International Council for
the Exploration of the Sea

C.M. 1993/C:1
Hydrography Committee

Report of the

WORKING GROUP ON MARINE DATA MANAGEMENT

Aberdeen, 22 - 24 April 1993

This is a report of a Working Group of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council. It should therefore not be quoted without consultation with the Secretary.

*General Secretary
ICES
Palægade 2-4
DK 1261 Copenhagen K
Denmark

Table of Contents

1. Opening of the meeting

The meeting was opened at 0930 on 22 April 1993, hosted by the SOAFD Marine Laboratory, Aberdeen. Participants were welcomed by the Chairman and E. Henderson of the Fisheries Laboratory who also explained the local arrangements. H. Dooley welcomed the participants on behalf of ICES.

Members of the Working Group present were: J. Atkinson, UK, J. Blindheim, Norway, R. Gelfeld, USA, N. Håkansson, Sweden, B. Hansen, Faroes, E. Henderson, UK, H. Loeng, Norway, K. Medler, UK, S. Narayanan, Canada, P.B. Nielson, Denmark, R. Olsonen, Finland, L. Rickards, UK (Chairman), J. Szaron, Sweden, H. Valdimarsson, Iceland and J. Wallace, Ireland. The Council was represented by H. Dooley. Apologies for absence were received from L. Smit, the Netherlands and C. Wulff, Germany. The Chairman had been in contact with D. de Armes from Spain (which is not currently represented on the WG), inviting him to attend the meeting. Unfortunately he was unable to do so, but he had provided a report of activities.

2. Adoption of the Agenda

The agenda for the WG meeting was adopted as a resolution of the Statutory Meeting in Rostock (C.Res. 1992/2:35, Annex 1).

3. Reports of activities of Data Centres in the ICES area

WG participants reviewed activities at their own data centre/laboratory over the past year and looked to developments in the future. A summary of these activities can be found in Annex 2(a) and the reports were distributed to WG members. Several laboratories are establishing or re-establishing their databases using commercial packages like Oracle or Ingress and this is proving more complicated than initially expected. In addition, several laboratories are working on data archaeology projects. R. Gelfeld described work at the US NODC/WDC-A currently being undertaken in this field which includes much cooperation with ICES. J. Wallace reported on activities in Ireland, where a Marine Data Centre is currently being established. This Centre officially came into existence on 1st April 1993. The WG welcomed this development and looked forward to hearing about progress made in the future.

4. Assess the state of oceanographic data quality control and data processing, taking into account the present performance of instrumentation

In Canada, a data collection and processing manual exists, which technicians are encouraged to use. S. Narayanan reported that her laboratory has a PC based data acquisition and processing system for CTDs, which is used as the standard for Canada. It was designed for use with Guildline CTDs, but can also be used with SeaBird. It is a user friendly package, which may be marketed by a commercial firm in the future. She agreed to send further details to WG members.

Within BODC most of the data quality control is done using high speed graphics workstations and a software package (SERPLO = SERies PLOtting) which was written in-house. Any problems in the data are identified and discussed with the data supplier, spikes are flagged, series header details are compiled and checked and data documentation assembled. This documentation, in addition to describing methods of data collection and processing, also notes any unresolved problems with the data. BODC are presently re-writing their Laboratory's sea level data processing software. Other quality control packages held by BODC include ADCP and sea level data processing and quality control software from the University of Hawaii.

J. Wallace asked what commercial packages were in use for data display and for statistics. Those packages used by WG members included SASS, Uniras, Surfer, PV-Wave and S-Plus. Overall it was felt that there was no substitute for visually inspecting the data. Packages can be used to automatically detect problems in the data, but they may remove good data or miss bad points. Some WG members eliminate data that are thought dubious; others favour flagging them. Discussion followed on how to flag data, what sort of accuracy of data are sought, etc., from which it emerged that the key was the importance of flexibility. Climatologies were felt useful to help with checking, particularly the Levitus Atlas. In addition the Canadians have built their own climatology for the shelf area around eastern Canada. Within the Baltic, there is a manual for data collection and also Baltic station specific models are used to check data. The package in use at MEDS for GTSPP, for data quality control is also in use at other centres, for example BIO (Canada), US NODC and IFREMER (France).

The expected accuracy of data from different locations varies, for example data collected in the North Sea can be measured to a greater accuracy than that collected in the Baltic, or the west coast of Scotland sea lochs. It was recommended that scientists/data collectors be encouraged to use manuals where they are available

and there was some discussion of sampling procedures and whether they should be reviewed.

Discussion followed about whether all quality control procedures have been written up and whether guidelines and details of recommended software are available. Manuals of use for data processing and quality control include the JPOTS manual, the SCOR Working Group 51 Report (IOC), WOCE Hydrographic Programme guidelines, various MDM WG recommendations (current meter and CTD data) and the forthcoming IOC/CEC Quality Control Manual. It was noted that the CEC intend holding a Quality Control Workshop in Dublin, prior to this year's Statutory Meeting.

5. Further develop guidelines for the management of ADCP and SeaSoar data

H. Dooley reported that the draft guidelines produced by the WG had been noted at the IOC International Oceanographic Data and Information Exchange (IODE) meeting last December. The ADCP guidelines had been passed on to Japan Oceanographic Data Centre (JODC), who are the RNODC for shipborne ADCP data, and the SeaSoar guidelines had been passed on to the IODE Group of Experts on the Technical Aspects of Data Exchange (GE-TADE). L. Rickards reported that she had also passed on the ADCP draft guidelines to the Japanese, who had not responded with any comments. She was also waiting to receive their manual for routine processing of ADCP data. This would be followed up soon when a Japanese colleague visited BODC. She had also sent the guidelines to D.Hamilton at the US NODC and he had responded with some comments. Specifically, he had recommended that the following information should be included in the guidelines:

* ADCP data should be accompanied by information about the transducer installation (eg depth, location on hull)

* Slight variances in PC clocks can introduce substantial errors (in navigation data), so attention is needed to that.

