC/WINDOWS, allows:

- extraction and display of any subset of data selected according to any combination of the following criteria: geographical location, data type (bottle, CTD, bathythermograph), cruise name and reference, time period, month, ship, country, parameter, quality flags
- extraction at three output formats: MEDATLAS, Comma Separated Values for spreadsheet, Ocean Data View (WOCE/Bremerhaven University visualisation software)
- interpolation at pre-defined standard levels or at user defined standard levels
- visualisation of the selected data on parameter/ parameter plots

Such an integrated database facilitates the access to data. It is expected that this integrated data product with meta-data, observed data, gridded data and software, will be a be a valuable tool for all the scientists, engineers and teachers of the Mediterranean and Black Sea region.

T. de Bruin (The Netherlands)

National Oceanographic Data Committee / Royal Netherlands Institute for Sea Research

The National Oceanographic Data Committee (NODC) of the Netherlands is the national platform for oceanographic data management.

The NODC:

- consists of 8 participating institutes from the public and private sector
- covers 90% of all oceanographic data, held in The Netherlands
- provides the infrastructure for the exchange of

  - information
  - knowledge
  - expertise
  - data

The 8 participants are:

1) National Institute for Coastal and Marine Management (RIKZ)
2) Directorate-General of Public Works and Water Management - Directorate North Sea (RWS-DNZ)
3) Hydrographic Service of the Royal Netherlands Navy
4) Royal Netherlands Meteorological Institute (KNMI)
5) Royal Netherlands Institute for Sea Research (NIOZ)
6) Netherlands Institute of Ecology - Centre for Estuarine and Coastal Ecology (NIOO-CEMO)
7) Netherlands Institute of Applied Geoscience TNO - National Geological Survey (NITG-TNO)
8) Delft Hydraulics (WL)

The NODC acts as a virtual data centre through the NODC website (www.nodc.nl) with general information, news, products and online databases of:

- dataset descriptions (EDMED)
- research project descriptions (EDMERP)
- planned cruises
- research reports

The NODC represents the Netherlands in IODE and in the European Euronodim/SeaSearch project.

Activities in past 12 months:

NODC

In November 2001 the NODC organised a very successful national symposium ‘Data, Management and Data management’. Its threefold aim was to:

- present the NODC to a larger audience
- emphasize the role of an institute’s management in oceanographic data management
- get input from the users for future developments of the NODC

RIKZ

- continues to participate in the Commission Integrated Water management
- has furthered the development of Taxonomica: a very flexible system for taxonomic coding
- has brought the DONAR database online at: www.waterbase.nl
- has started working on a new system, to be implemented within 5 years
- will organise the Monitoring-Tailor-Made symposium in June 2002

NITG-TNO

The DINO database of marine geological data was made available on the web: dinoloket.nitg.tno.nl

NIOZ - Data Management Group (DMG)

- has included many new (CTD) datasets in the online database
- is expanding databases to geological and biological data
- maintains active participation in research projects through development of project websites
- plays an important role in public relations of NIOZ
- is the National Antarctic Data Centre

Future Developments of the NODC

- Include more Participants in order to cover 100% of all Dutch oceanographic data
- National Platform
    - ➢ Coordinating Role (e.g., national input to SeaSearch and EDIOS projects)
    - ➢ Start a National Oceanographic Data Management Programme
    - ➢ Establish close cooperation with Belgium
    - ➢ Represent The Netherlands in ICES-WGMDM
    - ➢ Participate in international projects like SeaSearch-II

S. Jans (Belgium)

IDOD: *Integrated and Dynamical Oceanographic Data management*

Management Unit of the Mathematical Models of the North Sea (RBINS)1
University Centre for Statistics (KULeuven)2
Laboratory SURFACES (ULg)3

The existence of a structured and validated knowledge base is an obvious need for any scientific work, especially when dealing with the marine environment. Any policy to be defined or decision to be taken in the perspective of a sustainable management of the North Sea would be meaningless without a background of validated and readily accessible measurements or experimental data.

In the scope of a National “Sustainable Management of the North Sea” Programme, the set–up of an integrated oceanographic database (IDOD) was thus a key action.

The IDOD project was meant to establish, to manage and to promote a data base of marine environmental data, ensuring a smooth and scientifically sound data flow between the data producers (routine monitoring, field and laboratory experiments, mathematical models,...) and the end users (scientists, sea professionals, policy makers,...).

The project was split into five different –but highly inter–dependent– tasks:

As a basis, an inventory of the relevant data sets and databases was undertaken, in order to make them ready for incorporation in the database (a. o. with respect to current standards on data quality and on data documentation).

The procedures pertaining to the incoming flow of data were defined and implemented. This covers not only the practical aspects of the transfer of information but also the very important point of data quality control.

The design of the data base itself has been deeply analysed in function of the intrinsic characteristics of the data and in order to meet the present and future needs, ensuring the viability and the usefulness of the tool over the years.

In order to understand the processes driving the marine phenomena “hidden” in the data, a set of data analysis tools have been developed and are now tuned. Various approaches were used: statistical techniques, geostatistics and spatial analysis, space and time “corrections” of data sets by means of advection-diffusion models. Part of the information given by these tools is also used to improve the quality control on the incoming data.

Finally, as one of the most important objective of this project was to provide useful and scientifically sound information to a wide range of users, derived products (maps, tables, reports,...) that meet the specific requirements and level of expertise of the various categories of users were designed and currently being made available to the users.

The global methodology applied to reach our objectives reduces to the following words: analysis, design, implementation and production.

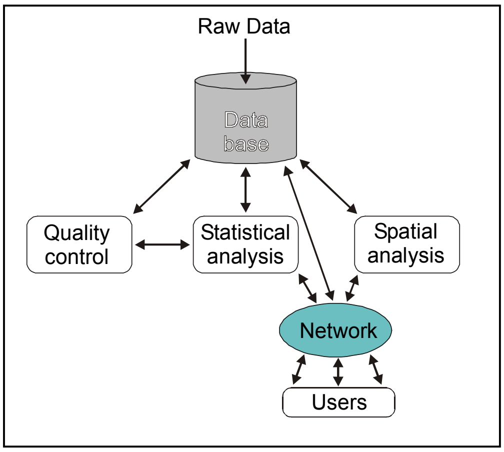
The *analysis* phase went into the details of the data, their structure and intrinsic characteristics, together with a deep insight into the sampling methods, the laboratory practices and the needs and requirements of the potential users.

During this process a special attention was paid to the rights of the parties involved: the data centre, the data producers and the financing authority. After some negotiations, a formal convention specifying the ‘rights and duties’ of each of the parties could be agreed.

---

1 G. Pichot, K. De Cauwer, M. Devolder, S. Jans, M. Moens, L. Schwind, S. Scory
2 J. Van Dyck, B. Plevoets, G. Dierckx
3 J.–P. Donnay, M. Binard, Y. Cornet, F. Muller

During the *design* phase the results of the analysis were translated into functional description: *how would data be entered?*, *how would their quality be checked?*, *how would they be retrieved and analysed?*,..., leading finally to the *implementation* of the information system.



The resulting system, now entering its 'production' phase, consists of a relational data base (running under *Oracle 8*i), quality and statistical analysis tools (based on *SPlus*) and visualisation and spatial analysis tools (developed with the help of *ArcView* and other *ESRI* software packages). As far as the technique makes it possible, processed data are made available to the users via the Web.

Content

The database mainly contains values of the concentrations of numerous substances in the air, the water, the sediment and the biota. These values result from measurements taken in *situ* and analyses carried out in laboratories. In addition to the concentrations, quantitative (biodiversity) and qualitative (pathology) information on the biota is also stored.

These values would be pointless if they were not accompanied by precise information about the circumstances in which they were measured. This is what is known as 'meta-information', a term that covers information such as the position in which samples were taken, the date, the time, the weather conditions, the sampling and analysis methods used, etc.

The database already contains several tens of thousands of items. All these data, documented and verified, constitute a coherent and unique source of information for scientists and other users.

Future

The IDOD project has given the scientific community, the Belgian authorities and other potential users the opportunity to dispose of an up–to–date information and management system about the quality of the marine environment. The sea evolves continuously as do the needs and the demands of the society. The data base, together with its query and analysis tools, opens the door to new scientific investigations and to better policy choices. The onus lies on all the interested parties –the DB managers, the data producers, the authorities and the community of users– to keep this knowledge base useful, *i.e.,* alive and up–to–date.

A4.2 Day 2 Presentations:

G. Dawson (UK)

SUMMARY OF PRESENTATION BY GARRY DAWSON UKHO TO ICES MDM APRIL 2002

The Maritime Environment Information Centre (MEIC) is part of the UK Hydrographic Office which is based at Taunton, Somerset in the SW of the UK. The presentation will review several areas of the branch's work and also introduce a project which, whilst being managed outside of MEIC, should be of interest to ICES MDM.

SOOP: Aside from its obvious close ties with the Royal Navy MEIC collects data from a wide variety of sources. The UK Ministry of Defence fund XBT probes for vessels through the Ships of Opportunity Programme (SOOP). These vessels are mostly UK Research vessels operated by organisations such as the Southampton Oceanographic Centre (SOC) and the Aberdeen Marine Laboratory. Operators are provided with the probes on the basis that the data collected is supplied to the UKHO for inclusion in its databases and that the data will, in some cases after an elapsed period of time, be released to data centres around the world. As UKHO had not released any SOOP data since 1995, the decision was made in 2001 to release UK SOOP data collected from 1995 to 2000. The data, which consists of about 7,000 observations, will be supplied to UK BODC, US NODC and ICES for use by the oceanographic community.

ARGO: UKHO is actively engaged in the UK's contribution to the ARGO programme. UKHO databases are used to propose float launch locations for UK floats and UK survey ships have been employed to launch floats. Turning to data, UK has proposed that Regional Data Centres can make a valuable contribution to the ARGO programme by providing regional expertise to quality control data received from ARGO floats. For the UK BODC, with UKHO assistance, has agreed to maintain a Regional Data Centre for the Southern Oceans and, as part of quality control, float data will be supplied to UKHO for comparison with its databases.

Marine Biology: To support its aims in environmental protection the UK Government requires that environmental assessments are completed for all trials and exercises. MEIC are building a database of observations of Cetaceans that will be available to inform this activity. Marine life observation forms have been distributed throughout the Royal Navy to encourage information of sightings to be collected and contacts are being developed with several civilian conservation organisations who have collected large numbers of observations. An atlas will be produced to assist those at sea to provide accurate information.

Environmental Products: UKHO not only collects and databases environmental data for use in products for the Royal Navy but also for use in collaborative projects. Recently, following a collaboration with Fugro GEOS, a Regional Environmental Summary for West Africa was produced which is aimed at the oil and gas exploration industry there. For the subject area this product provides textural descriptions, maps and diagrams covering a wide range of environmental parameters. Copies are available for purchase at £1000 (approx 1600 Euros) each !

Coastal Zone Map: Finally, although not directly an MEIC project MDM members may find the Coastal Zone Mapping Project (CZM) of interest. Jointly with the British Geological Survey and the UK mapping agency (the Ordnance Survey), UKHO is aiming to produce a "map" of part of the coastal zone of southern England. This is to meet the requirement of users (oceanographers, planners, leisure users) of the littoral zone to have one unified geographic base of the area for their work. Traditionally maps stop at the low water line, charts stop close to the coastline and geological maps terminate at various points but environmental factors know no such artificial boundaries. Furthermore different agencies on land and water use different datums from which to measure both horizontally and vertically. Although the term "map" is used to describe the final product, which is provisionally planned for release in 2003, it is unlikely to simply be a map in conventional terms rather it will be a digital product giving the user scope to display the data how they wish. To ensure that the interests of the maritime community are met in this project the UK's Inter-Agency Committee on Marine Science and Technology (IACMST) is acting as the intelligent customer.

Garry J Dawson
Data Manager
Maritime Environment Information Centre (MEIC)
UK Hydrographic Office
Taunton, Somerset
UK
+44 (0) 1823 337900 x 3225
garry.dawson@ukho.gov.uk
18/03/02

P. Ennet (Estonia)

Coupled Atmospheric-Sea-Ecosystem model

Peeter Ennet, Estonian Marine Institute

The aim of our work is to create a tool for operational prediction of aquatic ecosystem state. For solving this task we are linking three models – the operational weather forecast model HIRLAM (**HI**gh **R**esolution **L**imited **A**rea **M**odel), air pollution transport model Hilatar and aquatic ecosystem model FinEst. In this joint work the HIRLAM model will produce meteorological forcing data for Hilatar and FinEst. The Hilatar and FinEst models will be coupled to change data between each other. The coupling of these two models allow:

a) in FinEst model more exactly predict the plankton community evolution in the sea using atmospheric deposits data from Hilatar;
b) in Hilatar model to get sea surface data from FinEst for sea-atmosphere interaction calculations.

The structure of jointly work of HIRLAM, Hilatar and FinEst models is presented in Slides 1,2 and 5. During run of these three models the HIRLAM model don't use any returning data from other models. So during run of coupled Hilatar and FinEst models it is possible either to use the archived HIRLAM data or in case of operational prediction all these three models will run simultaneously.

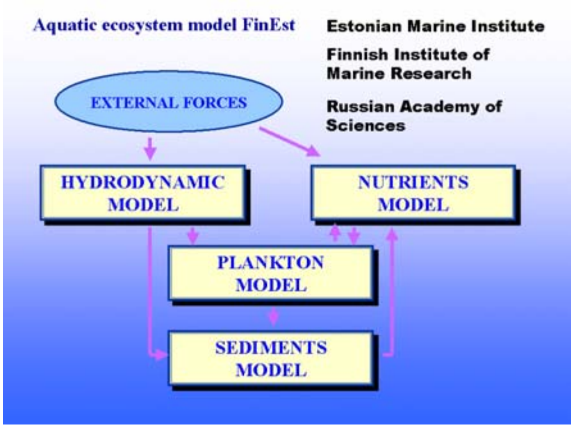
The FinEst model is the 3D coupled hydrodynamic and ecosystem model. The hydrodynamic part of the model describes water temperature, salinity,

suspended material, density, currents and water level. The ecosystem part describes the phosphorus (PO4, DOP), nitrogen(NO3, NO2, NH4, DON) and carbon (CO2, DOC) recycling in the water bodies, sediments and plankton community.

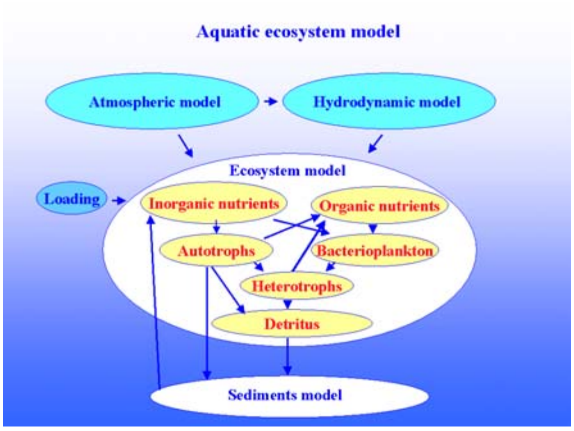
Five categories of autotrophs and five categories of heterotrophs are classified: netphytoplankton, nanophytoplankton, phytoflagellates, picophytoplankton, blue_green algae, mesozooplankton, microzooplankton, nanozooplankton, zooflagellates and bacterioplankton. The parameterization for biochemical reactions based on the Barrenblatt (1966) similarity theory supported by the recent experimental evidence (Moloney and Field, 1991).

One of the most important problems is solving of open boundaries. We are using multi step calculations where open boundary conditions are taken from the results of bigger area calculations (Slides 3,4).

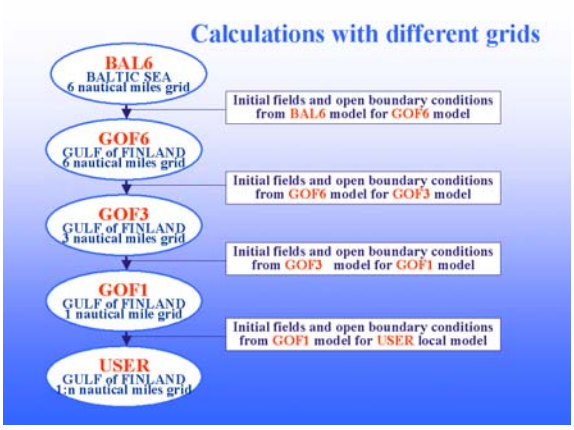
To take in the fish stock models into account the environmental conditions we will try include in our model also fish.



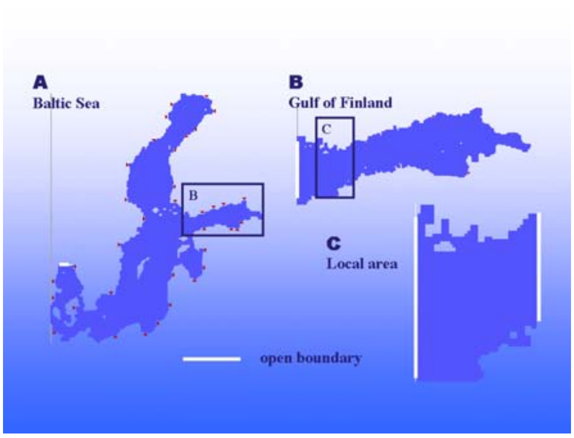
Slide1. Aquatic ecosystem model FinEst



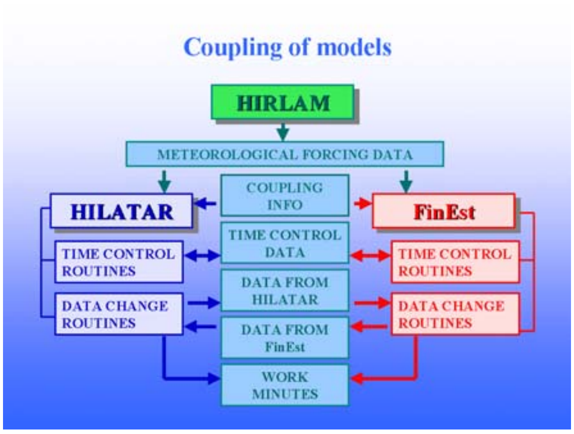
Slide2. Plankton community in FinEst



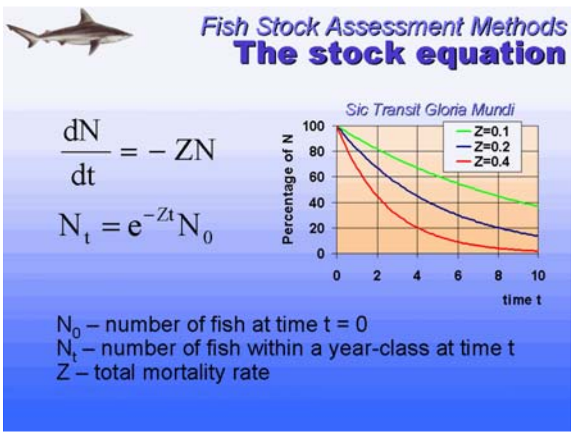
Slide3. Calculations with different grids



Slide4. Handling of open boundaries



Slide5. Coupling of models



Slide6. The fish stock equation

M. Garcia (Spain)

MARINE DATA MANAGEMENT WORKING GROUP

Report from Instituto Español de Oceanografía
Helsinki, Finland 17–19 April 2002
María Jesús García

During the last year, the work has been focus in two main purposes, the Mediterranean Data Archaeology and Rescued (**MEDAR**) project which has been finish in December 2001, and, the quality control of the historical sea level data.

