International Council for
the Exploration of the Sea

C.M. 1994/C:2
Hydrography Committee

Report of the

WORKING GROUP ON MARINE DATA MANAGEMENT

Bergen, 21 - 23 April 1994

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

1. Opening of the meeting

The meeting was opened at 0930 on 21 April 1994, hosted by the Institute of Marine Research, Bergen. Participants were welcomed by the WG Chairman and the Director of the Institute, Dr. Roald Vaage. In his welcome, the Director noted that the Institute had been setting up a database, in which all types of data would be linked together. It was intended that this would be completed very soon, together with new routines for the data quality assurance. This would help the Institute to fulfil its obligations to ICES. He continued by saying that the Institute monitored physical and biological conditions by means of acoustic methods, CTDs, current meters and trawls in the Barents Sea, the Norwegian Sea and the North Sea. He hoped that the WG discussions would provide ideas on how the Institute could proceed with its work of storing and presenting its data. The Director concluded by wishing the WG an enjoyable stay in Bergen and hoped that they would return in the future. H. Loeng then provided a brief introduction to the Institute of Marine Research and also explained the local arrangements.

Members of the Working Group present were:
S. Almeida, Portugal, J. Atkinson, UK, M. Fichaut, France, R. Gelfeld, USA, K. Jancke, Germany, N. Kaaijk, the Netherlands, L. Lastein, Faroes, H. Loeng, Norway, K. Medler, UK, S. Narayanan, Canada, P.B. Nielson, Denmark, R. Olsonen, Finland, M. Ostrowski, Poland, L. Rickards, UK (Chairman), H. Sagen, Norway, J. Szaron, Sweden, H. Valdimarsson, Iceland and J. Wallace, Ireland. The ICES Oceanographic Secretary, H. Dooley, was also present for part of the meeting. Apologies for absence were received from J. Blindheim, Norway, N. Hakansson, Sweden, E. Henderson, UK, and L. Smit, the Netherlands. S. Lygren, K. Bakkeplass, K. Seglem and O. Strand from the Institute of Marine Research also attended the meeting.

2. Adoption of the Agenda

The agenda for the WG meeting was adopted as a resolution of the Statutory Meeting in Dublin (C.Res. 1993/2:14, 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 and the reports were distributed to WG members, together with the report of the ICES Oceanography Secretary. The WG noted with pleasure the re-establishment of data centre activities in France, and the continuing work in Portugal after a period of reduced activity. Members of the WG were pleased to report that there was much activity on the data management front, both in terms of the assembly of high quality data sets and in the establishment of data bases.

4. Assess the 1990 oceanographic data sent to ICES by each Member State, identify problems and suggest solutions.

H. Dooley distributed an extract from the ICES ROSCOP database, updating the previous one circulated in December to WG members, indicating the status of data collected in 1990. He noted that the situation was quite good with 14184 stations supplied to ICES. However there was still quite a large amount of data outstanding; in particular there were major gaps in the German data set. Also, where no ROSCOPs were supplied, it was difficult to chase up the data.

Some discussion took place relating to country codes, as there was some confusion over the code for the old German Democratic Republic. H. Dooley confirmed that a revision of some country codes had been agreed by the IOC's committee on International Oceanographic Data and Information Exchange (see Report of IOC Committee on International Oceanographic Data and Information Exchange Fourteenth Session, Paris December 1992). The new set of codes would soon be available via ftp from ICES. Not many codes were affected.

Much variation in data supply was noted between the individual countries. The Faroes and Ireland were both in the process of compiling ROSCOP forms for their cruises. The Faroes have not yet supplied any data and Ireland has only supplied data from its joint work with Russia. The Fisheries Laboratory on the Faroes has been moving its CTD data holdings from an HP1000 to a PC and converting them to the same format. Quality checks have indicated unexplained non-linearities in the temperature sensor on the CTD and calibration is still in progress. The Irish Marine Data Centre will start sending data before too long, starting from the most recently collected data.