* Some scientists may have made sound speed corrections to improve the data, since the instrument is calibrated to standard values of sound speed in the water column. These corrections can be made from climatology or actual conditions, but whatever is done, a flag is needed to indicate whether or not these corrections were made, so that the same correction is not made twice.

The US NODC and US scientists are presently trying to work together to set up a system for processing and storing ADCP data. Following on from this, L. Rickards reported that she had discussed ADCP data with Patrick Caldwell (US NODC and University of Hawaii), who was setting up an ADCP processing system and database. He has available a software package, both for IBM PCs and Unix workstations. She has a copy of the PC version but has not yet tested it out. There is a comprehensive manual to accompany the package. The USA are trying to share technology and anyone interested would be welcome to try out the package. The University of Hawaii was keen to allow others to use package and to get some feedback.

There had been a paper at last year's Statutory Meeting from Bo Lundgren at the Danish Institute of Fisheries relating to ADCP data and how to process and archive them. In particular he posed questions about the choice of averaging interval (for both time and depth), choice of parameters to store, choice of reference level, recommendations regarding the use of tilt sensors, cleaning periods for transducers, calculation of vessel speed over the ground, suggestions for a database format, establishment of a mathematical model, deduction of tidal currents. L. Rickards agreed to write to him to keep him up to date with developments and inform him of the University of Hawaii system.

The WG agreed that a good way to test out the MDM guidelines would be to try exchanging some data using them; this would help to identify where extra information was needed and to try to solve any problems encountered. This would be carried out during the next year for both ADCP and SeaSoar data. Those countries collecting ADCP data could exchange data with one another and those collecting SeaSoar data could try sending these data to ICES. L. Rickards agreed to coordinate this and report back to next year's meeting, with a view to updating and amending the guidelines where necessary.

6. Examine oceanographic data flows in ICES Member Countries and make recommendations to improve the situation

The ICES oceanographic data bank contains approximately 500000 profiles, but there are many gaps in the data holdings. H. Dooley had previously distributed a report of 1988 data held at ICES to WG members. 1988 was the most recent year for which it could be said that a large volume of data were at ICES. However, although the shelf sea data submission was fairly complete, little data from the North Atlantic had been received. The WG first briefly reviewed the situation in each country, both from WG members and from information extracted from the ICES ROSCOP data-

base, to get an overall indication of the status, before going on to consider in more detail what could be done to improve the situation. Briefly the status for each country represented is as follows:

Belgium: Not many cruises reported on ROSCOP forms, little data with ICES.

Canada: Data are sent to MEDS from Canadian laboratories, but they do not seem to be sent onward to ICES. ROSCOPs are filled in for a few cruises, but ROSCOP type information is supplied to MEDS for all cruises.

Denmark: Denmark reports cruises on ROSCOP and sends data from standard sections to ICES. ICES has acted as NODC for this type of data for Denmark.

Finland: Data up to 1989 have been sent to ICES. Recently all data have been sent back to Finland from ICES.

France: ROSCOPs were sent to ICES in the past, but very few had been received recently. IFREMER are presently reestablishing their French data centre, and intend to send ROSCOPs once they have reestablished their database. Very little data have been received.

Germany: Germany have submitted ROSCOPs regularly and supplied data from about half the cruises reported on the ROSCOPs.

Iceland: Good relations are maintained with ICES and data have always been sent to ICES in the past. Iceland are a little behind with data supply at the present. ICES have recently sent all the Icelandic data back.

Ireland: Little activity in the past, although a few Lough Beltra cruises have been reported on ROSCOP. There are no Irish data at ICES from the last 25 years. However, with the setting up of the Marine Data Centre, it is hoped that both ROSCOPs and data will be sent to ICES in the future.

Netherlands: ROSCOPs are forwarded to ICES, but data supply is patchy. There is no data centre in the Netherlands; some institutes supply data others do not.

Norway: Regular reports of cruises on ROSCOP, but data supply has been erratic in the past. The Norwegian Oceanographic Data Centre has sent data, but these have not been calibrated or corrected. CTD data have also been sent. Much sea surface temperature data have been collected. Data supply to ICES has been discussed at the Institute of Marine Research, Bergen and will improve from 1994.

Poland: ROSCOPs have been submitted fairly regularly and data have been supplied from over half of these cruises.

Portugal: Very few ROSCOPs reported and no data.

Russia: Supplied ROSCOPs and data in the past, but both have decreased over the last 10 years.

Spain: Some ROSCOPs supplied; no data.

Sweden: The data supply was good, but has fallen behind recently. ROSCOPs were sent up to 1988. Data will now be sent annually again, or as quickly as possible for particular experiments. Data are sent from research ships, coastguards and ice-breakers.

U.K.(MAFF): MAFF submit their data regularly to ICES annually, usually in March: 1992 data have been sent. Hydrography and chemistry data are combined. ICES do not receive any of the biologists' data at present. The biologists feel that they do not get anything out of ICES and have not been specifically asked for their data. An inventory of the data available at ICES might encourage them to submit their data.