As result of the MEDAR project in the IEO we have rescued and qualified not only the data from the IEO but also a very important data set from the Centro de Estudios Avanzados de Blanes (CEAB) which is one of the main Spanish centre involve in nutrients data. By the way as Western Mediterranean Regional Centre we have qualified the data from Algeria and Morocco, mainly hydro-chemical coastal bottle data. Most of the cruises have been codified as public availability and will be published in the MEDAR CD_ROM. Only some coastal data and other recent cruises still limited to the project. An inventory by country with general information of the cruise is presented at the **http:/www.ieo.es/medar/** The data set compiled by IEO (Spain), INRH (Morocco) and ISMAL (Algeria) during the MEDAR project is:

Type	Cruises	Profiles	Samples
Bottles	227	5,085	26,899
CTD	31	1,097	561,423
XBT	5	164	83,862

Concerning to the quality control, we have notice that the big variability of the chemical data, in particular in coastal stations, made difficult to qualified the data profile by profile and very fine references are needed. In order to have as much information as possible useful for quality control, statistics on ranges of variability have been perform.

In relation with the sea level data, we have qualified the complete data set (1994–2001) for the 3 tide gauge stations in Santander, Coruña and Vigo and we have perform a study of the monthly mean sea level variability for the period (1994–2000). The monthly mean sea level have been sent to the Permanent Sea Level Services. During quality control we have used the Standard Normal Homogeneity Test (SNHT) and intercomparison with the data from stations in the same oceanographic region as Northern Spanish Peninsular Coast, Mediterranean and Canary Islands.

Talking about the data from the project RADIALES, the IEO structural project around the Spanish coast. The number of stations carry out up to date is 3,542 but only some data are already included in the database **SIRENO.** Although, some of this data in the Mediterranean region are included in the MEDAR data set but only the hydro-chemical variables, not the biological ones.

In relation with developing software, new version of the software for quality control of oceanographic data for PC platform was developed. Source codes were rewritten and more friendly graphics user interface was introduced. Besides, some new modules were included both to convert data from external sources (spreadsheet ASCII formats) into MEDATLAS format for exchange and quality control according to the MEDAR/MEDATLAS II protocols (1) and to export data to flat format files compatible with widespread free and commercially available software, such as database managers, spreadsheets, graphics, statistics and other analysis packages.

All modules composing the QC system can be used as stand-alone applications. The QCMEDAR software will be included and distributed as a product within the final CD-ROMs of the MEDAR/MEDATLAS II MAST Concerted Action. It will also be available for downloading to MEDAR partners through IEO web/ftp server (www.ieo.es/medar).

MARINE DATA MANAGEMEMT WORKING GROUP
Report from Instituto Español de Oceanografia
Helsinki, Finland 17-19 April 2002

María Jesús García

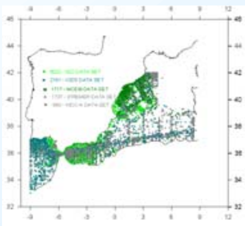
1. **Data set compiled during MEDAR project**
2. **Large variability of the chemical variables. Remarks on the QC.**
3. **Status of the Sea level Data**
   - **Tide gauge Network. Data Inventory**
   - **Sea level time series quality control**
   - **European Sea Level Service (ESEAS)**
4. **Status of the data management from RADIALES project**
5. **Software**

ICES Marine Data Management Working Group   Helsinki, 17-19 April 2002

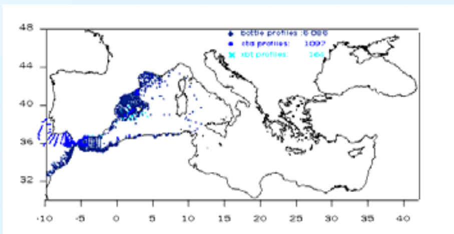
Activities related to MDM in 2001. Institute Español de Oceanografía

Data set compiled during MEDAR project

Data set from MEDATLAS I



New data set compiled by IEO, IHNR & ISMAL during MEDAR project



Type	Cruises	Profiles	Samples
Bottles	227	5,085	26,899
CTD	31	1,097	561,423
XBT	5	164	83,862

⊗ Many chemical data rescued from CEAB. Main Spanish institute involve in chemical data.
⊗Data quality control of the IHNR (Morocco) and ISMAL(Algeria)

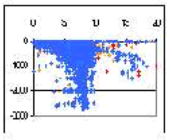
ICES Marine Data Management Working Group   Helsinki, 17-19 April 2002

Activities related to MDM in 2001. Institute Español de Oceanografía

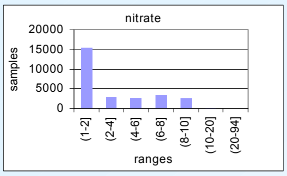
Large variability of the chemical variables. Remarks on the QC.

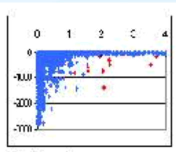
- Check spikes or gradients with large threshold values
- Large broad ranges in particular in coastal stations.

Ranges

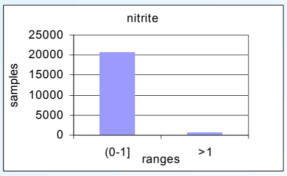


1. Nitrates	samples
(1-2]	15341
(2-4]	2965
(4-6]	2686
(6-8]	3361
(8-10]	2559
(10-20]	166
(20-24]	108





Nitrites	samples
(0-1]	20675
>1	665

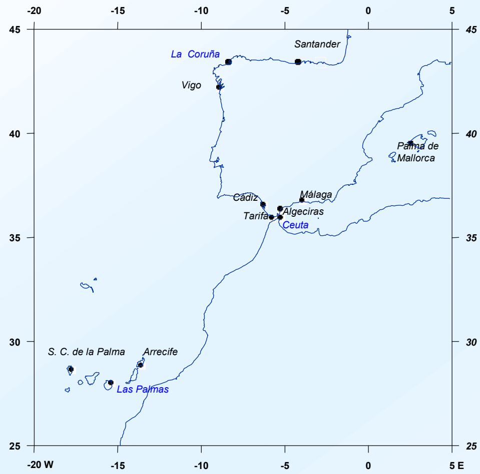


ICES Marine Data Management Working Group

Helsinki, 17-19 April 2002

Activities related to MDM in 2001. Institute Español de Oceanografía

Tide Gauge Network. Data Inventory



Stations	Years
Santander	1943-ongoing
* La Coruña	1943-ongoing
Vigo	1943-ongoing
Cadiz	1945-ongoing
Tarifa	1943-1961
	1963-1989
	1991-ongoing
Algeciras	1943-1955
	1961-ongoing
Malaga	1943-1959
	1961-ongoing
P. Mallorca	1963-1982
	1989-1993
	1996-ongoing
* Ceuta	1943-ongoing
Arrecife	1949-1975
	1980-ongoing
* Pto. Luz(Las Palmas)	1949-1956
	1971-1989
	1991-ongoing
S.C. Palma	1949-1989
	1997-ongoing
* GLOSS stations	

ICES Marine Data Management Working Group

Helsinki, 17-19 April 2002

Mean Sea Level along the Northern Iberian Peninsular Coast.

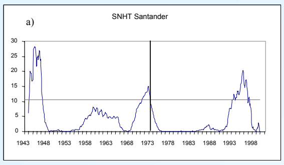
Sea level time series Quality Control — E.Tel & M.J García

•The **Standard Normal Homogeneity Test** gives the points where an inhomogeneity exists and provides information about the probable break magnitude. But this non-homogeneity could be due to an error or to an anomalous but real behaviour of the variable.

$$T_v = v \cdot (\overline{z_1})^2 + (n - v) \cdot (\overline{z_2})^2$$

$\overline{z_1}$ is the mean for the series from data point 1 to v,
$\overline{z_2}$ is the mean from v to n (the end of the series).

• The series are corrected only after comparison with other series in the same oceanic region and supported by historical information about the incidences on the tide gauge.



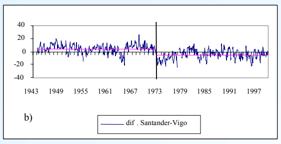


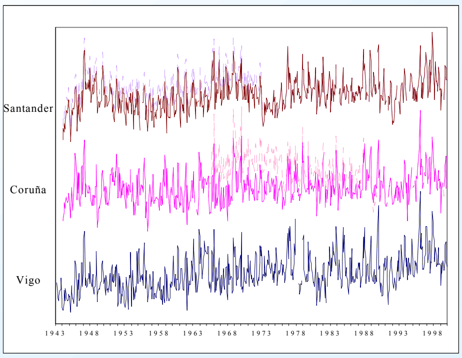
Fig. 2a. Test results for Santander.
2b. Diff. between Santander and Vigo anomalies.

ICES Marine Data Management Working Group

Helsinki, 17-19 April 2002

Activities related to MDM in 2001. Institute Español de Oceanografía

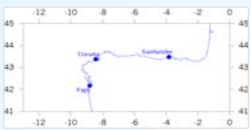
Long-term trend. Northern Iberian peninsular coast — E.Tel & M.J García



Original (dashed lines) and corrected (continuous lines) time series

a)	begin	end	n	mean	trend	trend**
				(cm)	(mm/y)	(mm/y)
Vigo	jan-43	dic-99	676	252.21	**2.66**	**2.88**
	jan-80	dic-89	119	254.09	5.41	5,77
	jan-90	dic-99	119	257.93	15.54	15.24
Coruña	jan-44	dic-99	660	263.69	**1.34**	**1.49**
	jan-80	dic-89	119	265.14	4.43	4.56
	jan-90	dic-99	111	266.99	10.6	10.44
Santander	jan-44	dic-99	651	278.63	**1.75**	**2.05**
	jan-80	dic-89	113	282.14	3.71	4.10
	jan-90	dic-99	109	281.66	9.86	10.08
Regional*	jan-44	dic-99	672	—	**2.08**	**2.32**

b)	Vigo	Coruña	Santander
Vigo	1,000		
Coruña	0,805	1,000	
Santander	0,724	0,750	1,000
(significant at 0,01)			



ICES Marine Data Management Working Group

Helsinki, 17-19 April 2002

Activities related to MDM in 2001. Institute Español de Oceanografía

European Sea Level Service (ESEAS)

Stablished in 2001 as final recommendation of European Sea Level Observing System (EOSS)

Major objectives:

- Co-ordinate the observation of sea level and derived product.
- Improve the exploitation of the available sea-level information.
- Implementation of the European component of the GLOSS

Members:

- Almost all the European Countries.

Information: http://www.gdiv.statkart.no/eseas/

ICES Marine Data Management Working Group

Helsinki, 17-19 April 2002

Activities related to MDM in 2001. Institute Español de Oceanografía

ESEAS - RI (European Sea Level Service Research Infrastructure

Proposal presented to the Fifth Framework Program

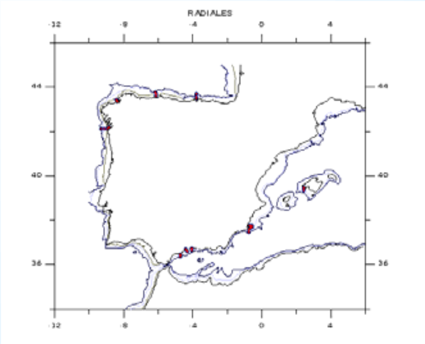
Workpackage

- Quality Control and sea leevel observation
- Absolute sea level variation
- Decadal to inter-decadal sea level variation
- Improving the sea level observing system
- Project management

ICES Marine Data Management Working Group

Helsinki, 17-19 April 2002

Status of the data management from RADIALES project



Time Series	Period	Cruises	Profiles
ECOMALAGA	1992-ongoing	29	261
ECOMURCIA	1996-ongoing	17	150
RADBAL	1994-ongoing	69	204
RADCOR	1989-ongoing	255	1070
RADGIJÓN	2001-ongoing	10	36
RADSANTANDER	1990-ongoing	167	890
RADVIGO	1987-ongoing	159	931

•This data are organized in the Relational Data Base "SIRENO": Not many and not all the variables
•The data from the Mediterranean (1-3) are included in the MEDAR data set . Hydro-chemical

Software

New version of QCMEDAR. Included in the MEDAR CD_ROM
A program to input data in Medatlas format and quality control according to the MEDATLAS protocols
This version is composed of several independent programs for WINDOWS (32 bits).

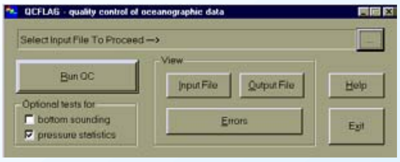
Main Modules:

- ***"Input Data"***
- ***"Quality Control":***
  - Check Format
  - Plot with statistics
- ***"Visualization"***
- ***"Output format"***
  - Med →ODV
  - FlatOutput"

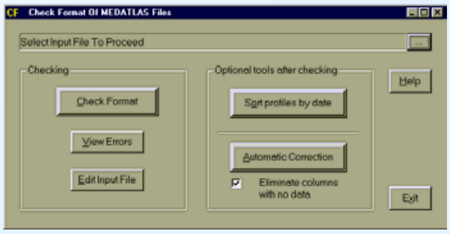
This version can be download from http://www.ieo.es/medar

Software. Quality Control Screenshots

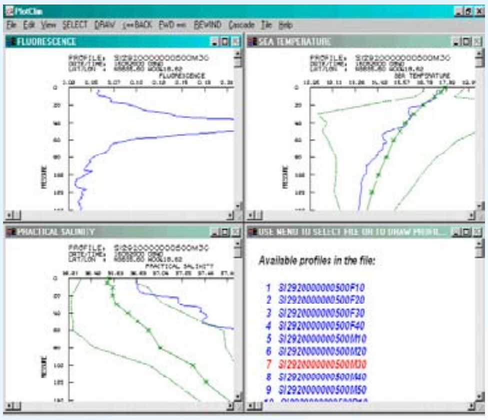
• ***"Quality Control"***



- *Check format*



-*Plot with statistics*

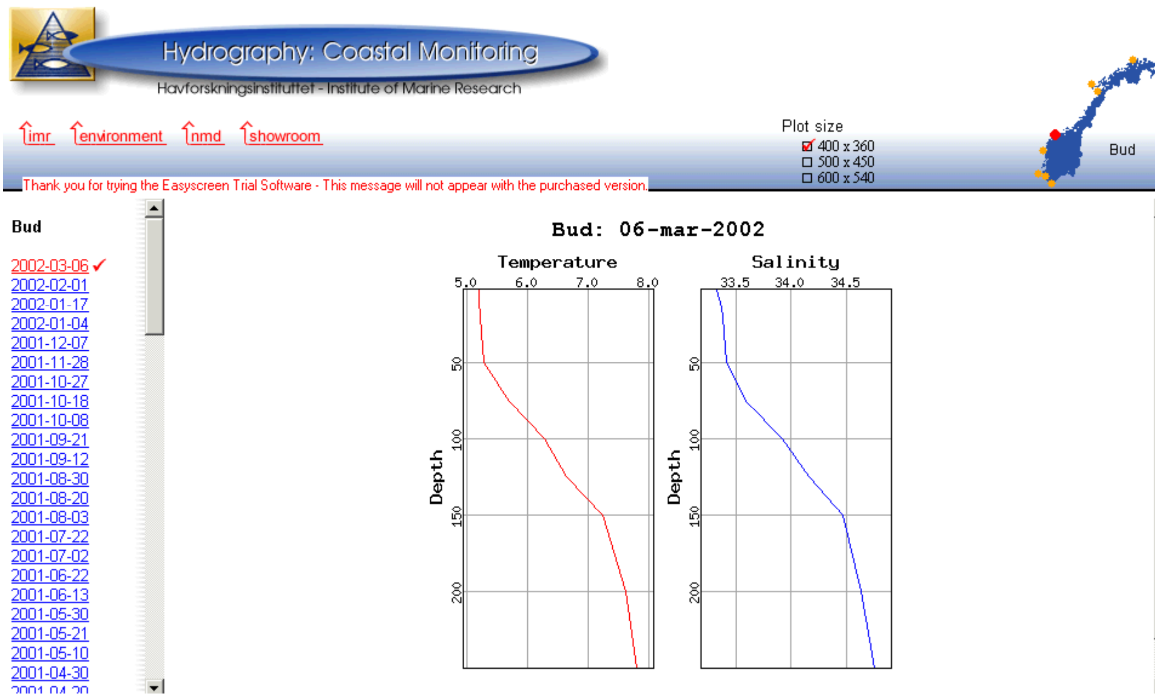


H. Sagen (Norway)

Fixed hydrographical stations along the Norwegian coast
ICES WGMDM Presentation, Helsinki, April 2002
Helge Sagen, Norway

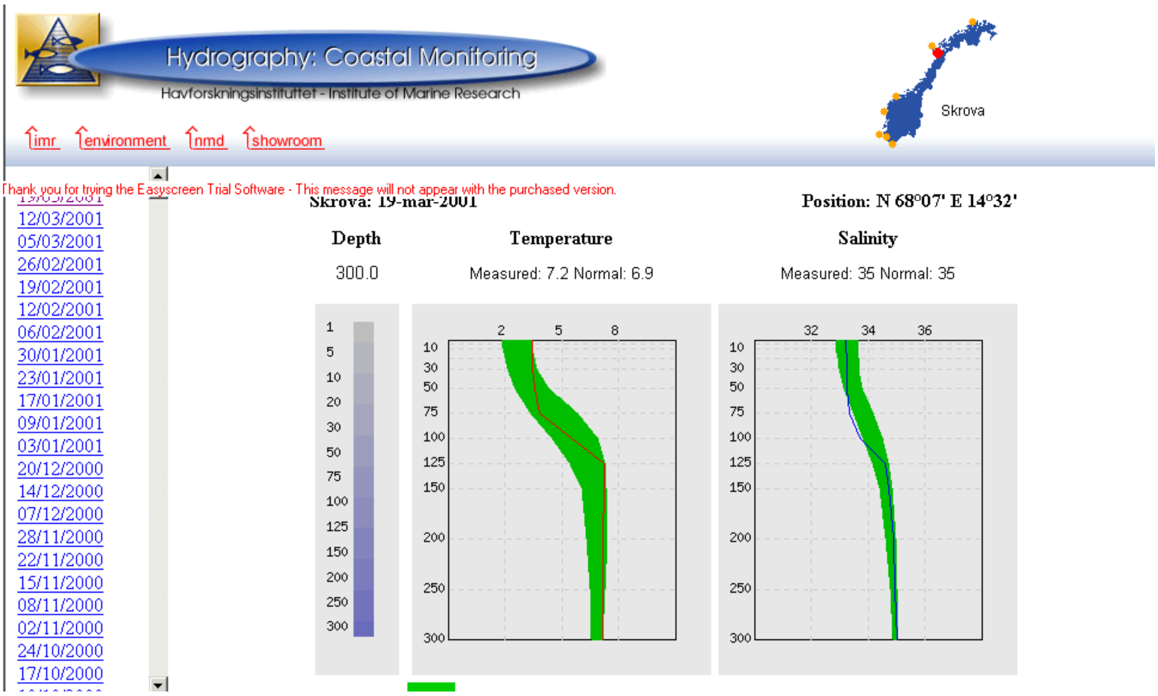
Along the Norwegian coast 8 fixed stations are operational. Local fishermen are operating the stations using their own private fishing boat every second week. The boat is equipped with a "mini-CTD" system measuring pressure, temperature and salinity. The instrument is developed by on of the researchers at the Institute of Marine Research. The data are transferred to the institute using regular telephone lines. The mini-CTD is docked in a communication box equipped with a modem. At the institute the data faces a quality control system and are being formatted into a locally defined format with measurements at standard depths.