The situation for Norway, Denmark, Sweden, Iceland, Poland and the UK Fisheries Laboratories was good, although Iceland does not submit ROSCOPs for non-physical data; the Norwegian universities are not very good at supplying data and some data from polar cruises were missing from the Polish data. Germany has good coverage with ROSCOPs but is currently lacking resources to process the data. ROSCOP and data submission from Spain is very patchy, usually only for EC MAST cruises; they also have a lot of uncalibrated data. The Netherlands and Belgium supply data for the

Joint Monitoring programme (and the Young Fish Survey - the Netherlands) but with the exception of one scientist in the Netherlands, little data (or ROSCOP forms) are submitted. Russia is currently not releasing data. In Canada, the universities are bad at submitting information and data to MEDS, but the Department of Fisheries and Oceans laboratories do submit their data to MEDS. S. Narayanan expressed some surprise that more information was not available from Canada at ICES. R. Olsonen agreed to check why there were data reported for only four Finnish cruises for 1990. The UK submission was not good for NERC research ship cruises; BODC was making progress in obtaining the data, but had not yet quality controlled them. The USA, while cooperating closely with ICES, has only submitted directly requested data to the ICES data bank. R. Gelfeld said that some effort would be made to do this for CTD data.

The WG felt that real progress had been made during the year, resulting in data being forwarded to ICES. It was decided to continue close monitoring of the data flow, concentrating on data collected post-1990 - while not forgetting to continue efforts to chase up outstanding 1990 data. At the next meeting WG members should report back on further progress and problems, perhaps with a target of getting up-to-date by the time the ICES Data Bank reaches its 100th anniversary in eight years time. H. Dooley reminded the WG that nutrient and other chemical data should be submitted as well as temperature and salinity. He also noted that the appropriate parameter for the z-coordinate was pressure, not depth, as recommended by the Joint Panel on Oceanographic Tables and Standards (JPOTS). ICES had been receiving data in a mixture of units, which had been causing some problems.

5. Review progress with the implementation of IOC's Global Oceanographic Data Archaeology and Rescue (GODAR) Project in each ICES Member State.

R. Gelfeld introduced the GODAR Project discussions by giving a review of progress achieved so far; further details of this are given in Annex 3(a), together with a map showing the distribution of chlorophyll data received. In summary, digital oceanographic data covering long time spans are a requirement for climate and global change research, and at present much historical data exist only in manuscript form at the collecting institutions. In order to address this problem the US NODC and WDC(A) have initiated the Global Oceanographic Data Archaeology and Rescue (GODAR) project. This is being organised and conducted in cooperation with the Intergovernmental Oceanographic Commission (IOC). ICES is also playing a major role. Some workshops have been held and more are planned and a report of the initial results has been published. Information on which data are held by GODAR can be obtained from R. Gelfeld in the forms of maps or lists of cruises. Data will be made available in the near future on CD-ROM and possibly over computer networks. Data from the Data Archaeology project are in the public domain.

L. Rickards suggested that WG members could consider the following questions when they were starting data archaeology projects: Do you have data which has not been sent to WDC(A)? How do you find out if you do not know this? What data are required? How do you track them down? R. Gelfeld had answered some of these questions in his introduction and various publications including IOC/UNESCO Report No. 58 'Sources for Historical Oceanography' and an ICES report to the Ocean Data Archaeology workshop in 1990 are also useful pointers. Entries in the EDMED directory may also prove of use.

The status of data archaeology in member states was reviewed country by country: a summary of this is included in Annex 3(b). Some discussion followed about whether the data are quality controlled by NODC/WDC(A) and whether metadata are included. J. Szaron asked about problems with ships which had been coded as 'unknown'. R. Gelfeld replied that he was trying to sort out this problem with ICES and eliminate any duplicates. Both ICES and WDC(A) have software for checking for duplicate data series. H. Dooley said a list of ship codes would soon be available on the ICES public directory. New ship codes can be obtained through ICES - the ship name and call sign is required for a new code to be assigned.