U.K.(SOAFD) Submit their water bottle and CTD data annually to ICES. The data from the February IYFS are sent as soon as possible. 300-500 stations are submitted annually; feedback is good and queries are resolved quickly.

U.K.(HO): Send XBT data on cartridge annually to ICES and to the US NODC. CTD/water bottle data from OWS Lima are sent to BODC for onward transmission to ICES.

U.K.(BODC): Behind with data submission from NERC laboratories and Universities due to lack of funding and resources over the last few years. However the situation should improve as BODC has a contract with the UK Hydrographic Office to collect in, quality control and bank civilian CTD data. Once this is underway data will be sent to ICES regularly.

U.S.A.: Relations are good with ICES, but data flow varies from very good to very bad. ROSCOPs have been a problem in the past, but these are now being digitised for inclusion in the ROSCOP database. Much use is being made of communication and data transfer by e-mail and ftp. There is cooperation over the data archaeology project.

A sub-group considered the problems of data flow to ICES. L. Rickards had produced a table showing how many cruises reporting CTD and classical water bottle had passed these data on to ICES. This covered the last 25 years. One major problem seemed to be the lack of awareness of what data were required, how they should be sent, etc. Although this information was undoubtedly available it did not seem to be reaching the right people or was not passed on as people moved to new jobs. The sub-group agreed that the first step should be to increase awareness of what was required, linked to showing what was already available. To do this the WG will send out a package to all members of the Hydrography Committee and its WG Members:

a) Summary table showing the present status of data submission
b) A statement detailing what data types are required (classical water bottle, CTD/STD)
c) A statement of how the data can be submitted (ASCII files, ICES format, ftp, floppy disk etc)
d) Stress that data can be restricted for a period and that ICES will contact scientists about any requests for data collected over the past 10 years
e) Stress that timely submission of data are important for climate studies and global change
f) Summary maps showing the data held by ICES in each country
g) Advertise and offer the ROSCOP database on floppy disk or by ftp

In addition, it was agreed, at the suggestion of R. Gelfeld, that a presentation should be prepared for the Statutory Meeting in Dublin, presenting the summary table of the status in each country, including the date of last submission of data, with a brief explanation. However this was only scratching the surface of the problem, and it should be vigorously followed up in subsequent years. With this in mind it was agreed to take 1990 as a sample year for each WG member to chase up the data from their own country and get it to ICES by the time of the next meeting. WG members could then report back their success (or failures and problems encountered). Of course, data submission from all years should be actively encouraged; choosing one year as an example should in not restrict data submission to that one year.

7. Consider ways of enhancing the utility of data archived at the ICES Service Hydrographique

This item again related to improving communication between scientists and ICES; and, in particular, of demonstrating what data are available. A sub-group looked at two products currently available to assess their usefulness in advertising data and products available from the ICES Oceanographic Data Bank.

The first of these was the UK Digital Marine Atlas (UKDMA), which includes a set of charts showing temperature, salinity and nutrients in the North Sea from the ICES Young Fish Survey Data. A useful feature of this package was the link to a description of the data, but UKDMA does not cover a wide enough area for ICES and there are no analysis facilities (i.e. you can view the charts, but cannot access any data). However it was felt that this was a very useful way of raising the profile of ICES.

The Annual Cycles package developed by the SOAFD Marine Laboratory was also reviewed. This was developed to include different features from UKDMA. The Marine Laboratory had taken all of their hydrographic data, dividing it into boxes and 10 year periods. The package then allows the display of parameters against time Monthly means and standard deviations are plotted and the data are also available in tabular form. It is a useful package for displaying data in a highly compressed form.

The sub-group reported back that these packages were both useful, but it would be a major task to develop either system for ICES. L. Rickards explained that the next phase of UKDMA was under discussion and one possibility for future development was to divide the product into 'chapters'. It currently contained 463 charts and the next version would inevitably contain many more, and it was thought that dividing it up into specialist chapters would allow users to select those of interest. Thus, it would be quite possible to include an ICES chapter and expand the coverage of ICES data sets. For example, location information showing where data had been collected, colour coded by country to show which country had collected the data. This could

be included in a series of (queryable) charts covering, say, periods of 10 years per chart.

There was some discussion of producing a CD-ROM containing all of the ICES profile data and the ROSCOP database. ICES receives approximately 50 requests a year, about half of these are for products rather than the data, and the requests themselves are very varied. It was thought unlikely a single product would help. In addition new data are continually being added to the data bank. However the Group agreed that it would be most useful to know what data were available (i.e. in the form of a list of stations). H. Dooley demonstrated a PC package which he has developed which plotted out station location for a particular year (or years) or area. This was felt to be an excellent way of indicating what data was available, and the Group agreed that this development should be greatly encouraged and the product distributed as widely as possible. The software package and data occupies about 8 Mbytes of hard disk.

8. Evaluate the utility of different software packages for databases in oceanographic data management

A sub-group of the WG reviewed the OCEAN-PC package, which is being developed by IOC, with extensive contributions from the ICES Oceanographic Secretary. The package currently comprises a series of programs for data entry (including ROSCOP entry and search), conversion between different formats, display and quality control of profile data and plotting locations of stations. This software, provided by ICES, is used by them in-house. It was felt timely to test out and review what was currently available in the package, as there was to be a meeting in Copenhagen in the next month or two relating to OCEAN-PC development.

Overall comments about the package from the sub-group were very favourable. The main omission is a manual (including a general overview), which will be dealt with at the forthcoming meeting. Those in the sub-group had not used the package before, and despite the lack of a manual were able to test out the package successfully. They found that, after some initial teething problems, the package was easy to use, and generally self-explanatory.