The system has been operating since 1996 on a standalone PC running Windows NT. The first generation Pentium PC runs at 200 MHz (recently upgraded from 133 MHz) using the Microsoft Visual FoxPro 6 relational database system. The operational system is in 2002 facing the fact that it is growing to be outdated. The system was developed at the institute by on of the technical staff who ended his relationship with the institute in 2000. The system has been running without new development since then.



Profiles were presented graphically by using GIF on the fly generated in PERL which both are freely available. This part of the software was running on a Hewlett Packard Unix server, the official web server of the institute. Overnight one weekend the web server changed from Unix to Linux without any notice.

Facing these facts a small group was formed to look into developing a new interface to the data. The official web pages of the Institute of Marine Research were revised in 2001 and a new technical solution was recommended. The relational database system was to run on Windows 2000 using the Microsoft SQL server 7. The data and the web pages were to be stored in the database system and PHP was chosen as the scripting language.



The original web pages have been reconstructed using HTML and PHP. The measured data was dumped from Visual FoxPro and reloaded into SQL server. All activities were executed within one month. A lot of minor problems occurred, mainly due to lack of knowledge of the new software. In 2002 the new system is up and running in a prototype release.

The free software was successfully installed. The older graphical presentation system was tried ported to Windows, but was abandoned because of personal development style and lack of documentation. Graphical presentation of data is now being performed using simple graphic routines in PHP. The lack of a graphical library limits the available figures. It is now possible to extend the functionality of the presentation system. In the end of 2002 we again face the problem of finding funds for the technician who develop the new system.

J. Szaron (Sweden)

SMHI: Towards Operational Oceanography

Jan Szaron, SMHI Oceanographic Services

SMHIs Oceanographic Services runs a successful monitoring programme that covers eastern Skagerrak, eastern Kattegat, the Sound and the Baltic Proper. Personnel from the Services are also present on almost all cruises conducted by the National Board of Fisheries. In all this means that the major stations are visited at least once a month. Cruise reports, based on data subjected to a number of QC-procedures, are published 1–2 days after the cruise on the web. All data from on-board analysis are fed into the main database right after the completion of a cruise and hence are available for operational use (the duration of a routine monitoring cruise is 5–6 days). The next planned step is to transfer data via satellite to further decrease the time delay.

SMHI has started to reactivate and develop the SST/SSS-network. A thermosalinograph with a fluorescence sensor and an automatic water sampler has been installed on the R/V Argos and is now being tested. The system is designed to work on a number of identified and suitable ferry- and shipping lines and will send data in real time to SMHI.

To further strengthen SMHIs operational oceanography capacity two meteorological/oceano-graphical buoy moorings have been deployed (see Annex 1). The buoys will hopefully be fully operational during 2002.

Computer based models are now important instruments in operational oceanography. SMHI has devoted considerable resources to develop the Baltic HOME (Hydrology, Oceanography and Meteorology for the Environment) expert system. HOME will identify the processes involved when chemical substances are transported through air, soil and water. All environmental variables will be taken into consideration and treated. The system will supply information on existing conditions and form prognoses of physical and biochemical conditions in arbitrary points in the seas surrounding Sweden (see Annex 2). Baltic HOME is presented in detail on http://www.smhi.se/sgn0102/n0205/baltichome/baltichome_en.htm.

Annex 1

Towards a new real time data acquisition system for the Kattegat and the Baltic Sea Area

Karlson B., Borenäs, K., Broman B., Dahlin H., Henriksen J., Lindh H., and Nygren R.
Swedish Meteorological and Hydrological Institute

Fig 1. The buoy southeast of Läsö has just been deployed by R/V Argos.

Summary

In an effort to improve the Swedish oceanographic observation system two buoy moorings were deployed in May 2001 by SMHI, in co-operation with the Swedish Armed Forces. The buoys are located in the Kattegat (southeast of Läsö) and in the Baltic (east of Huvudskär) as indicated on the map.

The systems consist of a surface unit and an underwater part. A large number of oceanographic, as well as meteorological, variables are measured every hour and the data are transmitted via satellites in order to offer real time presentations.

The Swedish buoys will be part of an international network of oceanographic observational systems. The mooring in the Kattegat is also a Danish-Swedish contribution to the HELCOM COMBINE Programme (the Helsinki Commission monitoring programme), and thus partly financed by the Danish and Swedish Environmental Protection Agencies.

Fig. 2. The position of the buoys.

Technical description of the Läsö buoy

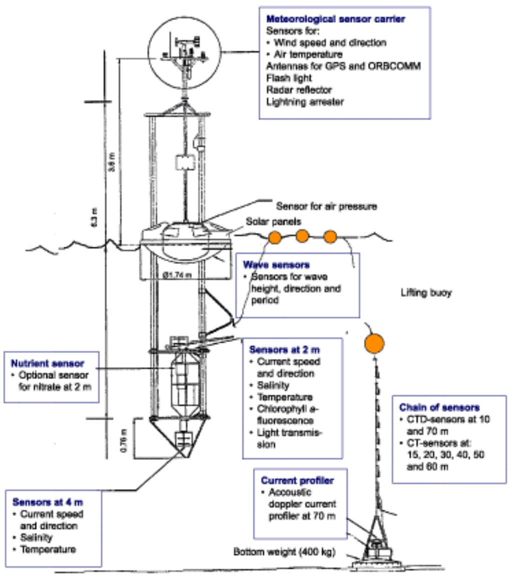


Fig. 3. Detailed description of the Seawatch buoy system.

Examples of time series

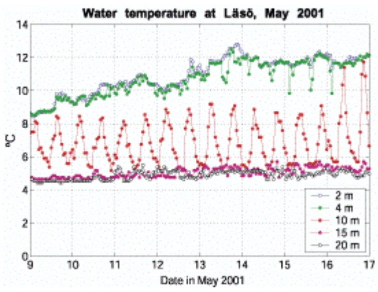


Fig. 4. The temperature at various depths during the period 9-17 May, 2001. The upper layer is warming up and the interface between the upper and lower layer seems to be oscillating at the tidal frequency.

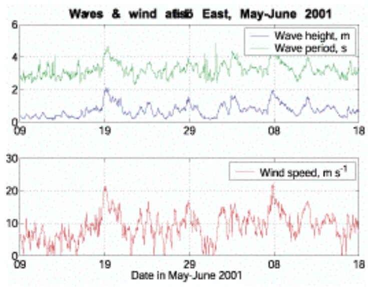


Fig. 5. A coupling between the wind speed, the wave height and the wave period is in evidence from measurements obtained during early summer 2001.

System for data transmission, storage and presentation

Data are transmitted every hour via the ORBCOMM satellite system.

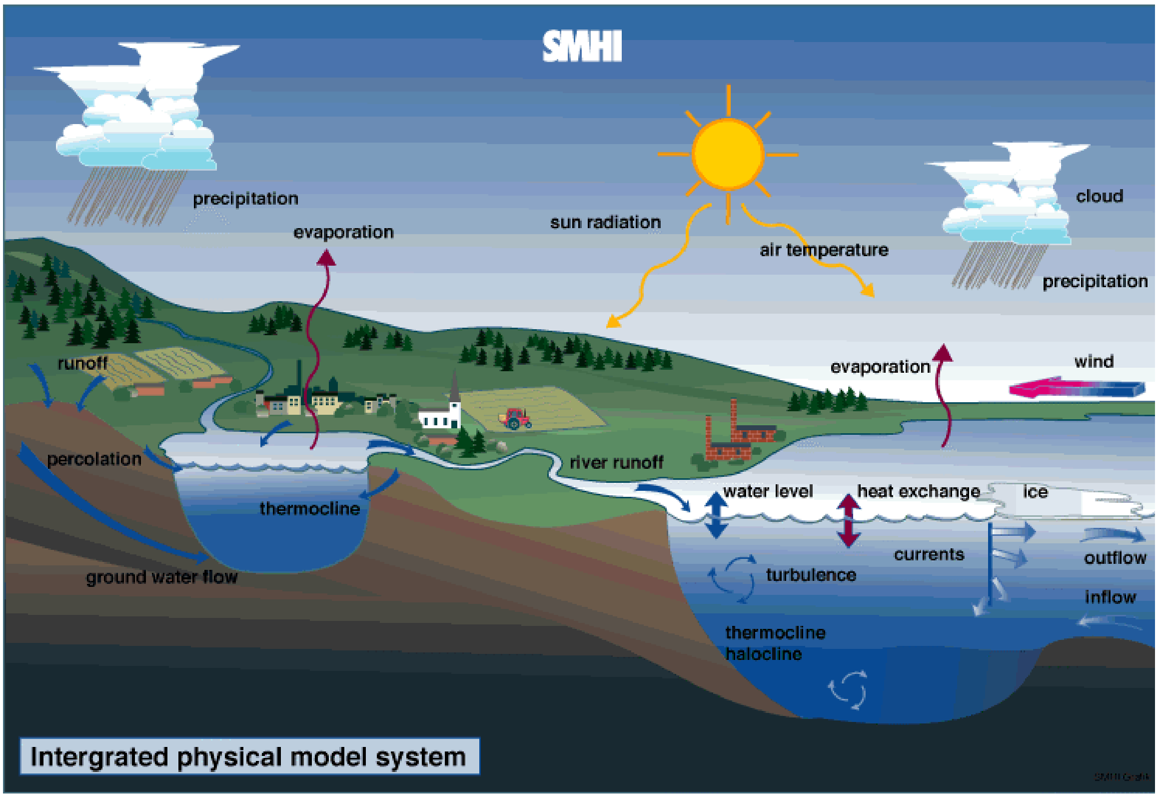
SMHI

A receiving station in Italy transfers the data to a database at SMHI for further publication on the web.

Contact
Bengt Karlson
SMHI, Oceanographic Services
Nya Varvet 31
SE-426 71 Västra Frölunda
Phone: + 46 31 751 8958
E-mail: bengt.karlson@smhi.se

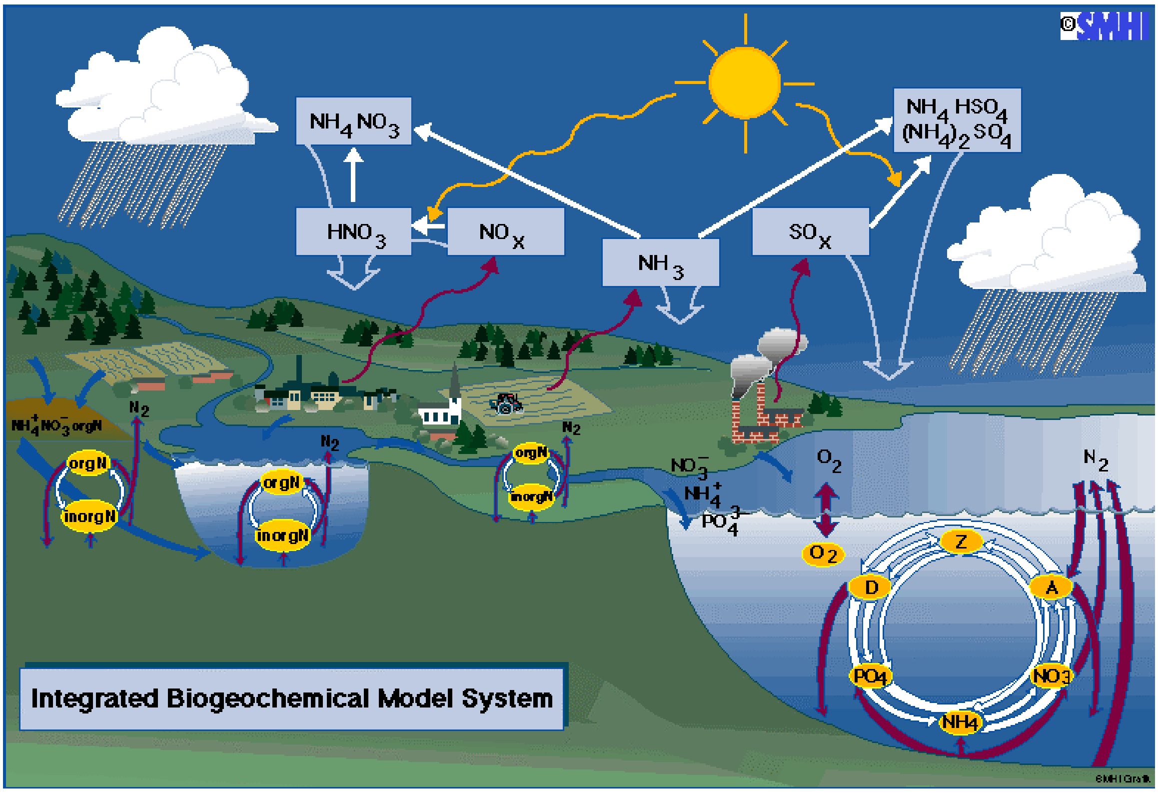
Annex 2

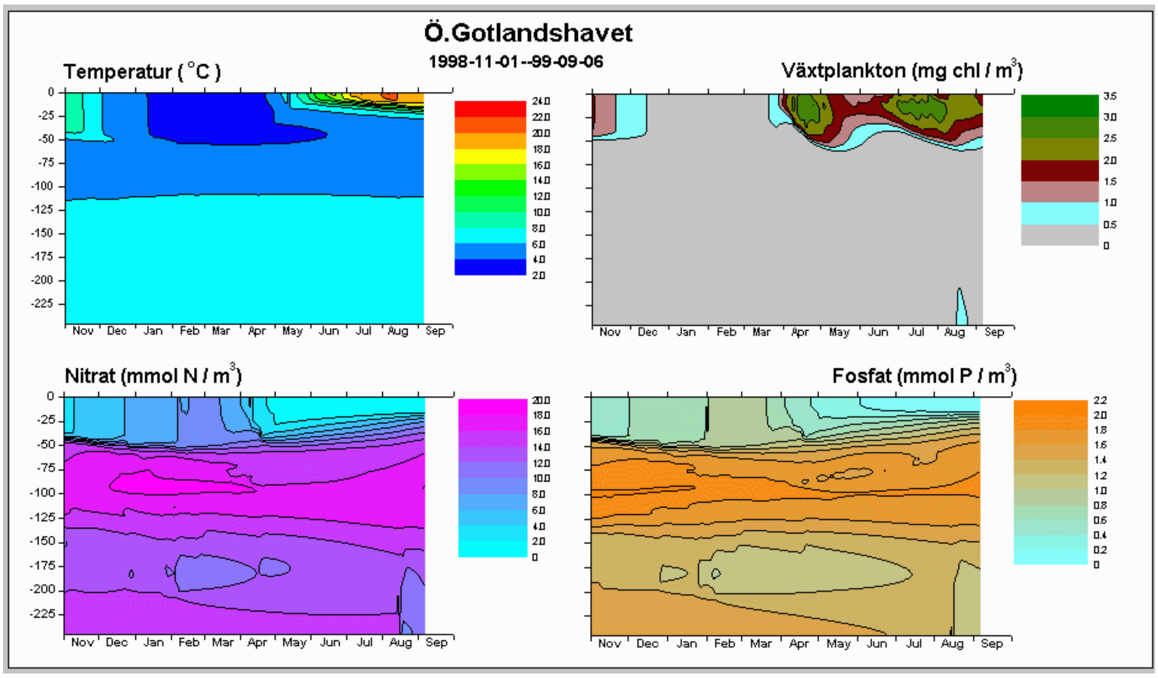
Baltic HOME



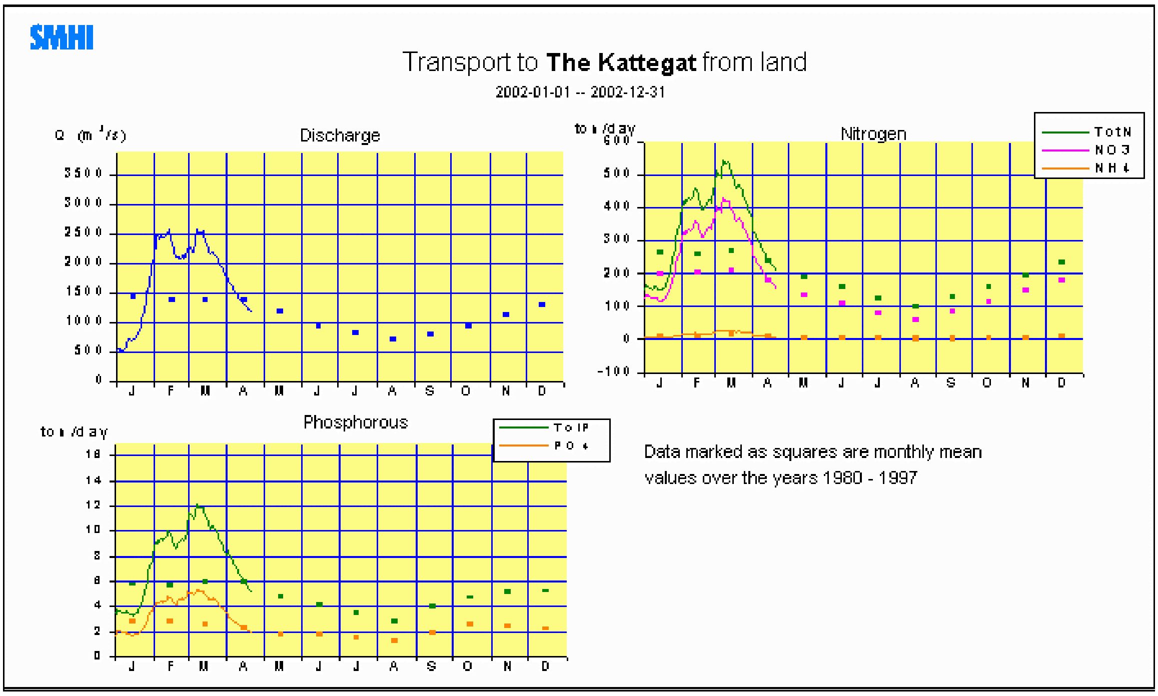
Annex 2

Baltic HOME





Ö.Gotlandshavet
1998-11-01--99-09-06
Temperatur (°C)
Växtplankton (mg chl / m³)
Nitrat (mmol N / m³)
Fosfat (mmol P / m³)
Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep



SMHI
Transport to The Kattegat from land
2002-01-01 -- 2002-12-31
Discharge
Q (m³/s)
Nitrogen
ton/day
TotN
NO3
NH4
Phosphorous
ton/day
TotP
PO4
J F M A M J J A S O N D
Data marked as squares are monthly mean values over the years 1980 - 1997

Edward Vanden Berghe (Belgium)

E. Vanden Berghe began by describing the work of the Flanders Marine Data Centre (VMDC). The data centre has six staff and relies on standard technologies such as SQL, Apache and Visual Basic. VMDC's area of interest is the Southern North Sea and in particular the Belgium continental shelf. They also retain a scientific interest in the Western Indian Ocean (off Kenya) and the Antarctica. Concerning habitats, VMDC's interests are both coastal and open waters and all marine sciences although all their staff are marine biologists.