There was a problem of duplication of effort, in keying in manuscript data. It was always worth checking with ICES or WDC(A) to make sure that data were not already available in digital form. R. Gelfeld encouraged MDM WG members to contact him and H. Dooley for details of data already held by WDC(A) and ICES. H. Dooley reported that he had completed reviews of the data held by ICES/WDC(A) (for data collected up until 1968) for Denmark, Iceland, Finland, Sweden, Spain and Portugal. Reviews of the UK and the Netherlands are in progress, while Norway is proving more problematical. He has been quality controlling and combining the WDC(A) water bottle/nutrient data set with the data held by ICES. Data held by ICES are noted in the ROSCOP database for data collected after 1967. Pre-1967 data will have ROSCOP entries created from the data holdings.

R. Gelfeld recommended that MDM members chased up data from the 1960s, when a lot of data had been collected but had not been passed on to the appropriate data archives. There was some discussion as to whether the data should be sent to ICES, to WDC(A) or to both. R. Gelfeld recommended that data from ICES member

states should go to ICES first and H. Dooley agreed. If this was likely to cause any problems, then the data supplier should discuss this with WDC(A) and ICES first. Other problems included the potential loss of data when scientists moved on or retired, especially when data collection and processing methods are not stored alongside the data.

The WG were pleased to hear of the progress of GODAR and members will continue to search out data to forward to ICES/WDC(A). The WG looked forward to further updates on the progress of data archaeology.

6. Report on experiences in exchanging ADCP and SeaSoar data, between data centres/ laboratories and ICES, using the ICES guidelines.

The guidelines previously compiled by the WG were distributed, prior to sub-groups considering them in more detail and also to provide a starting point for the discussions. These may be found in Annex 4. L.Rickards reported that BODC had distributed some ADCP data on the North Sea Project CD-ROM, and were now looking in more detail at ADCP data as they were storing them for UK WOCE cruises. In addition some SeaSoar data had been forwarded by BODC to the UK Hydrographic Office, as pseudo-CTD casts. She also reported that a reply had been received from the Japan Oceanographic Data Centre, who have been designated RNODC(ADCP), with regard to their plans for ADCP data. This contained an outline of the database they were setting up, which includes meteorological parameters as well as the ADCP data. The RNODC has recently set up an expert group, comprising 7 Japanese ADCP experts, to develop guidelines for collection and data processing for obtaining high quality ADCP data. They were interested in the developments in the MDM WG and wished to be kept up to date with progress.

a) Shipborne ADCP data

A sub-group comprising L. Lastein, K. Medler, S. Lygren, H. Valdimarsson, P.B. Nielsen and S. Narayanan considered the guidelines for the management of shipborne ADCP data which have been developed by the MDM WG over the last few years. To help with this S. Lygren demonstrated the system presently set up at the Institute of Marine Research, Bergen, for ADCP data. Input from the RNODC(ADCP) in Japan was also considered.

ADCP data are collected from vessel mounted units in Norway, Canada, Denmark, Faroes and other countries. In Norway (Institute of Marine Research), an Ingres database on a Unix platform is used to archive binned ADCP data (8m depth bins, 10 minute averages), but this stores only horizontal current velocities (u and v) and per cent good together with position, date/time, ship velocity and water depth. Pitch/roll, automatic gain control (AGC), error velocity, etc., are discarded, as are data files when bottom tracking is lost. In Canada, the averaging and storing depend on individual scientists, but an initiative is underway to organise a database. In the Faroes, all original data are kept. The Japanese have drawn up specific data formats for ADCP data, however these are restricted to data collected at four levels. The University of Hawaii has developed a 'Common Oceanographic Data Access System (CODAS)' which produces quality controlled ADCP data on a Unix machine or PC. This is available over Internet.

It was the view of the sub-group that it was still premature to store data at national data centres. The sub-group recommended:

* Each regional centre/institute should be encouraged to keep all parameters collected, but ensure that header information is complete in terms of position, date/time, etc. If no bottom tracking is available, then the navigation files should be appended.

* Accompanying information on the instrument, collection method and data quality should be in an accompanying computer file.