However, there was inconsistency in designation of keys for help and/or information in different programs within the package. This should be standardised. (At present F5 is used in ROSEARCH, format conversions and TSLOOK, and F6 is used in ROSIN and ENTICE.)

ROSIN, the ROSCOP entry program was tested out quite thoroughly. The sub-group reported that it would be useful to have help available for individual fields, as in the help screens presently available, individual items tended to get lost. For example, J. Atkinson reported that the sub-group had problems in finding out the format of latitude and longitude fields, although this information is available. Some problems were encountered in entering a file id., but this was put down to inexperience (and was thought to have been caused by accidentally including a hidden character in the file name). The sub-group were in favour of field-by-field validation, and some clues when incorrect information was entered. Obviously not all fields could be checked (e.g. plain text including principal scientist's names, laboratories, etc.,), but latitudes, longitudes, data type codes, dates and Marsden square numbers could all be validated. There were also comments that with the grid used for calculating Marsden squares, a land outline would be a useful addition.

A few problems were encountered using the menuing system, but this was an early version of the software, which has now been changed. However on the (new) IOCOPC front end, highlight bar does not appear when you first go to a menu, although it appears when the cursor keys are pressed.

In addition to the OCEAN-PC appraisal, some general discussion took place in respect of the different databases and systems in use at the various centres and laboratories for storage of data. Much use is made of the Oracle and Ingress relational database management systems. They are easy to use, but H. Dooley pointed out that users should be aware of some of their pitfalls. He said that the WG should be aware of the implications of using such database management systems for data exchange. For example, exporting data using report generators may miss out index information and the resulting file may have little structure.

R. Gelfeld reported that, in the past at the US NODC, different data sets were held in different formats, making search and retrieval of data time consuming. They are now moving to modern technology, and will be using SYBASE (a relational database system) to store all their data in one database on Unix workstations. The database will be on-line, with data stored on optical platters (500Mbytes per side) in a jukebox. Data will be put into appropriate formats on extraction.

One further consideration was the use of electronic communication for both mail and file transfer. Compression techniques ('compress' on Unix and 'pkzip' on DOS machines) were very powerful and useful for the transfer of data. Those connected to Internet can make use of these facilities for quick and easy transfer of data files.

9. Review and report on available coastline and bathymetric data sets

There is a need for digitised bathymetric and coastline data sets for the North Atlantic and Nordic Seas for both modelling work and oceanographic work in general. The Oceanic Hydrography WG (OHWG) had asked MDM to provide some information on the available data sets. The Group identified a few global data sets and most WG members knew of, or had access to, coastal and bathymetric data sets covering their regions. A summary of these is given in Annex 2(b) and will be forwarded to the OHWG. In addition to these there is a Russian data set for use with some PC software which may be available, and also a Baltic coastline and bathymetry data set held by the German Bundesforschungsanstalt für Hydrologie.

Limitations of some of the global data sets were noted. These included the World Vector Shoreline, which does not go in to estuaries and the DMA (CIA WDB 2) data set which has significant landmarks missing, for example one of the Faroe Islands. R. Gelfeld agreed to check and let the WG know when a new version becomes available.

B. Hansen asked whether it would be possible to export bathymetric data from the planned GEBCO CD-ROM in a form suitable for use with other applications (for example, Surfer, PV-Wave, Uniras). L. Rickards promised to investigate this. H. Loeng and J. Blindheim thought that the Hydrographic Office in Stavanger had bathymetric information for the North Sea and, if required, would investigate this further.

10. Any other business

i) International Current Meter Inventory

The Second Edition of the International Current Meter Inventory and software was distributed nearly two years ago and it is now time to update it once again. L. Rickards requested that members of the WG, who had put in a lot of effort to get this project off the ground, should update their entries. She would be writing to remind them of this very soon. Version 2.0 contains references to over 29000 current meter series from 16 countries and has been distributed to 84 scientists in 31 countries. It is hoped that the number of countries included will increase for the third edition. H. Dooley asked how easily accessible the data referenced in the inventory were. L. Rickards replied that this varied from country to country, and when she wrote to ask for updates she would also request the status of the data so that an indication of availability could be included in Version 3.0.

ii) Thermosalinograph data

K. Medler posed the question of collecting underway data, primarily temperature and salinity data, asking particularly what accuracy of data was obtained, and whether these data were archived at data centres or at ICES.

E. Henderson reported that the SOAFD Marine Laboratory collects temperature and salinity data *via* hull intakes on both of its research vessels. In the past a Bisset Berman thermosalinograph recorded data onto charts; now TSG103s are used with the data logged on floppy disk. Little checking has been done on the data and the accuracy is unknown. However the Laboratory have now built up a large archive of data which they are trying to check out.

J. Szaron reported that thermosalinograph data are collected on Swedish Baltic shipping routes; the data are used to produce maps and are archived at the SMHI head office. H. Loeng reported that thermosalinographs are used on Norwegian coastal ships and vessels crossing the North Sea. The instrument used was manufactured by IMR, Bergen, and thermosalinograph units are exchanged when the ship returns to port. The accuracy of the data is unknown. R. Olsonen added that for the last year Finnish passenger ships and research vessels have been fitted with instruments measuring temperature, salinity and chlorophyll whilst the ship is underway. J. Wallace said that Lough Beltra, the Irish Research vessel, has a thermosalinograph whose accuracy is unknown.