Activities

The VMDC is accessible via a website (http://www.vliz.be/Vmdcdata/index.htm) and holds Belgium and international data. The Marine Information and Data Acquisition System (MIDAS), which captures data from research ships and databases and makes it available via the website, is under development and will be completed in less than a year's time. VMDC run a Monitoring Network of oceanographic and meteorological parameters for the Government that is aimed at increasing the safety of shipping. The International Marine Information System (IMIS) provides information on publications and projects. The Library module is a copy of OSPAR and can exchange data using their own version of MEDI allowing them to integrate datasets.

Another activity is Alphia. Aphia is species register with taxonomy and simple distribution to be used as reference literature. It is available via the web and contains 12,000 records with all North Sea families. Great care is needed with taxonomy as in the past different names have been used for the same species. Aphia stores the multiple names where they exist.

North Sea Benthos Survey

This was conducted by the ICES Benthos Ecology WG and was a sampling campaign with 9 cruises in April-May 1986 collecting from 235 stations. The result was 1004 taxonomy and 16500 distribution records. The shortcomings of this survey were that, although workshops were held to standardise taxonomy, different names still were used and checks could not be made as specimens were used to determine dry weight and thus not preserved.

North Sea Benthos Project

This new project is a collaboration with the Netherlands Institute for Sea Research (RIVO) and the Netherlands Institute of Ecology, Centre for Estuarine and Coastal Ecology (CEMO) who were involved with the previous survey. The object is to collect and correlate all Benthos Samples from the North Sea collected in 2000 and to set up a procedure for data exchange. Counties involved were Germany, Holland, Belgium, UK and France. Challenges faced included no obligation but just a common interest, no dedicated sampling, no standard methodology, no regular grid and no regular distribution. The project collected data at 1329 locations with 58,000 records of which 29,000 were in the Southern North Sea. It covered 1494 taxonomic names but the distribution density of records was varied.

There were many problems with differences in taxonomic names. Of the 1426 species, 814 occurred in one dataset and these numbers are after cleaning up spelling errors. The primary cause of this is different interpretations of the taxonomy. The solution to these problems is to flag the source used to identify the species and also who carried-out the identification in the databases. Workshops can also be used to harmonise identification.

A4.3 Day 3 Presentations:

L. Rickards (UK)

The UK Marine Environmental Data Network
Lesley Rickards

The UK Inter Agency Committee on Marine Science and Technology (IACMST) has set up an Action Group and Co-ordinator to improve communication between those government departments and agencies which have an interest in marine data. The aim of this group is to provide an enhanced system of networking amongst sources of marine environmental data. This network facilitates scientific research and underpins wealth creation. The initial core activities were to:

- To develop, maintain and make available inventories of data
- To develop guidelines for data management.
- To improve mechanisms to facilitate data exchange (including contributing UK data to global databases).

Over the last year there have been several new developments. Work was undertaken to produce a 'Climate of UK Waters at the Millennium - Status and Trends' Report. This was done in collaboration with a UK GOOS Action Group, which also operates under the auspices of IACMST. The Report was published by IACMST in its Information Document series in July 2001. It concentrates on the following parameters: temperature, salinity, sea level, waves, weather and climate and plankton, with some representative samples of nutrient data. The intention is that a further report might be produced in 2004, in collaboration with MPMMG and some preliminary discussions have taken place, and a meeting will be held in May to discuss this further.

The contents of the Report are available as a series of web pages from the OceanNET web site at www.oceannet.org/UKclimate-status/. An Adobe Acrobat version is also available for download from the Web site. The report is now being developed into a prototype web-based time series atlas, which could be expanded into a more comprehensive time series atlas for UK waters.

In addition, the Marine Environmental Data Network has undertaken a feasibility study for a UK Moored Current Meter Data Product. In the area 48°N to 62°N, 20°W to 10°E, the BODC Current Meter Inventory includes entries from 8035 current meter deployments. Of these 3680 are held by BODC, 1576 are not held and the remaining 2779 returned no data or were lost. The plan is to produce either a CD-ROM or a web-based product, or both. The product should provide access to the data files in a user-friendly way. Thus there needs to be some sort of selection tool - ideally selection from a map, but also the ability to select by date range, length of data series, water depth, organisation, etc. The data themselves need to be made available in easily accessible formats that can be used with other software packages and tools. This could either be ASCII or netCDF. In addition to the data files and accompanying documentation, statistics relating to the data can be complied. Various visual presentations of the data should also be available (time series, scatter plot, progressive vector diagrams). It is intended that work on this will be carried out over this year (i.e., 2002).

Finally, following on from discussion relating to UK coastal zone activities, a joint project under the British Government's Invest to Save Programme has been set up to Integrated Coastal Zone Map (ICZMap) project. The intention is to produce a coherent base map for the coastal zone going from land to sea, including bathymetric information in addition to land based contours, all to a common datum.

R. Gelfeld (USA)

World Ocean Database 2001

Robert D. Gelfeld
Ocean Climate Laboratory
National Oceanographic Data Center / NOAA
Silver Spring, MD

INTRODUCTION

The oceanographic databases described by this paper greatly expands on the *World Ocean Database 1998* (WOD98) product. We have expanded these earlier databases to include data from new instrument types such as profiling floats and new variables such as pCO_2 and TCO_2. Previous oceanographic databases including the NODC/WDC profile archives, and products derived from these databases, have proven to be of great utility to the international oceanographic, climate research, and operational environmental forecasting communities. In particular, the objectively analysed fields of temperature and salinity derived from these databases have been used in a variety of ways. These include use as boundary and/or initial conditions in numerical ocean circulation models, for verification of numerical simulations of the ocean, as a form of "sea truth" for satellite measurements such as altimetric observations of sea surface height, and for planning oceanographic expeditions. Increasingly nutrient fields are being used to initialise and/or verify biogeochemical models of the world ocean. The databases, and products based on these databases, are critical for support of international assessment programmes such as the Intergovernmental Programme on Climate Change (IPCC) of the United Nations.

It is well known that the amount of carbon dioxide in the earth's atmosphere will most likely double during the next century compared to CO_2 levels that occurred at the beginning of the Industrial Revolution. Regardless of one's scientific and/or political view of a possible “enhanced greenhouse warming” due to the increase of carbon dioxide, it is necessary that the international scientific community have access to the most complete historical oceanographic databases possible in order to study this problem, and other scientific and environmental problems.

The production of oceanographic databases is a major undertaking. Such work benefits from the input of many individuals and organizations. We have tried to structure the data sets in such a way as to encourage feedback from experts around the world who have knowledge that can improve the data and metadata contents of the database. It is only with such feedback that high quality global ocean databases can be prepared. Just as with scientific theories and numerical models of the ocean and atmosphere, the development of global ocean databases is not carried out in one giant step, but proceeds in an incremental fashion.

The *World Ocean Atlas 1994* (WOA94) represented the first database analysis product of the National Oceanographic Data Center (NODC) Ocean Climate Laboratory (OCL). WOA94 included vertical profiles of six variables including temperature, salinity, oxygen, phosphate, nitrate, and silicate and objective analyses of these variables at standard depth levels. *World Ocean Database 1998* (WOD98) updated WOA94 to include additional data for these six variables and additional variables such as chlorophyll, nitrite, pH, alkalinity and plankton. High resolution CTD (conductivity-temperature-depth) and high resolution XBT (expendable bathythermograph) profiles are also included. Products derived from this database, such as objective analyses of the variables that comprise WOD98 were made available as a separate atlas and CD-ROM series entitled *World Ocean Atlas 1998* (WOA98). *World Ocean Database 2001* (WOD01) now includes data from new instrument types such as profiling floats (P-ALACE, SOLO, APEX), Undulating Ocean Recorders (*e.g.,* towed CTDs), and Autonomous Pinniped Bathythermographs (instrumented Elephant Seals).

As with our previous work, users can obtain the latest information on WOD01 (*e.g.,* Errata sheet, Frequently Asked Questions and Updates) via the NODC Home Page, **http://www.nodc.noaa.gov/** (click on *Ocean Climate Laboratory*, then click on *World Ocean Database 2001*). The purpose of the WOD01 atlas series is to describe the database and show the historical distributions of profiles made using the various instrument types included in WOD01 some specific variables that comprise WOD01. This is accomplished through the display of global data distribution plots for individual years. These plots and the year-season data distribution plots are available in colour via the NODC Home Page. This provides users with basic information about the data in the historical ocean profile archives of NODC/WDC. In this atlas, “WDC” stands for the World Data Center for Oceanography, Silver Spring which is collocated with NODC. WDC, Silver Spring was formerly known as “WDC-A for Oceanography”. More information about the World Data Center System can be found at www.ngdc.noaa.gov/wdcnain.html.

a) Goals for *World Ocean Database 2001* (WOD01)

Our goal in developing and distributing WOD01 is to make available without restriction, the most complete set of historical ocean profile data and plankton measurements possible in electronic form along with appropriate metadata and quality control flags set for:

each individual observed level measurement,
"data" values at standard depth levels derived from vertical interpolation,
data representing an entire cruise.

As with earlier versions of NODC/WDC databases, we expect the data contained in WOD01 will find use in many different areas of oceanography, meteorology, and climatology. Whether studying the role of the ocean as part of the earth's climate system, conducting fisheries research, or managing marine resources, scientists and managers depend on observations of the marine environment in order to fulfil their mission. Oceanography is an observational science. Because of the importance of understanding climate variability, in attempting to forecast both natural and anthropogenic climate variations, it is necessary to study the role of the ocean as part of the earth's climate system (IPCC (1996); WCRP (1995)).

It is important to note that WOD01 is a product based on data submitted to NODC/WDC by individual scientists and scientific teams and institutional, national, and regional data centres. A major contribution of NODC/WDC to the field of oceanography has been to provide centralized databases where all data and metadata are in the same format. This has allowed investigators such as Wyrtki (1971) and Levitus (1982) to construct atlases that have proven to be of utility to the scientific research and the operational forecasting communities.

b) Data and instrument (probe) types in WOD01

WOD01 consists of profile data from several oceanographic instrument (probe) types. We present a brief description of some of the major instrument types and/or systems that are (or were) used to make measurements which are included in WOD01. NODC instrument codes are presented by Conkright *et al.* (2002) and are also available via the NODC Home Page.

A description of various oceanographic instruments may be found in the recent publication by Emery and Thomson (1997). Detailed descriptions of instruments and measurement techniques can be found in the scientific literature, some of these sources are given in the bibliography of this atlas.

Ocean Station data (OSD)

Ocean Station Data (OSD) has historically referred to measurements made from a stationary research ship using reversing thermometers to measure temperature and making measurements of other variables such as salinity, oxygen, nutrients, chlorophyll, etc. on seawater samples gathered using special bottles. The two most commonly used bottles are the Nansen and Niskin bottles. Data that are in the OSD files are frequently referred to as "bottle data" and the entire collection of data from these file may be alternatively referred to as the "Bottle Data File". WOD01 includes measurements of temperature, salinity, oxygen, nitrate, phosphate, silicate, pH, alkalinity, chlorophyll, pCO_2, TCO_2 and plankton.

ii) Conductivity-Temperature-Depth data (CTD)

Conductivity-Temperature-Depth (CTD) instruments measure temperature and conductivity as a function of pressure (depth) at relatively high (often referred to as "continuous") vertical resolution. Salinity is computed from the conductivity measurement. CTD data may be submitted to NODC/WDC at sub-meter vertical resolution. These data are now archived at this resolution whereas in the past, electronic storage limitations resulted in only selected levels being stored. An earlier version of the CTD instrument was the STD (salinity-temperature-depth) which computed salinity from a conductivity sensor as the instrument was moving though the water column. Because of instrument problems that led to erroneous data values (spikes), this method was replaced by the CTD method for which conductivity measurements are recorded from the instrument and then salinity computed with appropriate calibration information. Dissolved oxygen content and chlorophyll can now be measured "continuously" with sensors placed on CTD instruments. New sensors are being developed to make "continuous" measurements of other variables. We refer to CTD "stations" or "casts" to recognize that more than one variable is being measured when a CTD instrument is deployed.

iii) Mechanical Bathythermograph data (MBT)

Mechanical Bathythermograph (MBT) instruments were developed in their modern form around 1938 (Spilhaus, 1938). The instrument provides estimates of temperature as a function of depth in the upper ocean. Earlier versions of the instrument were limited to making measurements in the upper 140 m of the water column. The last U.S. version of this instrument reached a maximum depth of 295 m. A temperature profile as a function of depth is traced on a smoked glass slide which is digitised. Pressure is determined from a pressure sensitive tube known as a Bourdon tube. One advantage of the MBT compared to lowering a reversing thermometer is that the MBT could be dropped from a moving ship and then winched aboard ship again. The accuracy of the MBT instrument is generally acknowledged to be about 0.5°F (0.3°C) (Levitus *et al.*, 1998; Boyer *et al.*, 2002). The Digital Bathythermograph (DBT) instrument is a version of the MBT that reports data electronically rather than mechanically and may reach depths deeper than 295 m. DBT profiles are included in the MBT files.

Expendable Bathythermograph data (XBT)

The Expendable Bathythermograph (XBT) was deployed beginning in 1966 and has replaced the MBT in many measurement programmes. There are different models of XBT instruments which have different maximum depth penetration and/or other different characteristics. The T-4, T-6, and T-7 probes reach maximum depths of 450, 750, and 750 m respectively. The T-7 probe differs from the T-6 probe in that it can be dropped from a faster moving ship and still maintain certain accuracy standards. The T-5 probe reaches a maximum depth of about 1800 m.

The depth of a temperature measurement from an XBT instrument is determined using the time elapsed between when the probe enters the water and the time each temperature measurement is made. A vendor supplied drop-rate equation is utilized. However, the vendor supplied drop-rate equation for T-4, T-6, and T-7 probes was found to have a systematic error and a new equation has been developed by the international research community (Hanawa *et al.*, 1995; UNESCO, 1994). The recommended practice regarding exchange and archiving of XBT profiles is that XBT profile data be exchanged or sent to data centres without correction for the systematic depth error, until an "international mechanism is established to implement the general use of the new equation" (UNESCO, 1994). This policy is to avoid double corrections.

v) Moored Buoy data (MRB)

WOD01 includes moored buoy data from the TAO (Tropical Atmosphere-Ocean, central and eastern Pacific), TRITON (Japan-JAMSTEC, western Pacific), PIRATA (tropical Atlantic), and MARNET arrays. We have included only real-time MRB observations reported over the Global Telecommunication System (GTS) and stored in the Global Temperature Salinity Profile Programme (GTSPP) database. Temperature and/or salinity are reported by these instruments. For a description of the TAO buoy network we refer to the work by Hayes *et al.* (1991) and McPhaden (1993, 1995).

Profiling Float data *(PFL)*

WOD01 includes data from profiling floats which drift at subsurface levels and are pre-programmed to rise to the sea surface at pre-set intervals and record temperature and/or salinity during their ascent. Float types include Profiling Autonomous Lagrangian Circulation Explorer (P-ALACE); PROVOR (free-drifting hydrographic profiler, www.ifremer.fr/coriolis), SOLO (Sounding Oceanographic Lagrangian Observer), and APEX (Autonomous Profiling Explorer). We have included only real-time PFL observations reported over the Global Telecommunication System (GTS) and stored in the Global Temperature Salinity Profile Programme (GTSPP) database (www.nodc.noaa.gov).

Drifting Buoy *(DRB) data*

WOD01 includes data from buoys that drift at the sea surface with thermistor chains extending vertically from them that record temperature. We have included only real-time observations DRB reported over the Global Telecommunication System (GTS) and stored in the Global Temperature Salinity Profile Programme (GTSPP) database.

Surface only data (SUR)

WOD01 includes "surface only" (SUR) data from ship-of-opportunity programmes (SOOP) and/or research cruises. Variables in the Surface Only Data Files include Temperature, Salinity, pH, Chlorophyll, Alkalinity, pCO_2, and TCO_2.

Undulating Ocean Recorder data (UOR)

WOD01 includes Undulating Ocean Recorder (UOR) data from instruments (probes) mounted on a towed undulating vehicle. Data from the TOGA, JGOFS, PRIME and OMEX projects are included. Variables included are: temperature, salinity, oxygen, chlorophyll.

Autonomous Pinniped Bathythermograph (APB)

Marine scientists have recently initiated programmes to instrument marine mammals to record environmental data and transmit these data via satellite to on-shore receivers. WOD01 includes temperature from Elephant Seals instrumented with time-temperature-depth recorders (TTDR) and ARGOS satellite platform terminal transmitters. Details of such data included in WOD01 can be found in the work of Boehlert *et al.* (2001).

c) Economic justification for maintaining archives of historical oceanographic data: the value of stewardship

Oceanography is an observational science and it is not possible to replace historical data that have been lost. From this point of view, historical measurements of the ocean are priceless. However, in order to provide input to a "cost-benefit" analysis of the activities of oceanographic data centres and specialized data rescue projects, we can estimate the costs incurred if we wanted to resurvey the world ocean today, in the same manner as represented by the WOD01 Ocean Station Data (OSD) profile archive.

The computation we describe was first performed in 1982 by Mr Rene Cuzon du Rest, of NODC. We use an average operating cost of $20,000. per day for a medium-sized U.S. research ship with a capability to make two "deep" casts per day or 10 "shallow" casts per day. We define a "deep" cast as extending to a depth of more than 1000 m and a "shallow" cast as extending to less than 1000 m. This is an arbitrary definition but we are only trying to provide a coarse estimate of replacement costs for this database. Using this definition, WOD01 contains approximately 1.8 million shallow casts so that the cost of the ship time to perform these measurements is approximately $3.6 billion. In addition WOD01 contains 323,000 profiles deeper than 1000 m depth, so the cost in ship time to make these "deep" measurements is approximately $3.2 billion. Thus, the total replacement cost of the OSD archive is about $6.8 billion, a figure based only on ship-time operating costs, not salaries for scientists or any other costs.