* Compilation of an inventory of ADCP data; it may be worthwhile to get a list of all the vessels with ADCPs. The RNODC(ADCP) has already compiled an inventory of Japanese institutes collecting shipborne ADCP data. The ROSCOP database can provide information for this in ICES countries.

* The 1995 MDM WG agenda should have an agenda item to discuss ADCP data. The sub-group recommends that each country using shipborne ADCP prepares a paragraph on the header and data format and archival method.

M. Ostrowski mentioned the use of ADCP for investigating zooplankton, and also of calibrating the transducers for this purpose. S. Narayanan noted that integration of ADCP data with tidal models, in areas where tidal influences dominated, was not an easy matter. Other points made in the discussions were as follows: all parameters collected should be stored, a flag to indicate whether bottom tracking was used should be included with the data, precision of the positional information was important, the vessel and type of instrument should be incorporated. The Japanese have included meteorological data in their database scheme; this was not thought to be essential. The MDM guidelines were thought to be sensible; the header information appeared to be complete. It was not

possible to specify what bin average to use; one appropriate to the area being investigated should be used.

The discussion widened to a more general one on current meter data collection and quality control. Although most countries were satisfied with their procedures, it was felt useful for the MDM to discuss this further in the future.

b) SeaSoar (Batfish) data

A sub-group consisting of J. Atkinson, M. Fichaut, M. Ostrowski and J. Wallace considered the guidelines for the management of SeaSoar data previously drawn up by the WG. Their aim was to use some SeaSoar data together with data documentation describing the procedures used to collect and process the data to test out whether they complied with the guidelines and whether the data could be used by others without problem or confusion.

The sample data were supplied as pseudo-CTD casts, although the original 1 second times series was archived at the data centre providing the sample data. The sub-group felt that the guidelines were sufficient and did not require amendment.

Some general discussion followed primarily concerned with calibration methods, identification of SeaSoar data as distinct from CTD profiles and methods of processing the time series to station equivalents. S. Narayanan noted that her laboratory kept SeaSoar data as time series; in addition, they use CTDs towed by trawlers, which are flagged as such in the database. SeaSoar or towed CTDs can be calibrated by using them as conventional CTDs with a rosette sampler prior to towing and near surface water samples can be taken when the SeaSoar is near the surface for salinity determinations. ICES will accept SeaSoar data from areas or cruises where little CTD or classical bottle data are available. It was felt important to consider what is needed for climate studies. It is essential that SeaSoar data should not be confused with other data types (i.e. they should be flagged in some way to distinguish them).

7. Critically review operational procedures for oceanographic data centres in ICES member countries.

J. Wallace was requested to introduce this topic, partly because the Irish Marine Data Centre was fairly new and still developing its procedures, but also because he had visited several data centres to examine how they operated prior to setting up the Irish Marine Data Centre. He stated that procedures vary from place to place, depending on the size and national obligations of the data centre, but there is some commonality. Taking Ireland as an example, some of the procedures were examined. Procedures have been defined for the management of data collected on the research vessel, Lough Beltra, based on the methods developed by BODC during the NERC North Sea Project. In addition, software is supplied on-board to produce ROSCOP forms at the end of each cruise. This means that ROSCOPs and data will be available quite quickly.

It was noted that the IOC has published various guides to establishing and running national oceanographic data centres. Various different approaches were discussed. In some centres, notably BODC and the Irish Marine Data Centre, project oriented data management dominated. This approach often brought with it demands for results from the funding agencies. However, it also meant that those scientists participating in the project had good access to the data, and that the data centre worked alongside the scientists involved. But if the project contained many diverse data types, quality control was difficult to implement, as the centre was unlikely to have expertise in all the areas where data were collected.

H. Dooley stressed that data management should not be just the responsibility of the national centre; it is too important for that. The responsibility should begin with the data collecting scientists and their institute or laboratory. L. Rickards reported that one approach taken within the UK component of WOCE was to assign a scientist on board ship as the data manager for that cruise. That person would then have the responsibility for coordinating the data transfer to the data centre. Perhaps data managers by topic should also be appointed within large projects. M. Fichaut noted that, in the past, the French Data Centre (BNDO) had operated a data analysis unit, which included calibration of instruments.