In BODC, since the advent of Community Research Projects, surface underway data collected from the ship's non-toxic supply have been quality controlled and archived. The data collected include temperature, salinity, fluorescence, bathymetry, oxygen, transmittance and irradiance. The thermosalinograph used for most of the North Sea Project cruises was an autoranging TSG103, but for some of the follow-up cruises this was replaced by an older, manually ranged instrument which measured salinity directly and considerable amounts of data were lost when it was switched to the wrong scale. All of the data from 38 North Sea cruises, together with those from 11 UK JGOFS cruises, have been visually checked using BODC's screening package.

iii) Topics for the next meeting

The following items were suggested for inclusion in next year's agenda

a) Assess the 1990 oceanographic data sent to ICES by each member state, identify problems and suggest solutions
This follows on from this year's discussion and will allow members of the WG to assess the situation in their own countries. Although 1990 has been chosen as a sample year, the assessment is in no way restricted to this year.

b) Review progress in the implementation of IOC's Global Oceanography Data Archaeology and Rescue Project (GODAR) in each ICES Member State
Each WG member will report on progress in sending data to ICES, not just for recently collected data, but actively searching out valuable data sets collected in the past, but not passed on to any data centre for archiving. It was suggested that IOC might be invited to co-sponsor this and be represented at the next meeting. It might be particularly useful if IOC could fund the attendance of a data manager/scientist from, for example, Poland or Russia, who might not otherwise be able to attend. The intention would be to produce a list showing who has what data where, and plan how easy it would be to transfer those data to ICES and the appropriate national data centres.

c) Report on experiences in exchanging ADCP and SeaSoar data, between data centres/laboratories and ICES, using the ICES guidelines
To check out the guidelines produced for ADCP and SeaSoar data, those centres collecting or archiving these data should exchange samples of them. If this proves successful then the guidelines can be circulated more widely. If any problems occur, then solutions can be investigated.

d) Critically review operational procedures for oceanographic data centres in ICES Member Countries
There is a need to consider in some detail how data centres work and whether they can be improved. In addition, there is a need to consider such topics as the procedures for data acquisition, quality assurance and data flow between data centres and ICES. This is particularly important where new marine data centres, like the one in Ireland for example, are be established.

e) Consider the problems solved (and created) by the use of new technology and databases in member countries
As more centres and laboratories move to using commercially available relational database management systems and to new technology in the form of new computer facilities and exchange media (e.g. DAT tape, floppy disk, cartridge, optical disk, CD-ROM, anonymous ftp), the WG thought it timely to consider the progress made in member states.

iii) Time and place of next meeting

The WG expressed its wish that the next meeting should be held at the Institute of Marine Research, Bergen, between 21 - 23 April 1994. This follows on from the Oceanic Hydrography WG, allowing continued cooperation and interchange of ideas between the two working groups.

The Chairman closed the meeting by thanking the participants for their active and valuable contributions. On behalf of the WG, she also thanked E. Henderson for an efficiently arranged meeting.

Annex 1 Agenda

C.Res. 1992/2:35

The Working Group on Marine Data Management (Chairman: Dr. L.J. Rickards, UK) will meet in Aberdeen, Scotland, UK from 22-24 April 1993 to:

a) assess the state of oceanographic data quality control and data processing, taking into account the present performance of instrumentation;

b) examine data flows in ICES Member Countries and make recommendations to improve the situation;

c) consider ways of enhancing the utility of data archived at the ICES Service Hydrographique;

d) review and report on available coastline and bathymetric data sets;

e) further develop guidelines for the data management of ADCP and SeaSoar data;

f) evaluate the utility of different software packages for databases in oceanographic data management;

Annex 2(a) Highlights from the reports of the Data Centres

ICES

ICES was represented at the 14th Session of IODE by the Oceanographic Secretary. As a result of this session there will be two meetings at ICES in May; one concerned with IOC's OCEAN-PC and the other a sub-group of GE-TADE on modern formats. ICES is also being represented at the Global Oceanographic Data Archaeology and Rescue Project in Obninsk. Data archaeology is a topic that ICES has been actively involved with in recent years. ICES is now an active partner in the GLOBEC programme. In addition, scientific analysis of existing data sets has been carried out for the North Sea Task Force. During the year support has been given to a number of groups making use of data and products from the data bank.

Canada

Physical oceanographic data are collected routinely by the Institute of Ocean Sciences, Sidney, B.C.; Bedford Institute of Oceanography, Dartmouth, N.S.; Maurice LaMontagne Institute, Quebec and the North West Atlantic Fisheries Centre (NAFC), St. John's, Newfoundland. The Marine Environmental Data Service (MEDS) acts as the national archive centre for the Department of Fisheries and Oceans (DFO) and also participates in international data management through projects such as GTSPP. The physical oceanography section of DFO at NAFC is responsible for the purchase and maintenance of all oceanographic instrumentation at the regional level, as well as for quality control, processing and archival of all data.

Denmark

The primary task of the Oceanographic Department in the Royal Danish Administration of Navigation and Hydrography is to collect, analyse and distribute oceanographic information to shipping in Danish, Faroese and Greenland waters. In 1992 the Oceanography Department continued the activities initiated in 1990 and 1991. These include establishment of oceanographic sensors in Danish waters with data communication in real time, feeding a marine database with quality controlled data, tidal analysis and prediction and application and improvement of remote sensing techniques for producing sea surface temperature maps of Danish waters. The Department is also involved in Nordic WOCE and EDMED.