Data fusion

It is not uncommon in oceanography that measurements of different variables made from the same sea water samples, are often maintained as separate databases by different principal investigators. In fact, data from the same oceanographic cast may be located at different institutions in different countries. From its inception, NODC recognized the importance of building oceanographic databases in which as much data from each station and each cruise as possible are placed into standard formats, accompanied by appropriate metadata that make the data useful to future generations of scientists. It was the existence of such databases that allowed the *International Indian Ocean Expedition Atlas* (Wyrtki, 1971) and *Climatological Atlas of the World Ocean* (Levitus, 1982) to be produced without the time-consuming, laborious task of gathering data from many different sources. Part of the development of WOD01 has been to expand this data fusion activity by increasing the number of variables that NODC/WDC makes available as part of standardized databases.

e) Distribution media

WOD01 is being distributed on-line (www.noaa.nodc.gov) and on CD-ROMs with all data compressed in DOS format. Based on requests by users of our earlier products, the OCL developed a new ASCII format to make the most efficient use of space on storage media used to transfer data to users. To further minimize storage space requirements, the data have been compressed with the GZIP utility (Conkright *et al.*, 2002). Even with compression, there are seven CD-ROMs containing all profile data in WOD01 at observed levels and one CD-ROM containing all profile data in WOD01 at standard levels. Without compression, the number of CD-ROMs required to distribute the WOD01 database would total about sixteen.

f) Application software interfaces

We have included software conversion routines so that users of software packages, databases, and programming languages such as MATLAB, IDL, PC-Surfer, C, and FORTRAN can access the data in WOD01. An effort is in progress to develop a JAVA based interface for viewing data from the WOD01 CD-ROMs. In response to user requests, we have defined the WOD01 format to be as “self defining” as possible so as to eliminate, or at least minimize, the need for any structural changes to the format when new data or instrument types are added or increases in data precision occur. We do not envision any substantial changes to our present data format. We will use the Internet to make available additional converters, Graphical User Interfaces (GUI), and other software tools that become available.

2. COMPARISON OF WOD01 WITH PREVIOUS GLOBAL OCEAN PROFILE DATABASES

Table 1 shows the amount of data available from different instrument (probe) types that were used in earlier global oceanographic analyses. During the past three years, the archives of historical oceanographic data have grown due to special data management and data observation projects and also to normal submission by scientists and operational ocean monitoring programmes. With the distribution of WOD01 there are now approximately 7.1 million temperature profiles and 1.5 million salinity profiles (and other profile data and plankton data) available to the international research community in a common format with associated metadata and quality control flags. There has been a net increase of about 1.7 million temperature profiles since publication of *World Ocean Database 1998.*

Table 2 shows a comparison of the total number of Ocean Station Data variables at the sea surface with previous databases.

3 DATA SOURCES

The oceanographic data that comprise WOD01 have been acquired through many sources and projects and from individual scientists. Some of the international data exchange organizations are:

a) IOC Global Oceanographic Data Archaeology and Rescue Project;
b) World Ocean Database Project;
c) IOC Global Temperature-Salinity Profile Programme;
d) ICES;
e) International Research Projects data - including MEDAR/MEDATLAS, JGOFS, WOCE, BODC Project data sets (i.e., OMEX, BOFS);
f) Declassified naval data sets;
g) Integrated Global Ocean Service - Volunteer Observing Ship programmes
h) NOAA Ship-of-Opportunity Program (SOOP);
i) SURTROPAC;
j) Underway CO_2.

4 QUALITY CONTROL FLAGS

Each individual data value and each profile in WOD01 has quality control flags associated with it. A description of these flags and general documentation describing software to read and use the WOD01 database are found in the report by Conkright *et al.* (2002). WOD01 now includes Quality Control Flags assigned by principal investigators. Users can choose to accept or ignore these flags. It is clear that there are both Type I and Type II statistical errors (for normal distributions) associated with these flags. There are some data that have been flagged as being questionable or unrepresentative when in fact they are not. There are some data that have been flagged as being “acceptable” based on our tests which in fact may not be the case. In addition, the sparsity of data, non-normal frequency distributions, and presence of different water masses in close proximity results in incorrect assignment of flags.

The obvious advantage of flagging data is that users can choose to accept or ignore all or part of the flags we assign to data values. The most important flags we set are those that are set based on unusual features produced during objective analyses of the data at standard levels. This is because standard statistical tests may be biased for the reasons described above. Data from small-scale ocean features such as eddies and/or lenses are not representative of the large-scale permanent or semi permanent features we attempt to reproduce with our analyses and will cause unrealistic features such as bull's-eyes to appear. Hence, we flag these data, and other data that cause such features, as being unrealistic or as questionable data values. It is important to note that an investigator studying the distribution of mesoscale features in the ocean will find data from such features to be the signal he/she is looking for. As noted by Levitus (1982), it is not possible to produce one set of data analyses to serve the requirements of all possible users. A corollary is that it is not possible to produce one set of quality control flags for a database that serve the exact requirements of all investigators. As data are added to a database, investigators must realize that flags set for having violated certain criteria in an earlier version of the database may be reset solely due to the addition of new data which may change the statistics of the region being considered. Even data that have produced unrealistic features may turn out to be realistic when additional data are added to a region of sparse data. Conkright *et al.* (1994b) present the objectively analysed field of silicate at 1000 m depth using all silicate data available as part of WOA94 and using only data flagged as being acceptable. The differences are obvious.

5. OUTLOOK FOR FUTURE ACQUISITIONS OF HISTORICAL OCEAN PROFILE AND PLANKTON DATA AND INTERNATIONAL COOPERATION IN A "GLOBAL OCEAN DATABASE PROJECT"

Substantial amounts of historical ocean data continue to be transferred to NODC/WDC for inclusion into databases. The outlook for continuing to be able to increase the amount of such data available to the scientific community is excellent. Based on the positive results of the IOC/GODAR project and the Global Ocean Database Project, we have requested the continued cooperation of the international scientific and data management communities in building the historical ocean data archives. There is a particular need for high resolution CTD data so that we can resolve smaller scale features in the vertical and thus provide objective analyses of variables at greater vertical resolution than present. There is a need for additional historical chlorophyll, nutrient, oxygen, and plankton data so we can improve understanding of ocean biogeochemical cycles.

Improving the quality of historical data and their associated metadata is a important task. Corrections to possible errors in data and metadata is best done with the expertise of the principal investigators who made the original observations, the data centre or group that prepared the data, or be based on historical documents such as cruise and data reports (however, one has to also consider that these documents may contain errors). The continuing response of the international oceanographic community to the GODAR project and the Global Ocean Database Project have been excellent. This response has resulted in global ocean databases that can be used internationally without restriction for the study of many environmental problems.

As the amount of historical oceanographic data continues to increase as a result of international cooperation, the scientific community will be able to make more and more realistic estimates of variability and be able to place confidence intervals on the magnitude of temporal variability of the more frequently sampled variables such as temperature.

Table 1. Comparison of the number of stations in WOD01 with previous world ocean databases.

Data type	Climatological Atlas of the World Ocean (1982)	WOA94	WOD98	WOD01
Station data and low resolution C/STD casts	425,000	1,194,407	1,373,440	2,121,042
High resolution CTD casts	na	89,000	189,555	312,344
MBT profiles	775,000	1,922,170	2,077,200	2,376,206
XBT profiles	290,000	1,281,942	1,537,203	1,743,592
Moored Buoys	na	na	107,715	297,936
Drifting Buoys	na	na	na	50,549
Profiling Floats	na	na	na	22,637
Undulating Oceanographic Recorders	na	na	na	37,631
Autonomous Pinniped Bathythermograph	na	na	na	75,665
Total Stations	1,490,000	4,487,519	5,371,525	7,037,062
Surface only data* (cruises)	na	na	na	4,743*

* Surface data are represented differently in the database - all observations in a single cruise have been combined into one "station" with depth, value of variable measured and latitude, longitude, and Julian year day to identify data and position of individual observations.

Table 1a. Instrument types in the WOD01.

DIRECTORY	SOURCE
OSD	Bottle, low resolution *Conductivity-Temperature-Depth* (CTD), and plankton data
CTD	High resolution *Conductivity-Temperature-Depth* (CTD) data
MBT	Mechanical Bathythermograph (MBT) data
XBT	Expendable (XBT) data
SURF	Surface only data
APB	Autonomous Pinniped Bathythermograph - Time-Temperature Depth recorders attached to elephant seals
MRB	Moored buoy data from TAO (Tropical Atmosphere-Ocean), PIRATA (moored array in the tropical Atlantic), MARNET, and TRITON (Japan-JAMSTEC)
PFL	Profiling float data from Profiling Autonomous Lagrangian Circulation Explorer (P-ALACE) subsurface drifting floats; PROVOR (free-drifting hydrographic profiler), SOLO (Sounding Oceanographic Lagrangian Observer), and APEX (Autonomous Profiling Explorer)
DRB	Drifting buoy data from surface drifting buoys with thermister chains
UOR	Undulating Oceanographic Recorder data from a Conductivity/Temperature/Depth probe mounted on a towed undulating vehicle;

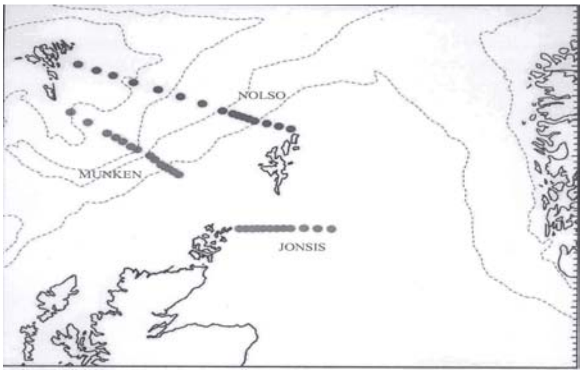
Table 2. Comparison of the number of sea surface observations in WOD01 of several Ocean Station Data variables with previous databases.

Variable	CLIMATOLOGICAL ATLAS OF THE WORLD OCEAN	WOA94	WOD98	WOD01
Temperature	425,000	1,194,000	1,439,209	1,951,170
Salinity	399,429	1,034,091	1,343,580	1,767,283
Oxygen	159,016	324,627	480,718	586,277
Phosphate	na	171,064	279,011	373,141
Silicate	na	80,235	186,226	261,774
Nitrate	na	61,817	144,523	208,573
pH	na	na	103,338	130,863
Alkalinity	na	na	6,759	22,268
chlorophyll	na	na	131,690	128,558
Plankton	na	na	na	148,243
pCO_2	na	na	na	2,159
TCO_2	na	na	na	6,018
Nitrate+Nitrite	na	na	na	9,382
Pressure	na	na	na	57,748

G. Slesser (UK)

Management of Long Term Monitoring Data
ICES WGMDM Presentation, Helsinki, April 2002
G. Slesser, UK

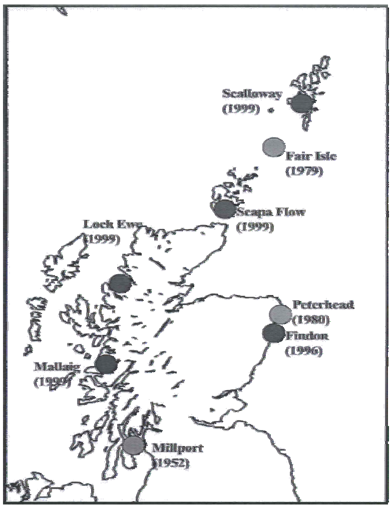
Long term monitoring carried out by the Marine Laboratory Aberdeen began in the early 1890's by the forerunner of the laboratory, the Fishery Board of Scotland (est. 1882). The first water bottle casts were carried out during 1893 and conducted by Dr H N Dickson on board HMS Jackal. Four of these stations were to become part of the now standard Nolso-Flugga Faroe-Shetland Channel section (Figure 1). In addition, at positions further south, he sampled at three stations and these were to become part of the standard Fair Isle-Munken section (Figure 1). A full set of stations (12) were first sampled along the Nolso-Flugga line in 1903, and since then have more or less been sampled annually or more except for the war years.



More recently, early to mid seventies, as part of the Joint Oceanographic North Sea Data Acquisition Project (JONSDAP), the first of a line of stations were carried out east of Orkney and this line is now commonly known as the JONSIS (Joint Oceanographic North Sea Information System) line (Figure 1). This line has also been sampled annually or more over the past 30 years.

Figure 1

At present, the Marine Laboratory samples these three sections three times a year, spring, autumn and winter. The data collected using CTD's and water sampling bottles are processed and entered into the FRS Hydrographic Station data base. In addition "mini" databases containing standard depth data for these data have been set up and are known as MUNKEN, NOLSO and JONSIS.



Latterly, when the idea of producing an Annual Scottish Ocean Climate Status Report was first mooted it soon became apparent that there were very little coastal time series of data available.

One of the longest, being sea temperature data being recorded daily since 1952 at the University Marine Station at Millport (Figure 2). Other sites where coastal sea temperature were being recorded are at Fair Isle, daily since 1979, and Peterhead, monthly/daily since 1980 (Figure 2). Therefore, five sites have been set up to give a wider coverage of Scotland. These are at Scalloway, Scapa Flow, Loch Ewe, Mallaig and Findon (Figure 2). Sampling was initiated during 1999. In addition to temperature being sampled at these sites by minilogger every 30 or 60 minutes, salinities and nutrients are being sampled weekly at Scalloway, Scapa Flow and Mallaig. Also, initiated in 2000 was the annual sea surface sampling of temperature, salinity and nutrient data every three miles around the coast of Scotland. At present the data from the coastal sampling sites and sea surface sampling positions are down loaded, processed, analysed and hence archived into the "Coastal" database.

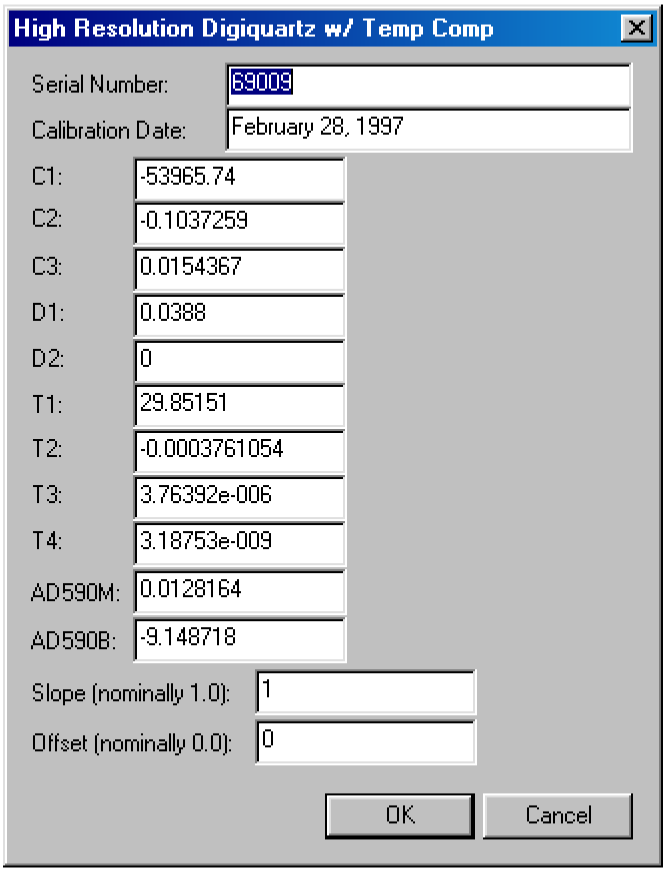
Figure 2

A. Isenor (Canada)

Automatic Generation of Seabird Configuration Files (CON Files)
ICES WGMDM Presentation, Helsinki, April 2002
A.W. Isenor, Canada

Seabird Electronic Inc. is a major manufacturer of ocean instrumentation. For many of the Seabird instruments, the conversion from frequency space to engineering units is conducted by applying a mathematical expression with multiple coefficients. The manufacturer, in a standard laboratory procedure, typically sets these coefficients. Within the Seabird software system, the coefficients are used for various computations during data acquisition and processing.

Within the Seabird environment, the coefficients are stored in a Seabird configuration file, more commonly known as a CON file. The CON file is accessible via Seabird software modules such as SEASAVE. The SEASAVE display of coefficients for a digiquartz pressure sensor is shown below.

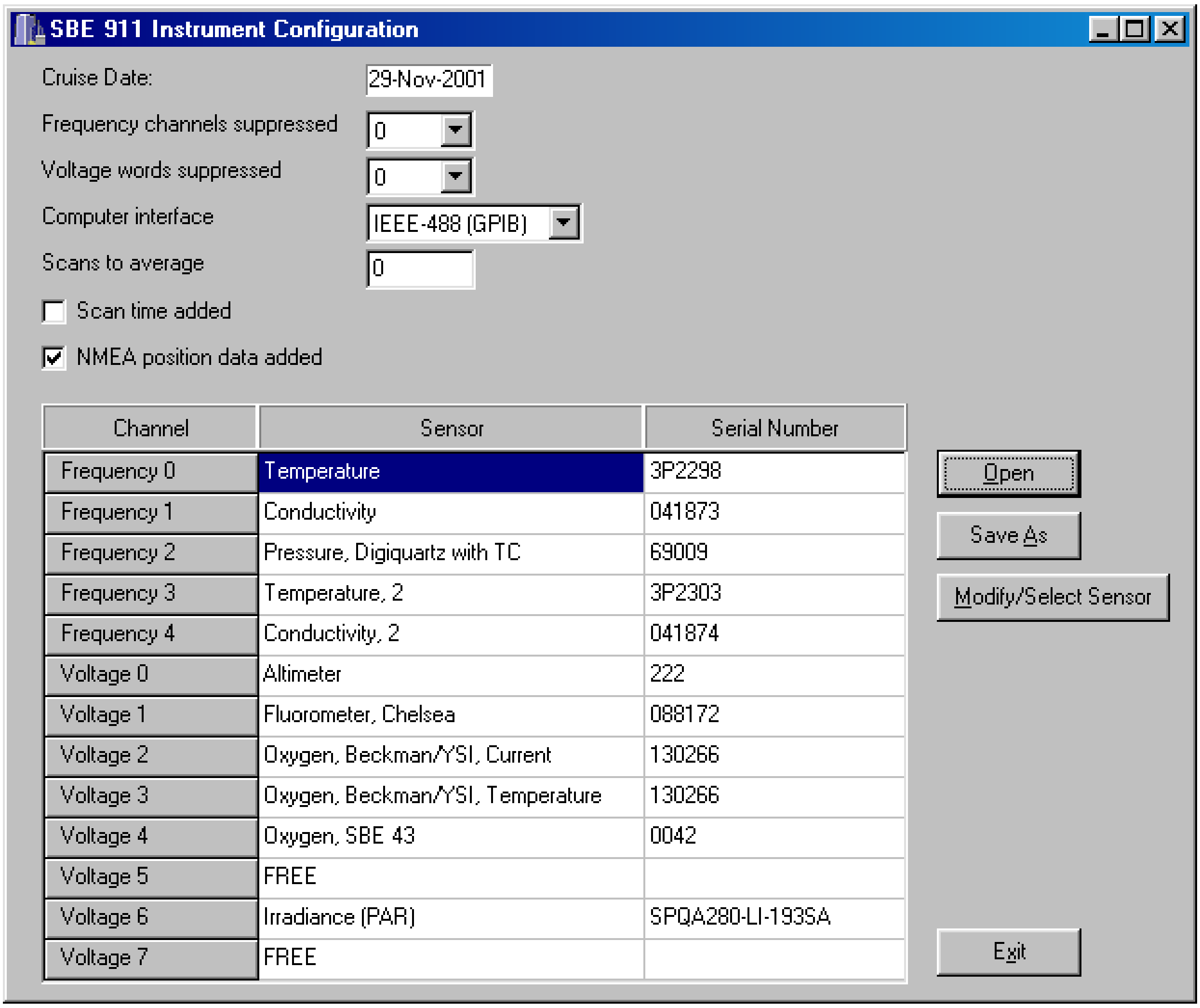


Note both the number of coefficients and the value of each coefficient.

In the past, these coefficients would be entered into the CON file appropriate to the assembled sensors. In larger laboratories, it is common to have many individual sensors. The exact arrangement of sensors for any given cruise is dependent on available sensors, sensor maintenance schedules, etc. This necessitates the entry and checking of these coefficients for each individual cruise set-up.

To reduce the level of effort and potential for errors in cruise set-up of the CON file, a data model and application have been developed to automatically build Seabird CON files.

The application requires the user to select the sensors based on serial number and sensor type, and constructs the CON file using a coefficient database. The selection component of the software looks like the following:



This application and data model will be briefly described. For any group using Seabird instrumentation and multiple arrangements of sensors, this application will be a valuable tool for the easy creation of CON files using stored coefficients. The software is freely available to the community.

A4.4 Written Submissions

T. O'Brien (USA)

World Ocean Database Plankton

- World Ocean Database 2001 (now available)
  - added 60,000+ stations since WOD-1998
  - added Biological Grouping Code (BGC)
  - added Common-unit Biological Value (CBV)
- World Ocean Atlas 2001
  - Plankton Atlas is in progress (due May 2002)
  - gridded mean fields of total zooplankton biomass and counts (e.g., bacteria, phytoplankton, zooplankton)
- NOAA Technical Journal
  - Plankton Data Management, due late summer (2002)
- WEB PAGE (www.nodc.noaa.gov)

S. Feistel (Germany)

Baltic Atlas of Long Time Inventory and Climatology
Summary for ICES WGMDM Meeting, Helsinki, April 2002
Sabine Feistel, Baltic Sea Research Institute Warnemünde (IOW), Germany

The Baltic Sea Research Institute Warnemünde (Institut für Ostseeforschung Warnemünde, IOW) is the successor of the former Institute for Marine Research (IfM Warnemünde), which was the most important East German oceanographic institute for several decades. Based on a long experience in Baltic Sea monitoring and corresponding long-time data records, IOW has undertaken the attempt to compile an improved climatological database for the Baltic in cooperation with other interested institutions.

BALTIC, a "Baltic Atlas of Long Time Inventory and Climatology"

The BALTIC project currently has got the status of a preparatory project of the IOW since January 2001. We will keep all interested parties informed about the progress of our activities on our website http://www.io-warnemuende.de/BALTIC. Please send suggestions and comments to sabine.feistel@io-warnemuende.de.

With the target to support e.g., climate-related investigations, interdisciplinary studies, numerical modelling and permanent monitoring, we intend to develop a comprehensive "climate atlas" for the Baltic Sea, similar to the famous Levitus global oceanographic data set, and going significantly beyond the already existing data collections of Bock/Lenz (1971) and Janssen/Backhaus/Schrum (1999).

In the recent years at IOW a lot of historical CTD and bottle data have been reconstructed in the HDR (historical data rescue) framework of the marine research institutes around the Baltic.

State of data holdings in the IOW data base as by March 2002:

- Monitoring (BMP) data from 1958 - 2001 (CTD, hydrochemical bottle data)
- current meter data Darss Sill (Baltic Sea) 1973 – 1992
- various (Baltic Sea) from 1970 – 2001
- various (Atlantic) 1976 - 1992
- data from 385 cruises, thereof 264 cruises in the Baltic Sea
- about 41 200 hydrographic stations (CTD profiles, hydrochemical bottle data), thereof 34 226 in the Baltic with about 7 800 000 measured samples (1 669 000 temperature, 1 290 000 oxygen, 1 635 000 salinity)

still in preparation:

- data from about 65 cruises

We began with the processing of temperature/salinity/oxygen measurements from our database in the surface layer (0–10 m) to gain experience in handling the data. A lot of new unexpected errors have been detected in our dataset despite of the systematic data validation and error checking during the routine data entry procedures, and we are still busy in correcting them. The majority of such cases appeared in early records prior to CTD use. We are currently specifying certain – as narrow as possible - validity limits separately for each parameter in certain regions and for various seasons. The first version of the BALTIC-software prepares a station map, generates an error file for each parameter (e.g., collects all values beyond the given validity limits), calculates Monthly Time Series (count, average and RMS), Long - Time and Monthly Climatology (count, average, RMS, min, max and trend) and draws monthly maps (like Figure 1) for the actually prescribed mesh size 1° x 1° x 10m x 1 month.

Output formats of processed data are presently lon x lat matrices or Surfer tables, see Figure 2.

As it can bee seen there are still a lot of white areas in the map. In a next step we will include the data of BSH and ICES. But we hope that the presentation of first results will lead to some cooperation with other institutes and additional data. We like to offer to interested participants the possibility not to provide us with original data but merely with statistical moments, in order to avoid potential user right problems and nevertheless improving the quantitative basis of BALTIC.

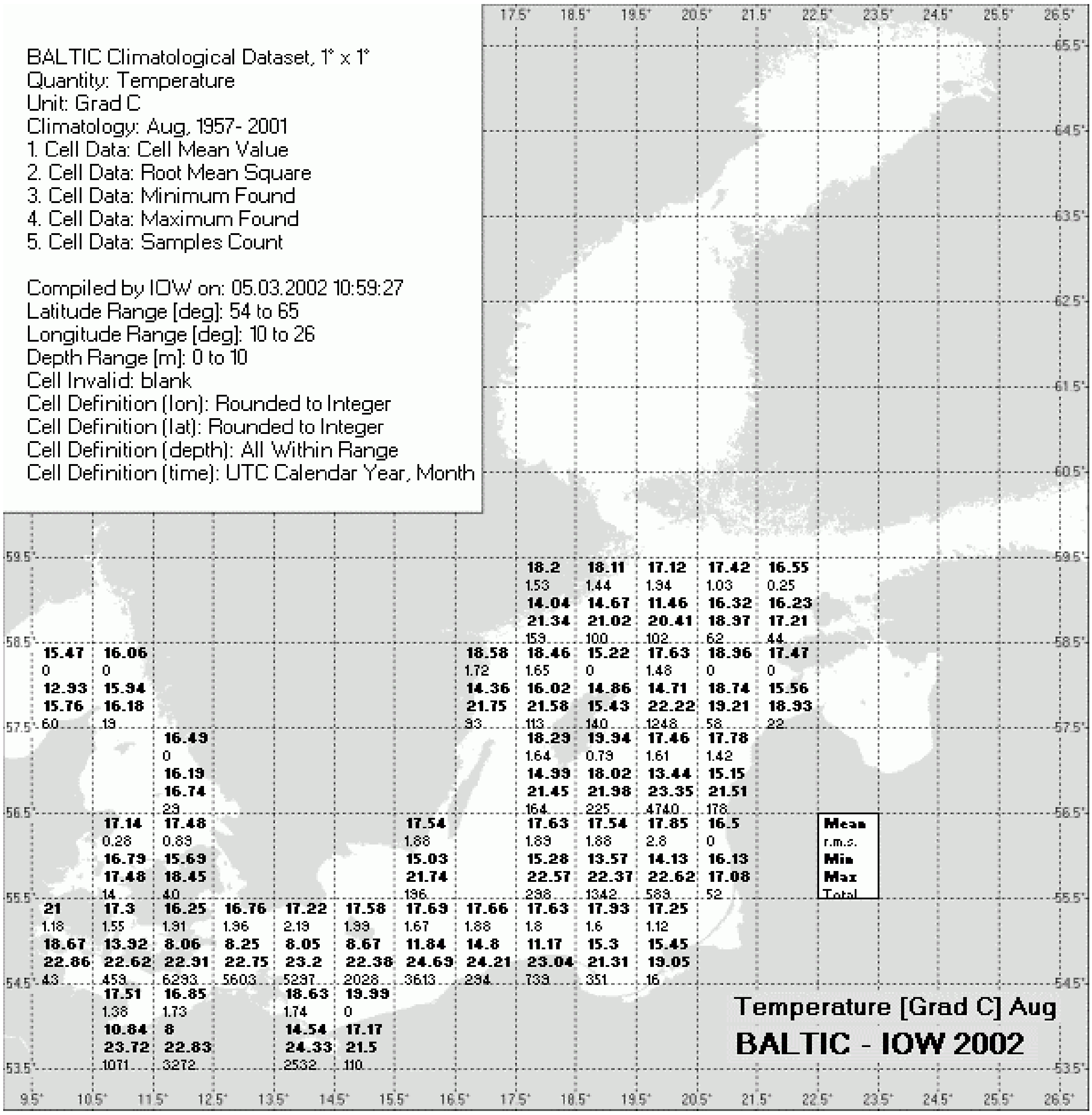


Figure 1

As a preliminary example, here the current spatial distribution of climatological surface layer temperatures is given for August. Within each cell, mean value, r.m.s., maximum and minimum ever measured, and the number of samples is shown.

You may well spot doubtful figures, still.

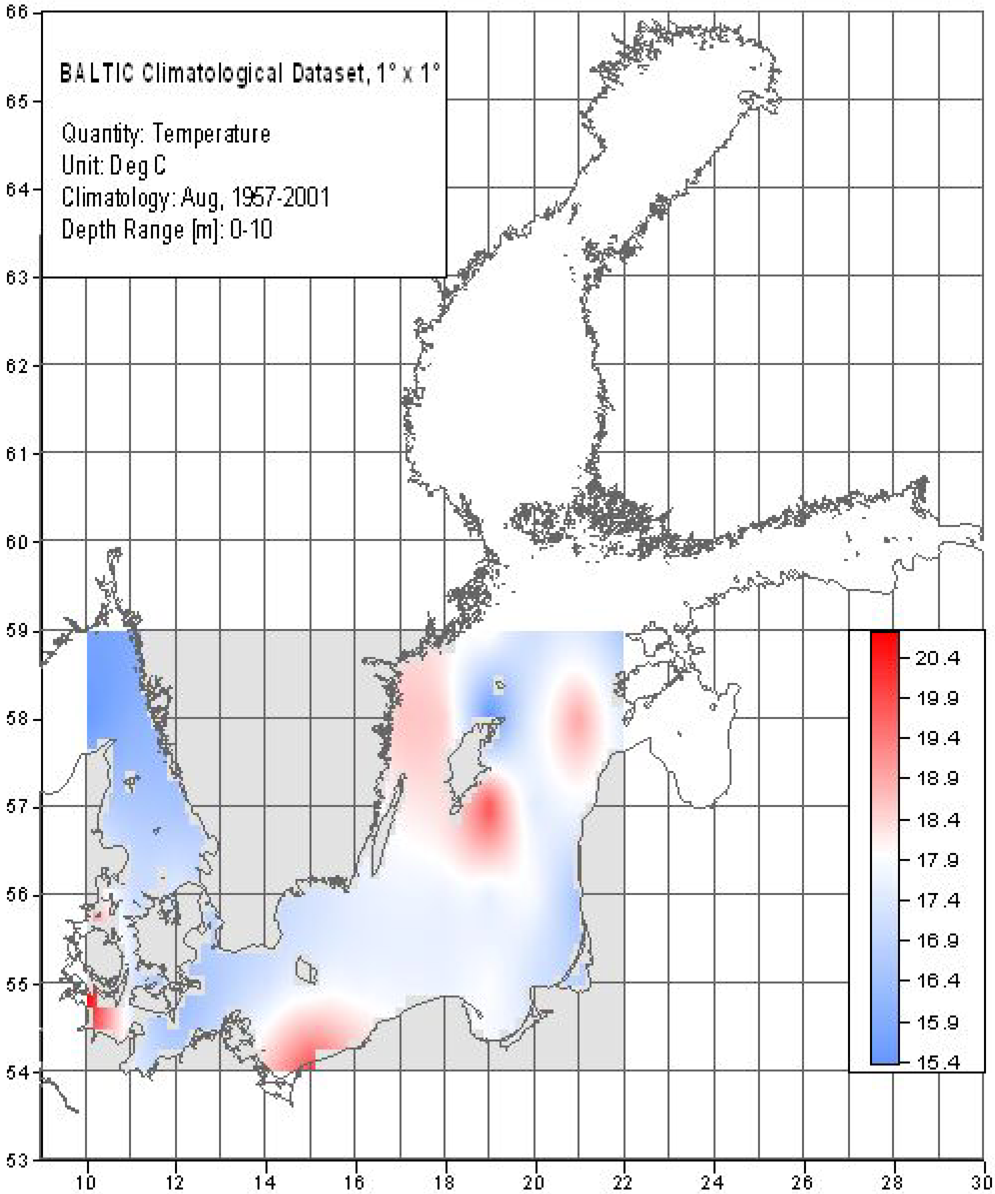


Figure 2

Distribution of climatological surface layer temperatures for August according to Figure 1.

ANNEX 6: DATA TYPE GUIDELINES

A5.1 Surface Drifting Buoy Data

ICES WGMDM
Guidelines for Surface Drifting Buoy Data
(Compiled 21 March 2001, revised August 2001)

Drifting buoys (UNESCO, 1988) have a long history of use in oceanography, starting in late 1978 with the First GARP Global Experiment (FGGE), principally for the measurement of currents by following the motions of floats attached to some form of sea anchor or drogue. Since 1988, over 1500 Lagrangian drifters have been deployed in the world oceans in the Surface Velocity Programme (SVP) of the World Ocean Circulation Experiment (WOCE) and the Tropical Ocean and Global Atmosphere Programme (TOGA). The buoys were standardised in 1991, with small spherical hull and floats, and large Holey-Sock drogue centred at 15 meters below the surface. Since 1993, Lagrangian Drifters with barometer ports and other sensors, including thermistor chains, have been in operation. Separate guidelines are available for Profiling Floats that record sub-surface variables.

1.0 RECEIVING DATA

The Data Centres require the following information to be supplied by the data supplier together with the data.

1.1 Data Standard

Delayed-mode quality controlled drifting buoy data provided by Principal investigators (PI) to the Data Centres should contain:

- A full description of the data format used for the data submission.
- Metadata information about the calibration (equations and coefficients) applied to the data set.
- Quality controlled data, reported at the original sampling interval of the instrumentation.
- A description of the quality control procedures applied to the data set.

All observed variables should be clearly specified and described. If parameter codes are used, then the source data dictionary must be specified in the metadata documentation. Variable units and precision must be clearly stated. If computed values are included, the equations used in the computations should be stated.

All relevant calibrations should be applied to the observed data including laboratory and field calibrations. Instrument calibration data should be included in the data file. The data should be fully checked for quality and flagged for erroneous values such as spikes, gaps, etc. An explicit statement should be made of the checks and edits applied to the data.

1.2 Format Description

Data Centres may receive drifting buoy messages in real-time coded formats transmitted on the Global Telecommunication System (GTS) and in delayed-mode quality controlled drifting buoy data formats from PIs. The WMO coded formats (WMO, 1995) used for GTS distribution of real-time drifting buoy messages are WMO FM 18-XI BUOY (ASCII), BUFR (binary), and CREX (coded BUFR).

The contents of the data and ancillary information should adhere to the Formatting Guidelines for Oceanographic Data Exchange (http://www.ices.dk/ocean/formats/getade_guide.htm) prepared by the IOC's Group of Experts on the Technical Aspects of Data Exchange (GETADE) and available from RNODC Formats.

The PIs data formats for the exchange of delayed-mode surface drifting buoy data set should include:

- A full description of the format used (preferably a fully documented ASCII format).
- The files should be homogeneous (i.e., each piece of information must always be in the same place in the file).
- Individual fields and units should be clearly defined.

- Ideally all of the data from the instrument should be stored in a single file.
- All data values should be in SI units.
- Time reported in UTC is strongly recommended.

1.3 Collection Details

Regarding the inclusion of metadata in real-time GTS reports, the meeting of the CBS Implementation Co-ordination Team on Data Representation and Codes, Geneva, 10–14 April 2000, accepted to recommend inclusion of certain metadata in the BUOY code for implementation in 2001. Proposed new fields relevant to drifting buoy data include:

- Drifting buoy type
- Drogue type
- If drogue was detached, date when the drogue was detached
- Agency responsible for the station (e.g., buoy operator)
- Instrument type, model, etc.
- Sensor calibration date

Metadata requirements for delayed-mode quality controlled drifting buoy data provided by PIs to the Data Centres include:

- Deployment platform name
- Country, organisation, Principal Investigator
- Project name
- Float number
- WMO number
- Sensor resolutions
- Information about the instruments and sensors (type and manufacturer, serial and model numbers, board type and serial number, software version)
- Information about the data precision and final accuracy.

Any additional information of use to secondary users which may have affected the data or have a bearing on its subsequent use.

2.0 VALUE ADDED SERVICE

When processing and quality controlling data, the Data Centres of the ICES community shall strive to meet the following guidelines.

2.1 Quality Control

The primary responsibility for data quality control lies with the PIs from which the observations originate. In many cases, drifting buoy data originate from a national meteorological service, an oceanographic institute or a PI for a particular research project.

Real-time surface drifting buoy messages entered on the GTS are quality controlled by the originator of the observation. To detect errors, which may escape national quality control systems and errors introduced subsequently, national services also carry out appropriate quality control of observational messages they receive. Because of the global nature of meteorological messages entered onto the GTS, special arrangements for quality control have been developed and implemented by Principal Meteorological or Oceanographic Centres (PMOC).

Thanks to several automatic Quality Control (QC) checks, erroneous messages are detected and removed from GTS distribution. Sensor data are compared with constant limits.

Variable	Lower limit	Upper limit
Sea level air pressure (hPa)	800	1080
Station air pressure (hPa)	400	1080
Air pressure tendency (hPa/3H)	0	100
Water temperature (deg)	- 1.8	+ 45
Air temperature (deg)	- 80	+ 50
Wind speed (m/s)	0	100
Wind direction (deg)	0	360

There are three main components to the quality control of delayed-mode surface drifting buoy data. All three components are used at the Data Centres or the PI to quality control the datasets.

The first component examines the characteristics of each float track looking to identify errors in either position or time based on calculated buoy speed versus time.

The second component is subjective as each observed variable is viewed independently to identify values that appear to be outside prescribed climatic limits for the area in question. A time series of one month of observations for each individual buoy is treated at one time. Knowledge of the different types of real and erroneous oceanographic and meteorological features is critical. This knowledge, when combined with a local knowledge of water mass structure, statistics of data anomalies, and cross validation with climatological data, ensures a data set of the best possible quality.

The third component is to identify and eliminate duplicate data. Duplicate observations occur either by having received the data more than once, or because real-time messages arrive before the delayed-mode data on which the real-time message was based.

To deal with both the real-time and delayed mode data, it is recommended that the Data Centre manage surface drifting buoy data in a continuously managed database. This will provide to the client those messages reported in real-time when these represent the only version available, or the delayed-mode data of higher quality which replace the original real-time data set. An overview of the data management practices for delayed mode drifting buoy data in a continuously managed database is provided in Annex A (Wilson, 1998).