Finland

The Finnish Institute for Marine Research have several data bases for hydrographic, chemical and biological data. Recently a new Paradox database has been obtained. The plan is to connect this database to the data collection system on-board ship to get all the observations into the database immediately. A checking system will be added to this. Hydrographic and chemical data from 1962 onwards are now being checked and comparisons are being made with the Finnish data held at ICES. In addition CTD and ADCP databases will be created in the near future.

Germany

Following the recommendation of the German Science Council, the Institut für Meereskunde, Warnemunde, was reestablished as the Institut für Ostseeforschung (IOW). This Institut is responsible for Baltic Monitoring Programme measurements. During 1992 entries were compiled for EDMED from IOW. DOD is responsible for archiving German marine data, so data collected by IOW are transferred to DOD. Data collected by the old Institut für Meereskunde will be transferred to DOD over the next few years.

Iceland

Approximately 400 CTD stations were made in 1992 on Icelandic research vessels, usually including chemical measurements. Current meter measurements have been made as part of the Greenland Sea Project and will continue as part of Nordic WOCE. Data management has centred around loading existing data into an Oracle database. This work is now well underway and will continue this year. Phytoplankton data are being treated in a similar fashion.

Ireland

The Irish Marine Data Centre was established on April 1st 1993. It is currently looking for suitable premises and will soon be recruiting staff. The Marine Data Centre intends to manage data in five areas as follows: data electronically recorded on-board Lough Beltra; the expansion and maintenance of the Irish edition of EDMED; the management, initially of key data sets, progressing to the management of a wide range of scientifically valuable data sets; the management of scientific documentation relating to Irish based research as well as general methodology and quality control procedures for marine research; the contract management of project specific data and information.

Netherlands

MARIS has no data of its own; all the Dutch data are scattered over various institutes. In 1992 MARIS established a working relationship with the Netherlands Foundation for Sea Research (which finances all major oceanographic expeditions), and with the Netherlands Institute for Sea Research to register all major scientific cruises in the MARIS database. Information from 1989 onwards is now being added to the database. Eventually MARIS will have a total overview of where data were collected and are stored.

Norway

During 1992 and 1993 an integrated model for marine research data was developed at the Institute of Marine Research. The intention was to establish an integrated database (using the Ingress relational database management system) with uniform data representation and structure that is flexible enough to represent all kinds of marine data. In connection with this, a project on quality assurance has begun. In addition, manuals for

operating instruments, handling and analysing samples will be prepared.

Spain

The Spanish Institute of Oceanography maintains an Oceanographic Data Centre, which was established in 1968. It is responsible for the compilation, storage and distribution of the data produced by the different research areas of the Institute. The Centre has full access to the mainframes of the Institute and maintains a databank or marine data (mainly hydrography and chemistry). At present the main purpose of the Data Centre is to identify, manage and store the data sets of interest, distributed around Spain, for the development of marine research. To support this, some reorganisation is taking place and a new structure will be developed over the next four years to meet the needs of Spanish marine investigators.

Sweden

The main activities over the year have included a change of computer system and conversion of the oceanographic databanks from random access files to a relational database system (MIMER). Historical data from Swedish lightships are being incorporated into the database. Much effort has been devoted to SKAGEX and the Swedish-Finnish project 'Gulf of Bothnia Year'.

U.K.(BODC)

During the year BODC has published Version 2.0 of the UKDMA and a CD-ROM containing the data collected during the NERC North Sea Project. BODC has been coordinating the CEC/MAST funded European Directory of Marine Environmental Data (EDMED). A digital version GEBCO (5th Edition) charts was now available and in the coming year a GEBCO CD-ROM will be published. Also in the coming year a BOFS (i.e. the UK contribution to JGOFS) CD-ROM would be published. BODC operates one of the WOCE Sea Level Data Assembly Centres.

U.K.(H.O.)

The Hydrographic Office (HO) acts as the UK National Data Centre for XBTs, and processes Naval bathythermograph data from the UK, Netherlands, Denmark and Norway, and also some civilian data. The HO have started a reprocessing programme for their data holdings, beginning with defence related areas.

U.K.(MAFF)

The emphasis of MAFF's work is on biological studies, including fisheries stock management. The Laboratory has a small oceanographic section. CTDs are frequently used; the data processed by the oceanographic group is forwarded to ICES. The data collected by the biologists may not be calibrated, as they are not always interested in absolute values. In addition, they often use the CTD mounted in a towed body, which makes the collection of calibration data difficult.

U.K.(SOAFD)

February 1992 saw the end of a 15 month study in Loch Linnhe, aimed at examining nutrient flux through the ecosystem for one year. A wide variety of data has been collected, which has brought its own set of problems in terms of data management. New software has been developed to screen, calibrate, edit and manipulate the data sets. A long term project has just started to develop more user friendly, graphical methods of accessing the current meter and hydrographic data collected by the Laboratory.

U.S.A.

The US NODC receives data collected by NOAA and other US federal agencies. NODC also acquires data through bilateral exchanges with other countries and through WDC(A)-Oceanography. NODC receives both delayed-mode and near-real-time data. Over the last few years, the NODC has begun a project to augment the historical digital data archives by seeking out and recovering manuscript and digital data not yet included in databases accessible to the world research community. As a result of this data archaeology, the NODC has begun to receive significant new data accessions.

Annex 2(b) Summary of bathymetric and coastline data sets available

ICES:
The DMA (CIA WDB 2) data set is used for OCEAN-PC; it has 1km resolution and is available as a binary file of size 1.2Mbytes. Software is included in OCEAN-PC to extract and plot areas. The data set includes rivers, coastlines, inland waters, country boundaries, etc. It is available over Internet.