2.2 Problem Resolution

The quality control procedures followed by the Data Centres will typically identify problems with the data and/or metadata. The Data Centre will resolve these problems through consultation with the originating PI or data supplier. The Data Centre may also consult the Marine Environmental Data Service of Canada, as the IOC Responsible National Oceanographic Data Centre for Drifting Buoys (http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Prog_Int/RNODC/RNODC_e.html), and also Joint IOC/WMO Data Buoy Cooperation Panel (http://dbcp.nos.noaa.gov/dbcp/) for advice when needed.

2.3 History Documentation

All procedures applied to a dataset should be fully documented by the Data Centre. These include all quality control tests applied and should accompany that dataset. All problems and resulting resolutions should also be documented with the aim to help all parties involved; the Collectors, Data Centre, and Users. A history record will be produced detailing any data changes (including dates of the changes) that the Data Centre may make.

3.0 PROVIDING DATA AND INFORMATION PRODUCTS

When addressing a request for information and/or data from the User Community, the Data Centres of the ICES community shall strive to provide well-defined data and products. To meet this objective, the Data Centres will follow these guidelines.

3.1 Data Description

The Data Centre shall aim to provide well-defined data or products to its clients. If digital data are provided, the Data Centre will provide sufficient self-explanatory information and documentation to accompany the data so that they are adequately qualified and can be used with confidence by scientists/engineers other than those responsible for their original collection, processing and quality control.

- A data format description fully detailing the format in which the data will be supplied
- Variable and unit definitions, and scales of reference
- Definition of flagging scheme, if flags are used
- Relevant information included in the metadata or data file (e.g., buoy type, drogue type, etc.)
- Data history document (as described in 3.2 below)

3.2 Data History

A data history document will be supplied with the data to include the following:

- A description of data collection and processing procedures as supplied by the data collector (as specified in Section 1.1 and 1.3)
- Quality control procedures used to check the data (as in Section 2.1)
- Any problems encountered with the data and their resolution
- Any changes made to the data and the date of the change

Any additional information of use to secondary users which may have affected the data or have a bearing on its subsequent use should also be included.

3.3 Referral Service

ICES member research and operational data centres produce a variety of data analysis products and referral services. By dividing ocean areas into regions of responsibility, and by developing mutually agreed guidelines on the format, data quality and content of the products, better coverage is obtained. By having the scientific experts work in ocean areas with which they are familiar, the necessary local knowledge finds its way into the products. Data and information products are disseminated as widely as possible and via a number of media including mail, electronic mail and bulletin boards.

If the Data Centre is unable to fulfil the client's needs, it will endeavour to provide the client with the name of an organisation and/or person who may be able to assist. In particular, assistance from the network of Data Centres within the ICES Community will be sought.

References

UNESCO. 1988. Guide to Drifting Data Buoys, IOC/WMO Manuals and Guides # 20.

Wilson, J.R. 1998. Global Temperature-Salinity Profile Programme (GTSPP) - overview and future, Intergovernmental Oceanographic Commission technical series 49, SC-98/WS/59.

WMO. 1995. WMO Manual on Codes No. 306.

UNESCO. 1991. Manual on International Oceanographic Data Exchange, IOC/ICSU Manual and Guides # 9, Revised Edition.

Annex A

Both the real-time messages and delayed mode surface drifting buoy data are available at the sample interval set by the manufacturer. This interval is typically hourly or synoptic hours depending on the requirements. The delayed mode data undergo calibration and quality control often incorporating site specific knowledge and experience of the PI. Real-time messages are often those surface drifting buoy data that undergo automatic, bulk quality control tests within operational time frames. The extensive quality control incorporating site-specific knowledge and experience of the PI often take longer. Real-time messages are most useful to those involved in operational forecasts, while delayed mode data are more useful to research.

To manage surface drifting buoy data, a Continuously Managed Database (CMD) system is implemented. As data are acquired in both real-time and delayed mode they are added to the database. Calibrated and quality controlled delayed mode data replaces the messages obtained in near real-time. The CMD therefore holds the most current and highest quality data set at all times. The database is continuously refined as additional quality checks are undertaken. Observations that have passed quality control and entered the database are not removed but are flagged to indicate that a higher quality version of the observation exists in the database.

A5.2 Biological Plankton Data

ICES WGMDM Guidelines
for Biological Plankton Data

(Compiled August 2001)

In the context of this guideline, phytoplankton or zooplankton sampling may be accomplished using either a vertical, horizontal or oblique tow of a net or from a rosette bottle.

In the case of a net, such a device would consist of frame which houses a mesh used in collecting the sample. An example maybe a square frame with multiple nets, of a single, conical shaped mesh with a circular ring opening. Typical mesh sizes would be less than 1000 μm (microns). At the mouth end the opening may be of up to 2 m. Attached at the small end of the net would be a jar or cod-end with a typical opening of about 10 cm.

1.0 RECEIVING DATA

The Data Centres require the following information to be supplied by the data supplier together with the data. When receiving data, the Data Centres of the ICES community shall strive to meet the following guidelines.

1.1 Data standard

The data set should consist of header and data information in one or more standard ASCII files. Each record should consist of date and time, navigation data, and measured parameters. It is recommended that each cruise constitute a single file. The navigation data should be in ASCII and in the form of latitude and longitude in degrees and decimal minutes, or decimal degrees. (Explicitly state which format is being used. It is recommended that N, S, E and W labels are used instead of plus and minus signs). Date and time must include month, day, year, hour, and minute. It is recommended that UTC be used.

All parameters must be clearly specified and described. If parameter codes are to be used, then the source data dictionary must be specified. Parameter units must be clearly stated. Parameter scales must be noted where applicable. If computed values are included, the equations used in the computations should be stated.

All relevant calibrations should be applied to the data including laboratory and field calibrations. The data should be fully checked for quality and flagged for erroneous values such as spikes, gaps, etc. An explicit statement should be made of the checks and edits applied to the data.

Sufficient self-explanatory information and documentation should accompany the data so that they are adequately qualified and can be used with confidence by scientists/engineers other than those responsible for its original collection, processing and quality control.

A brief description of the sample and data processing procedures must be included and should contain information regarding:

- Laboratory procedures and instrumentation
- Any species counts or mass measurements
- Description of any respiration, feeding or physiological experiments and results (e.g., carbon dioxide rates, carbon and nitrogen measurements)
- Report on corrections, editing or quality control procedures applied to the data
- Time reported in UTC is strongly recommended
- Estimate of final uncertainty in the data

Information about any supplementary/complementary data collected at the same time should also be supplied.

If a cruise/data report is available describing the data collection and processing, this can be referenced. If possible a copy should be supplied with the data.

1.2 Format Description

Data should be supplied in a fully documented ASCII format. Individual fields, units, etc. should be clearly defined and time zone stated. Time reported in UTC is strongly recommended. The contents of the data and ancillary information should adhere to the Formatting Guidelines for Oceanographic Data Exchange (http://www.ices.dk/ocean/formats/getade_guide.htm) prepared by the IOC's Group of Experts on the Technical Aspects of Data Exchange (GETADE) and available from RNODC Formats.

1.3 Collection Details

Other pertinent information to be included in the data transfer to the Data Centre includes:

- Project, ship, cruise identifier
- Country, organisation
- Date, time, latitude and longitude (for start and end if sampling via a net tow)
- Sounding, maximum and minimum pressure or depth of the tow
- Description of operational procedures such as tow orientation (vertical, horizontal or oblique), methods of position fixing (e.g., DGPS, GPS, etc.)
- Weather conditions (including sun and wind)
- Gear type (e.g., net mesh size, net mouth size, single or multi-net, etc.)
- Sample preservation method (e.g., pickling, frozen, etc.)
- Sample analysis/processing or data collection procedures (e.g., filtered size ranges, sub-sampling, etc.)

Any additional information of use to secondary users which may have affected the data or have a bearing on its subsequent use. An example field log sheet is included in Annex A.

2.0 VALUE ADDED SERVICE

When processing and quality controlling data, the Data Centres of the ICES community shall strive to meet the following guidelines.

2.1 Quality Control

A range of checks are carried out on the data to ensure that they have been imported into the Data Centre’s format without any loss of information. These checks should include:

- General check of accompanying information (e.g., tow dates within cruise dates, correct cruise identifier)
- Plot navigation to ensure no land points; compare with cruise report/CSR track chart if available
- Flag suspicious data or correct after consultation with the data supplier
- Checks on ship speed

If the navigation data are supplied separately, they will be merged with the individual tows.

2.2 Problem Resolution

The quality control procedures followed by the Data Centres will typically identify problems with the data and/or metadata. The Data Centre will resolve these problems through consultation with the originating Principal Investigator (PI) or data supplier. Other experts in the field or other Data Centres may also be consulted.

2.3 History Documentation

All quality control procedures applied to a dataset are fully documented by the Data Centre. Details of the quality control applied to a dataset should accompany that dataset. All problems and resulting resolutions will also be documented with the aim to help all parties involved; the Collectors, Data Centre, and Users. A history record will be produced detailing any data changes (including dates of the changes) that the Data Centre may make.

3.0 PROVIDING DATA AND INFORMATION PRODUCTS

When addressing a request for information and/or data from the User Community, the Data Centres of the ICES community shall strive to provide well-defined data and products. To meet this objective, the Data Centres will follow these guidelines.

3.1 Data Description

The Data Centre shall aim to provide well-defined data or products to its clients. If digital data are provided, the Data Centre will provide sufficient self-explanatory information and documentation to accompany the data so that they are adequately qualified and can be used with confidence by scientists/engineers other than those responsible for their original collection, processing and quality control. This is described in more detail below:

- A data format description fully detailing the format in which the data will be supplied
- Parameter and unit definitions, and scales of reference
- Definition of flagging scheme, if flags are used
- Relevant information included in the data file (e.g., ship, cruise, project, net tow deployment identifiers, start and end dates and times of tows, etc.)
- Data history document (as described in 3.2 below)

3.2 Data History

A data history document will be supplied with the data to include the following:

- A description of data collection and processing procedures as supplied by the data collector (as specified in Section 1.1 and 1.3)
- Quality control procedures used to check the data (as specified in Section 2.1)
- Any problems encountered with the data and their resolution
- Any changes made to the data and dates of these changes

Any additional information of use to secondary users which may have affected the data or have a bearing on its subsequent use should also be included.

3.3 Referral Service

ICES member research and operational data centres produce a variety of data analysis products and referral services. By dividing ocean areas into regions of responsibility, and by developing mutually agreed guidelines on the format, data quality and content of the products, better coverage is obtained. By having the scientific experts work in ocean areas with which they are familiar, the necessary local knowledge finds its way into the products. Data and information products are disseminated as widely as possible and via a number of media including mail, electronic mail and bulletin boards.

If the Data Centre is unable to fulfil the client's needs, it will endeavour to provide the client with the name of an organisation and/or person who may be able to assist. In particular, assistance from the network of Data Centres within the ICES Community will be sought.

Annex A
Example Net Tow Log Sheet

General

Project:_______ Ship:_________ Country:___________
Cruise Number:_________ Tow Number:___________ Event No.:_________
Location:____
Bottom Sounding:____________ Weather:________________________
Wind:______________________________

Start Tow

Date:______________ Time (UTC):____________ Twilight:______
Latitude:______ Longitude:______________ Method:______

End Tow

Date:______________ Time (UTC):____________ Twilight:______
Latitude:______ Longitude:______________ Method:______

Net Mouth Size:______________ Wire Angle:______________________

Net	Depth Range (Wire out)	Mesh Size	Volume of Water Filtered

Comments:

A5.3 Discrete Water Sample Data

ICES WGMDM Guidelines
For Discrete Water Sample Data
(Compiled December 1999, revised August 2001)

In the context of this guideline, discrete water sample data are considered to be any data that result from a single collection of water and so covers a huge variety of parameters. This collection of water must have a specific, identifiable time, position and depth. Such data could originate from a single bottle attached to a rosette or water drawn from a non-toxic supply.

No integrated samples are considered as part of discrete water sample data. Thus, tows that result in integrated data values are not considered in discrete water sample data. Nor are integrated samples from a pumping system or sediment trap.

1.0 RECEIVING DATA

The Data Centres require the following information to be supplied by the data supplier together with the data. When receiving data, the Data Centres of the ICES community shall strive to meet the following guidelines.

1.1 Data Standard

All parameters must be clearly specified and described. If parameter codes are to be used, then the source data dictionary consistency must be specified. Parameter units must be clearly stated. Parameter scales must be noted where applicable. If computed values are included, the equations used in the computations should be stated.

The data should be fully checked for quality and pre-edited or flagged for erroneous values. An explicit statement should be made of the checks and edits applied to the data.

A brief description, or a reference to the data collection and processing methods (e.g., reference to a specific technique or specific project protocols) must be included and should contain information regarding:

- Describe or reference full laboratory methods and procedures
- If sample was sent out for analysis, give laboratory name and accreditation level
- Describe or reference any internal or external quality assurance procedures (e.g., QUASIMEME, IAPSO)

A brief description of the data processing procedures must be included and should contain information regarding:

- editing/quality control methods
- how are trace values (values below the detection limit) identified
- how are missing values handled (null vs. zero, or "blanks")
- what is the precision of the methods (e.g., number of significant figures)

If a cruise/data report is available describing the data collection and processing, this can be referenced. If possible a copy should be supplied with the data.

1.2 Format Description

Data should be supplied in a fully documented ASCII format. Data Centres are capable of handling water sample data in a wide variety of user-defined and project formats. If in doubt about the suitability of any particular format, advice from the Data Centre should be sought.

Individual fields, units, etc. should be clearly defined and time zone stated. Time reported in UTC is strongly recommended. Ideally all of the data from the single water source should be stored in a single file. The contents of the data and ancillary information should adhere to the Formatting Guidelines for Oceanographic Data Exchange (http://www.ices.dk/ocean/formats/getade_guide.htm) prepared by the IOC's Group of Experts on the Technical Aspects of Data Exchange (GETADE) and available from RNODC Formats.

Often different groups or laboratories will analyse a single water sample for a multitude of parameters. In such cases, it is common for the data from the different groups to arrive at the data centre at different times. The receiving data centre may merge those data from a single water source. Thus it is crucial that the date/time, position and sample identifier accompany the data.

1.3 Collection Details

Pertinent information to be included in the data transfer to the Data Centre includes:

- Project, platform, cruise identifier
- Country, organisation, institute, PI
- Station number, sample identifier (or bottle number)
- Date and time of the start of the sampling (UTC is recommended)
- Position (latitude and longitude degrees and minutes or decimal degrees can be used. Explicitly state which format is being used. It is recommended that N, S, E and W labels are used instead of plus and minus signs.)
- Description of operational procedures including (where applicable) sampling rate, detection limits, standard analytic procedures, calibration of equipment, quality control of original data, methods of position fixing (e.g., GPS, DGPS)
- Details of the collection instrument and sensor (e.g., manufacturer, model number, serial number, and sampling rate)
- Sounding should be included for each station. The method and assumptions of determining the sounding should be included.
- Range of data values (desirable)

Any additional information of use to secondary users which may have affected the data or have a bearing on its subsequent use.

For additional information on quality control procedures, metadata requirements for particular parameters and collection instrumentation, see UNESCO (1996).

2.0 VALUE ADDED SERVICE

When processing and quality controlling data, the Data Centres of the ICES community shall strive to meet the following guidelines.

2.1 Quality Control

A range of checks are carried out on the data to ensure that they have been imported into the Data Centre’s format correctly and without any loss of information. For discrete water sample data, these should include:

- Check header details (station numbers, date/time, latitude/longitude, instrument type, data type/no. of data points, platform identifier)
- Plot station positions to check not on land
- Check ship speed between stations to look for incorrect position or date/time
- Automatic range checking of each parameter (e.g., WOD 1998, Maillard 2000)
- Check units of parameters supplied
- Check pressure increasing or decreasing as appropriate
- Check no data points below bottom depth
- Check depths against echo sounder
- Plot profiles (individually, in groups, etc)

- Check for spikes
- Check for vertical stability/inversions
- Check profiles vs. regional climatology
- Check calibration information available
- Compare parameters for predictable relationships (e.g., parameter ratios)
- Check for consecutive constant values
- Duplicate detection when comparing to archived data
- Flag suspicious data or correct after consultation with Principal Investigator (PI)

2.2 Problem Resolution

The quality control procedures followed by the Data Centres will typically identify problems with the data and/or metadata. The Data Centre will resolve these problems through consultation with the originating PI or data supplier. Other experts in the field or other Data Centres may also be consulted.

2.3 History Documentation

All quality control procedures applied to a dataset are fully documented by the Data Centre. Furthermore, All quality control applied to a dataset should accompany that dataset. All problems and resulting resolutions will also be documented with the aim to help all parties involved; the Collectors, Data Centre, and Users. A history record will be produced detailing any data changes (including dates of the changes) that the Data Centre may make.

3.0 REQUEST FOR SUPPORT

When addressing a request for information and/or data from the User Community, the Data Centres of the ICES community shall strive to provide well-defined data and products. To meet this objective, the Data Centres will follow these guidelines.

3.1 Data Description

The Data Centre shall aim to provide to its clients well-defined data or products. If digital data are provided, the Data Centre will provide sufficient self-explanatory series header information and documentation to accompany the data so that they are adequately qualified and can be used with confidence by scientists/engineers other than those responsible for their original collection, processing and quality control. This is described in more detail below:

- A data format description fully detailing the format in which the data will be supplied
- Parameter and unit definitions, and scales of reference
- Definition of additional quality control
- Flagging scheme, if flags are used
- Data history document (as described in 3.2 below)
- Accompanying data (e.g., CTD data at the time of bottle trip)

3.2 Data History

A data history document will be supplied with the data to include the following:

- A description of data collection and processing procedures as supplied by the data collector (as specified in Section 1.1 and 1.3)
- Quality control procedures used to check the data (as specified in Section 2.1)
- Any problems encountered with the data and their resolution and modification date
- Any changes made to the data and dates of these changes

Any additional information of use to secondary users which may have affected the data or have a bearing on its subsequent use should also be included.

3.3 Referral Service

ICES member research and operational data centres produce a variety of data analysis products and referral services. By dividing ocean areas into regions of responsibility, and by developing mutually agreed guidelines on the format, data quality and content of the products, better coverage is obtained. By having the scientific experts work in ocean areas with which they are familiar, the necessary local knowledge finds its way into the products. Data and information products are disseminated as widely as possible and via a number of media including mail, electronic mail and bulletin boards.

If the Data Centre is unable to fulfil the client's needs, it will endeavour to provide the client with the name of an organisation and/or person who may be able to assist. In particular, assistance from the network of Data Centres within the ICES Community will be sought.

References

Maillard, C., and Fichaut, M. 2000. Medar-Medatlas Protocol, Part I: Exchange Format And Quality Checks For Observed Profiles, IFREMER, June 2000 - R.INT.TMSI/IDM/SISMER/SIS00–084.

UNESCO. 1996. IOC-EU-BSH-NOAA-(WDC-A). International Workshop on Oceanographic Biological and Chemical Data Management Hamburg, Germany 20–23 May 1996, IOC Workshop Report 122.

WOD, 1998. World Ocean Database, Documentation and Quality Control, Version 2, Silver Spring, MD, December 1999.