Canada (East coast):
NAFC, St. John's has digitised the bathymetry of the Labrador region, with finer resolution for the shelf area, although this is incomplete in some places. The data are maintained in an ASCII file and also on a PC based GIS system. The data set is the property of the Government of Canada.

USA:
National Geophysical Data Center/WDC A Marine Geology and Geophysics, Boulder, Colorado publish material on CD-ROM. Further details are available from R. Gelfeld. The US NODC use the CIA high resolution landmass; the DMA data set is used also.

Iceland:
Part of CIA coastline is used; MRI also have depth contours around Iceland. Data from NOAA EPOCH5 on a 5' grid (available on DAT and cartridge) are also used.

Faroes:
Covering an area around the Faroes - high resolution bathymetry; a large effort has been put into digitising this; it is not yet finished.

Norway:
IMR Bergen has several packages for data display in map form. ITAKS, developed in-house, is the main one in use. Maps cover the Norwegian Coast, Barents Sea and the North Atlantic. The map databases available include the following:

In addition the Norwegian Hydrographic Office in Stavanger use the commercially available MacSea package, available for Macintosh. It covers the North Atlantic and is very flexible. It has the capability to import and export data, although importing data is not straightforward.

Denmark:
The coastline of Denmark is archived in digital form; digitised from Ordnance Survey mapsheets (scale 1:25000). The coastline of Greenland is kept in hardcopy form, but digitisation is in progress. The Royal Danish Administration of Navigation and Hydrography is responsible for all hydrographic surveying in Danish and Greenland waters. Since 1952 modern equipment has been used and to date 25% of Danish waters has been covered. In 1992 multibeam surveys were introduced. The bathymetric data are stored in scale free digital form. Information from older surveys is held in hardcopy form, at a scale of 1:20000. These sheets are being digitised. In 1978 hydrographic surveying in Greenland waters stopped, but it was resumed in 1989, with the highest priority given to the most heavily navigated areas.

Sweden:
SMHI has a high resolution coastline of Sweden, including rivers and lakes. The Institute of Marine Research, Lysekil has a fairly coarse bathymetric data set of the Baltic and Norwegian coastline.

U.K.:
World Vector Shoreline (WVS) from US Defense Mapping Agency: Scale 1:250000. Shoreline features are located to within 500m. (held by BODC, SOAFD and HO)

GEBCO 5th Edition: Published as 16 printed sheets on a Mercator projection at a scale of 1:10 million and 2 sheets on Polar stereographic projection at a scale of 1:6 million. All have been digitised and are available on magnetic tape (or floppy disk for small extracts). Contours are at depths of 0m, 200m, 500m, 1000m and at 500m intervals thereafter. Later this year the GEBCO data set will be available on CD-ROM.

Name	Area	Resolution	Topography	Scale
AKUP	69.5°N-74.5°N, 15°E-35°E	15'x 1°	CD	unknown
BARENTSHAV	62.0°N-82.0°N, 5°E-60°E	1° x 1°	CD	unknown
KARTDATA	45.0°N-81.0°N, 45°W-70°E	30'x 1°	C	variable
MASFJORDEN	60.5°N-61.05°N, 4.7°E-5.9°E	1' x 1'	C	1:50000
STURE	60.5°N-61.0°N, 4.55°E-5.08°E	1' x 4'	C	1:50000

The data set will be updated as new bathymetry becomes available. The GEBCO data set does not include the Mediterranean Sea, because a higher resolution data set is available (International Bathymetric Chart of the Mediterranean at a scale of 1:1 million). (held by BODC)

North Sea Bathymetry: Covering the UK sector. The data (800000 soundings) were supplied by the British Geological Survey (BGS). They have been tidally corrected using a model. Depths are to the nearest metre and positions are to the nearest 0.0001 degrees latitude/longitude. This has been gridded on a 1 minute by 1 minute grid using 'Interactive Surface Modelling (ISM) Version 7'. The data set is available on floppy disk. Future work on the data set will include refining the coastal/estuarine areas. The data set will be extended to cover the remaining areas of the UK shelf and checked against the German sector data set. It will also be gridded on a 1 minute by 1.5 minute latitude longitude grid. (held by BODC)

As part of the UKDMA Project 33 sea lochs on the west coast of Scotland were digitised. This was instigated by SOAFD, one of the UKDMA funding partners, and was required for the development of numerical models to study the circulation and flushing times of west coast sea lochs. The resolution is 100m. This has been patched in to the World Vector Shoreline where possible, although this gave rise to many problems. (held by SOAFD)

MAFF has recently purchased a GIS system (Laserscan HORIZON) which makes use of digitised UK Ordnance Survey and Bartholomew charts (at scales of 1:50000 and 1:250000 respectively). Vector charts are available for the whole of the UK, and some raster charts are available. The disadvantage is that the Laserscan formats do not lend themselves to non-Horizon applications.

Netherlands:

The mapping programme used by MARIS, called the Marine Geographical Information Manager, and developed by a subsidiary of SEATEAM in Den Helder contains digitised coastlines of the North Sea (average scale 1:200000, some areas down to 1:20000), the Baltic Sea (average scale 1:200000) and the Atlantic (average scale unknown). These maps have been digitised from existing hydrographical maps - some by SEATEAM, some obtained from third parties. These digitised maps are for sale from MARIS.