ANNEX 7: REPORT BY ICES OCEANOGRAPHER

Working Group on Marine Data Management
Helsinki, April 2002–03–22
Report on Oceanographic data activities in the ICES Secretariat

ROSCOP

Table 1 lists the number of ROSCOP forms received by the Secretariat and entered to the ICES ROSCOP database for the period since 1990.

The vast majority of ROSCOP forms are delivered in hard copy with very few being received digitally. France and UK provide most information using.our/.inf files. France exports into this format directly from its Oracle database.

On receipt, forms are quality controlled and these reveal many errors and inconsistencies, which have to be resolved before entering in the database. The Table shows that the overall health of the system from the Secretariat's perspective is not good. The continued decline in interest in the system arises probably for a number of reasons. One of these is that the Secretariat no longer actively seeks submissions.

The number of ROSCOPs in the database is some 20% more than indicated in the Table. This is because the Table does not include those ROSCOPs, which were created by the Secretariat following the receipt of data for which ROSCOP forms had not been submitted.

Table 1. Table showing ROSCOP form submission by country, 1990–2002 (as of 23 March 2002).

Country/ Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	Total
Australia	0	0	1	26	48	42	50	57	60	0	0	0	0	284
Belgium	26	24	25	30	31	26	29	33	29	28	32	0	0	313
Canada	16	17	28	44	31	4	27	2	2	64	0	0	0	235
Denmark	19	24	20	20	26	22	22	42	26	32	27	6	1	287
Estonia	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Finland	0	0	0	0	0	0	0	1	3	0	0	0	0	4
France	127	108	82	86	134	99	155	132	121	137	133	0	0	1314
FRG	175	0	0	0	0	0	0	0	0	0	0	0	0	175
Germany	0	174	149	198	209	197	168	150	130	106	83	35	0	1599
GDR	15	0	0	0	0	0	0	0	0	0	0	0	0	15
Iceland	9	11	8	16	16	13	4	27	8	5	2	5	1	125
Ireland	9	2	10	11	3	2	0	0	0	0	0	0	0	37
Japan	0	2	113	70	46	0	0	1	0	0	0	0	0	232
Latvia	0	0	1	3	2	8	4	2	11	11	10	0	0	52
Lithuania	0	0	0	0	0	0	0	0	0	0	0	0	0	0
New Zealand	0	0	2	8	10	5	12	0	0	0	0	0	0	37
Norway	40	42	71	56	63	70	85	102	91	77	69	51	0	817
Netherlands	82	81	13	10	17	23	14	9	8	3	2	0	0	262
Poland	14	15	7	4	10	6	20	17	12	1	0	0	1	107
Portugal	1	0	0	2	5	4	2	2	2	0	0	0	0	18
Spain	4	25	34	28	27	27	22	19	2	0	0	0	0	188
Sweden	17	14	14	20	22	19	19	20	19	47	46	18	0	275
UK	132	136	120	134	81	82	72	46	39	49	24	6	0	921
USA	120	143	112	115	81	47	44	50	46	38	30	0	0	826
USSR	8	2	0	0	0	0	0	0	0	0	0	0	0	10
Russia	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Other	1	7	1	5	4	1	0	1	3	1	0	0	0	24
Total	**815**	**827**	**811**	**887**	**866**	**697**	**749**	**713**	**612**	**599**	**458**	**121**	**3**	**8158**

The Secretariat currently runs two ROSCOP databases. The system continues to be managed via the.inf file system, but these files are copied weekly to the JET database which is the basis of an online search facility built up in much the same way as the online systems operated at BODC and SISMER. This is not a satisfactory situation, which should not be allowed to persist. But whatever the ultimate solution it is hoped that the user community will continue to be served by both online and offline (ROSEARCH) systems.

Data Submissions

According to Table 2, data submissions to the database continue in decline. The extent of the decrease appears to be about 75%. There are a number of reasons for this. There are a number of data sets in the "pipeline" and there are a number that have been received that have not yet been dealt with at all. These and published CD-ROM data sets which have been identified as suitable for inclusion in the database total about 50.

The situation with regard to nutrient data is even more disastrous, and even data submitted as part of the OSPAR monitoring programmes is in decline. Norway (Bergen) has not provided nutrient data since 1995, and receipt of this backlog is viewed with some trepidation.

Oceanographic data in support of the ICES IBTS surveys is also being significantly delayed, partly because of problems at the Institutions. In addition the Secretariat has had no opportunity to produce the standard oceanographic products for 2001 to support this survey.

Table 2. Number of Cruises by country for which data has been entered to the ICES oceanographic database, 1990–2002 (as of 22 March 2002).

Country/Year	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**Total**
Belgium	8	7	7	9	10	10	6	30	28	6	6	0	0	127
Canada	28	25	30	33	20	10	12	15	31	82	30	0	0	316
Denmark	31	18	19	19	26	24	24	29	30	15	9	1	0	245
Estonia	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Finland	7	15	15	14	18	22	19	13	17	15	0	0	0	155
France	21	28	20	22	29	14	10	15	7	6	3	2	0	177
FRG	75	0	0	0	0	0	0	0	0	0	0	0	0	75
Germany	0	93	65	75	60	55	49	39	37	10	10	3	1	497
GDR	6	0	0	0	0	0	0	0	0	0	0	0	0	6
Iceland	9	10	8	16	15	0	3	5	5	5	0	5	0	81
Ireland	0	0	1	6	1	0	0	0	0	0	0	0	0	8
Latvia	0	0	0	0	2	8	6	7	11	11	10	0	0	55
Lithuania	0	0	0	0	0	0	0	0	0	4	0	0	0	4
Norway	76	69	77	68	80	64	64	74	77	68	60	18	0	795
Netherlands	6	15	24	35	34	31	26	17	4	2	1	1	0	196
Poland	17	19	11	15	14	17	11	18	8	11	9	0	0	150
Portugal	2	0	0	2	4	4	1	0	2	0	0	0	0	15
Spain	1	11	17	17	24	31	11	13	12	12	13	0	0	162
Sweden	42	15	19	22	24	19	20	19	21	46	3	1	0	251
UK	58	36	45	56	48	54	36	24	22	19	24	25	1	448
USA	6	3	0	1	1	0	0	3	0	0	0	0	0	14
USSR	40	12	0	0	0	0	0	0	0	0	0	0	0	52
Russia	0	0	3	3	1	0	0	1	1	1	1	0	0	11
Other	2	1	1	0	1	0	0	0	0	0	0	0	0	5
Total	**435**	**377**	**362**	**413**	**412**	**363**	**298**	**322**	**313**	**313**	**179**	**56**	**2**	**3845**

Quality Control

Enough work is being done in the Group to draw the conclusion that the quality of some data submissions remains very patchy.

Figure 1 is a striking example of the kind of quality problems that are currently in the data exchange system. The apparently reduced salinity in 2001 arises from one data set produced from Ship-of-Opportunity observations. These data are also available in the various GTS centres, including Washington and Ottawa, but apparently no problems by either of these organisations, nor the ICES Oceanographic data centre, were initially noticed. Yet the observed and reported salinity values are in error by at least 20. It is a bit worrying (to the author at least!) that such data can pass through the QC net, including that set up by the scientist who had organised the collection of these data.

Fortunately, in this case, there was enough background and logged information available to allow a firm decision to be made on what if anything was wrong with these data. From entries in the ship's log it was possible to deduce that these and other data were indeed incorrect and should be rejected. Apparently the plastic impellor to the seawater intake pump had been broken when someone accidentally switched on a fire pump. Bits of this plastic had become lodged in the TSG housing, thus upsetting the conductivity loop. It was not until after the sensor housing was cleaned out 2 days later that it was possible to once again obtain reliable data.

But these data are already in the system in various places. How do we get them out of it?

Even though very few nutrient data sets are being acquired these days quality problems continue to proliferate some of those data sets that are submitted. One example is Figure 2 which are plots of nutrient values in the Northwest Atlantic. The analysis was set up in order to verify recently acquired Canadian nutrient data sets for this area.

The results here are certainly not satisfactory. The historical record for phosphate had largely been based on the Russian cruise activities of the 1960 and 1970s but this plot indicates that the Russian data differs from later data collected by Germany and Canada by a factor of more than two. Much of the historical phosphate data for the world's oceans and the OWS programme derives from Russian sources. This analysis begs the question whether any of these data can be trusted. The plot further indicates systematic, but smaller, differences of about 10%, the German data being systematically lower that Canadian. Figure 3, from the Greenland Sea, shows a similar problem. In this case UK nitrate data is systematically different from data collected by Norway and Germany.

These add to the growing list of anecdotes concerning nutrient data quality. In spite of international standards, carefully coordinated international projects, and the latest advanced technologies to provide more reliable and precise instrumentation, the level of precision of chemical measurements continues to fall short of what is required for international databases.

Although the examples in Figure 2 draw attention to systematic differences, which are fairly straightforward to detect, there are many further examples, which show very different noise levels from one data set to another. A notable example of this was the JGOFS data sets, which could not easily be combined across countries in spite of the fact that JGOFS data collection procedures were subjected to rigorous laboratory and at-sea inter-calibration exercises.

Problems of this nature have been around now for decades, but in the advent of global operational programmes such as GOOS and its regional components, the resolution of this problem has now become urgent. Should the data management community not now take the lead by again drawing attention to the international shortcomings of such data, and in particular now seek to draw attention to the problem in a near operational mode? In this way the science community can be more rapidly alerted to the fact that it really has to do much better that it has been until now.

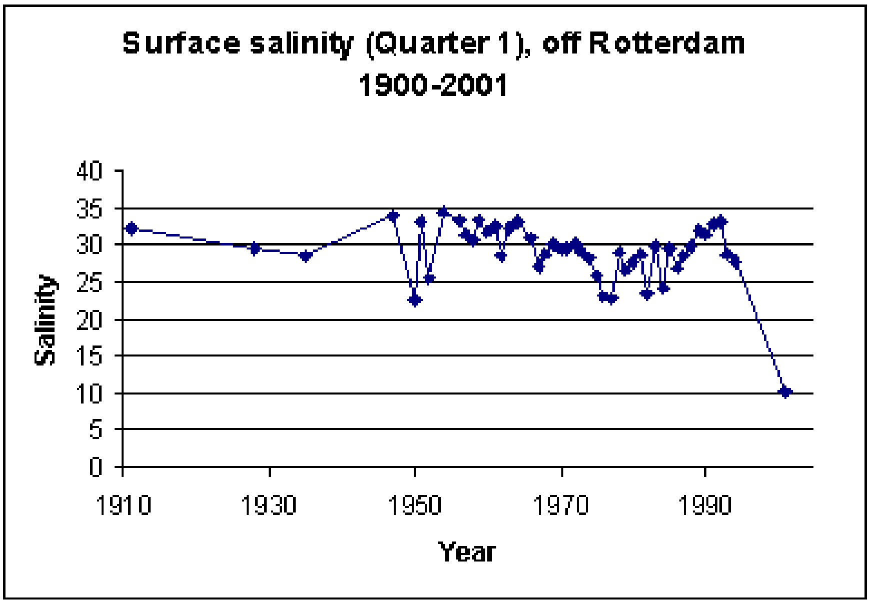


Figure 1. Example of Mean Surface salinity in Southern North Sea. The values in 2001 are very untypical but the existence of such data was not initially noticed.

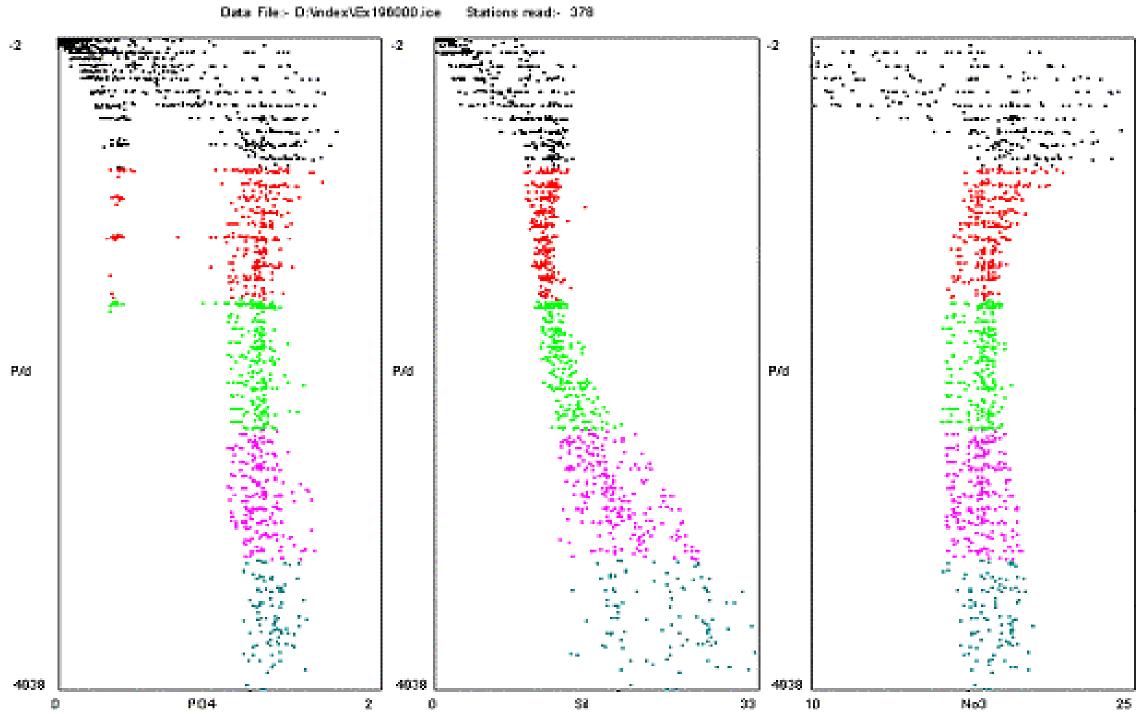


Figure 2. Plots of depth against phosphate, silicate, nitrate from historical data in the region of 45N 40W which indicates systematic differences in values between Russian (90), German (06) and Canadian(18) data. The published phosphate climatologies reflect the Russian values.

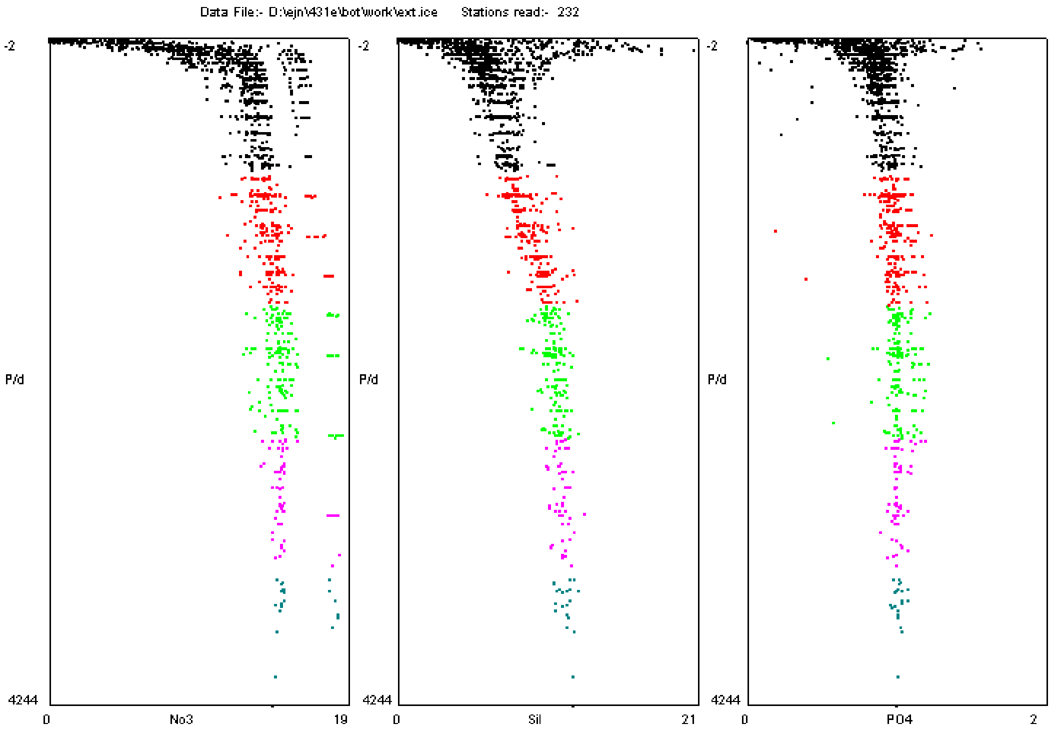


Figure 3. Plots of nitrate, silicate and phosphate in the Greenland Sea (1985–1995).

ANNEX 8: PROPOSED 2002/2003 TERMS OF REFERENCE

2CMD The **Working Group on Marine Data Management** [WGMDM] (Chair: A. Isenor, Canada) will meet in Gothenburg, Sweden from ?-? to:

a) Evaluate the use of the MDM guidelines for data management and exchange in response promotional activities.
b) Evaluate the results from SGXML regarding the cross parameter dictionary comparison and make recommendations regarding adoption in the oceanographic community.
c) In partnership with the IOC/IODE GETADE, further investigate details of the ITIS and actively promote the ITIS within the ICES and IOC community.
d) Identify problems in terms of both submission amount and quality of oceanographic data submitted to the ICES data centre and suggest solutions to member countries or international programmes as required.
e) Evaluate and develop future directions for oceanographic data management based on the results from SGXML.

WGMDM will report by 30 April 2003 for the attention of the Oceanography Committee

Supporting Information

Priority:	This Group flies the flag for ICES in setting standards for global databases. It also provides an important interface for oceanographic and environmental data management in ICES.
Scientific Justification:	a) Considerable inter-sessional effort will be made on promoting the data type guidelines. This effort will be assessed and feedback from other groups and organisations will be evaluated. Such feedback will help to establish future guideline activity. b) The efforts of SGXML have potential implications and application to general data exchange procedures. These efforts should be followed within the broader context of general oceanographic data flow. c) The cooperative environment between GETADE and WGMDM resulted in very constructive discussions and many collaborative action items. Of particular interest to both groups is ITIS and its role in the standardisation of data exchange. The ITIS should be actively promoted with the communities. d) The submission amount and quality of data sent to the ICES databank is an ongoing issue. The MDM needs to actively identify related issues and pass on recommendations to data centres or collectors as necessary. e) The data management community must explore the use of new technologies, such as XML, in a broader context. The WGMDM will attempt to integrate the efforts of SGXML into this broader context and develop possible directions for ocean data management in a distributed environment.
Relation to Strategic Plan:	ICES scientific objectives of understanding marine ecosystems must be underpinned by good data management procedures
Resource Requirements:	None
Participants:	Core Group of members of national oceanographic data centres ensure well attended meetings.
Secretariat Facilities:	None, apart from preparation of material by the Oceanographer.
Financial:	The Oceanographer should attend this meeting.
Linkages To Advisory Committees:	Group's report is seen by ACME.
Linkages To other Committees or Groups:	None, but links should be encouraged to broaden the scope of the group to more generic data management issues.
Linkages to other Organisations	IOC, especially its Working Committee on International Oceanographic and Information Exchange (IODE).