Spain:

The Geomining and Technological Institute of Spain is working on a systematic program of geological map making of the Spanish continental margins and adjacent areas. These data are available and the storage medium is mostly maps, charts and continuous feed paper. The areas covered include: Gibraltar Strait, Alboran Island, Almeria continental shelf (SE Spain), NW Spain continental shelf, Balearic Sea and Southern Spain continental shelf.

Annex 3 Joint meeting between the Oceanic Hydrography and Marine Data Management Working Groups

Several topics of interest to both groups were discussed during the joint session. R. Gelfeld described the data archaeology work currently being undertaken at the US NODC and WDC-A (Oceanography). This led on to a discussion of calibrations, quality control and information accompanying data.

Data archaeology began at the US NODC/WDC-A several years ago; the work so far has concentrated on physical oceanography, locating data sets which have not been sent to national oceanographic data centres and which may be in manuscript form or kept by individual scientists. Within the US NODC the archives have been searched, station location plots have been produced and the NODC has worked with scientists to identify gaps and to fill them.

The project now has international status with the backing of IOC and is known as the GODAR (Global Oceanographic Data Archaeology and Rescue) project. ICES is actively involved with this work and a close relationship is maintained between the NODC/WDC-A and ICES. Copies of the archaeology data set are kept at both ICES and WDC-A for security. The cost of adding data to the archive has been about 3 man years for 30000 stations; this includes not only digitising the data but also the quality control, which is time consuming. It is also important to check that data have not already been digitised to avoid duplication. The data from this project are freely available. Meetings such as the present one are valuable for exchange of ideas and to help track down data.

In response to a question from D. Ellett about metadata, R. Gelfeld said that WDC-A tries to recover as much information as possible. He asked if anyone knew of any large data sets, which may be at risk. G. Becker said that large volumes of data were collected by Germany prior to the formation of DOD, which were only partly at the data centre. R. Dickson volunteered to search for a data set which he remembered, but no longer knew the whereabouts of; this was a series of surface salinity values on a 5 degree grid for the North Atlantic (the data were derived from the continuous plankton recorder).

In reply to questions about restrictions on the data, H. Dooley said that data at ICES could be restricted, (i.e. data could be sent to ICES and no further). In addition, scientists were informed of any requests for data up to ten years old.

Over the last few years, interest in the high salinity anomaly (1989/90) meant that scientists were interested in comparing with other earlier data. Searches at ICES and WDC-A revealed many gaps in the data record which need to be filled.

Concern was expressed by D. Ellett about the information accompanying data, and he volunteered to search for copies of (old) papers with information about calibrations, intercalibrations and methods. The Working Groups agreed that this would be valuable but recognised that it is often hard tracking down the data sets - the metadata is even more difficult and harder to check. ICES does not store metadata with the data themselves, but includes comments in the ROSCOP database. J. Blindheim recalled how, in the past, standard reporting forms were used, but this seemed to have fallen into disuse. Information like the type of instrument, for example, should be included. S. Narayanan reported that, in Canada, each laboratory or scientist had had their own system, but now they had all decided on a standard form to be included in the header. Mention was also made of the WOCE procedures for hydrographic data. The data quality experts were provided with quite detailed cruises reports which were very useful and formed part of the quality control.

The WHP guidelines were thought to be useful and it was generally recommended that scientists should send information to accompany the data. In the USA, funding was often dependent on the appropriate documentation being produced. ICES recommended that documentation should be short and to the point. The use of the SCOR Working Group 51 Guidelines was encouraged. It was felt that part of the problem was that as people change jobs and move on their expertise often goes with them and guidelines for data submission and accompanying information would help to alleviate this.

H. van Aken was concerned about data submission to the WDC-A from ICES countries where NODCs did not function well. In cases like this, the data can be sent directly to WDC-A. There was also some discussion about how data should be sent ICES. Some years ago the MDM Working Group had produced guidelines for sending reduced resolution data to ICES; however ICES will accept high resolution data (1 or 2 dbar interval). Data sent to ICES are carefully quality controlled and any problems sorted out with data originators. Data quality control guidelines need to be set up and this is an important part of the processing. For the GTSPP an essential element is the quality control; however this is carried out by a panel, and the data originators are not consulted. It was felt that it would be a better approach if the data centres remained in close contact with the data originators when carrying

out their quality control and reported back any problems.

H. van Aken briefly asked about ways of organising CTD data on a PC. There was brief discussion of the value of relational databases, which can be very useful and easy to access, but can cause problems in the longer term when the database systems are updated or become obsolete - then access to the data can be very difficult. However, they are very useful as working databases.

Following on from the discussion at last year's joint meeting L. Rickards briefly described a software package, developed by the University of Hawaii for processing, quality control, display and storage of shipborne ADCP data. This package is available for Unix systems and IBM PCs. It has mostly been written in-house, but does use a commercial package (Matlab) for some functions. The University of Hawaii was keen to let others try out the package and obtain feedback. Further details can be obtained from L. Rickards. She has a copy of the software package and manual, but has not yet had the opportunity to test it out.

Annex 4 Recommendations

Proposed Agenda for next year's meeting

a) Assess the 1990 oceanographic data sent to ICES by each member state, identify problems and suggest solutions;

b) Review progress in the implementation of IOC's Global Oceanographic Data Archaeology and Rescue Project (GODAR) in each ICES member state;

c) Report on experiences in exchanging ADCP and SeaSoar data, between data centres/laboratories and ICES, using the ICES guidelines;

d) Critically review operational procedures for oceanographic data centres in ICES Memeber Countries;

e) Consider the problems solved (and created) by the use of new technology and databases in Member Countries;

A representative of IOC will be invited to attend.