It was noted that projects now often require access to historical data while the projects and individual cruises are in the planning stages. There is also a marked interest in gridded data sets, especially in relation to climate change work. The GODAR project is already providing benefits in this area.

The WG felt that it was difficult to have the same rules for all, but there should be certain standard procedures for quality assurance, header information, documentation to accompany data sets, etc. The guidelines for management of data produced by the WG provided valuable checklists for this. These had been produced for moored current meter and CTD data some years ago, as well as the more recent SeaSoar and ADCP guidelines.

H. Dooley mentioned an intercomparison of quality control procedures which he had been coordinating with

two other centres. A sample data set had been supplied in two formats. These were translated to the participating centre's in-house format, quality controlled and a brief report produced. This was thought to be a very useful exercise for MDM members to participate in and to report back on next year. H. Dooley agreed to coordinate this and all members of the WG who wished to take part in this should request a copy of the data set.

8. Consider the problems solved (and created) by the use of new technology and databases in member countries.

Two reviews were presented on databases which had been developed over the last few years. One by N. Kaaijk of the DONAR database set up by Rijkwaterstaat (RWS) and the other by H. Sagen illustrating the database set up by the Institute of Marine Research, Bergen. This latter was accompanied by a demonstration. Both of these databases are complex and are available to a large number of users. They have taken between 3 and 5 years to develop and make use of relational database technology.

A sub-group comprising H. Sagen, J. Szaron, N. Kaaijk, R. Olsonen, S. Almeida and H. Dooley considered relational databases in more detail and reported back to the WG.

The sub-group compiled the following checklist of items to consider when setting up databases:

* Involve users/scientists when the database is being set up
* Much care must be taken during the database design
* Try to avoid complex databases; do not try to do everything at once
* Be cautious of 'professionals' selling database systems
* Exploit colleagues so that you do not reinvent the wheel
* Use national/international guidelines when setting up the database tables
* Consider how data should be collected and stored so that you can get out of the database what you want
* It is costly to build complicated databases
* Consider carefully the storage of null data values, so that information will not be lost
* Complicated retrievals need very careful thought

In the ensuing discussion, H. Dooley noted that many databases do not handle numbers very well, which may lead to problems with precision. He also posed the question, what do we use databases for? Is it for ease of retrieval, and should we be using GIS systems? N. Kaaijk stated that you need to know what you want before you start and need to balance quality of service against economics (cost of the system). In the Netherlands the DONAR database is used as a tool, rather than as an archive. M. Ostrowski added that relational databases are more suited to business applications; oceanographic databases do not quite fit into existing systems. K. Medler noted that it had taken 3-4 years to produce the system now in use at MAFF. There were problems with the complex set up, which was serving a large number of groups. The database had been implemented by a central computing unit, and it was not always easy to explain what was required to the computer experts. One needed to know what was required at the beginning, but with long development times, scientists were likely to have changed their requirements by the time the database had been established. J. Wallace said that there was a need to design open ended architecture.

R. Gelfeld and L. Rickards both noted that one advantage of these databases was the quick and easy access to inventory information, which had previously not been easily accessible. N. Kaaijk added that users now knew where to find the data. H. Dooley was concerned about long term security of the data - what happens when Oracle/Ingres, etc. are replaced and the data is only available in that format. J. Wallace thought that the hardware changes were more likely to be a problem than the software.

9. Date and location of the next meeting; topics for discussion.

i) Topics for the next meeting

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

a) Assess the post-1990 oceanographic data sent to ICES by each member state, identify problems and suggest solutions;
This would follow on from this years effort to encourage flow of 1990 data to the ICES Oceanographic Data Bank, and ensure that momentum built up during 1993 would be continued in the future.

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

Much data has been recovered by GODAR already, but many valuable data sets still remain outside of established data banks and archives. WG members need to continue searching out old data sets and forwarding them to ICES and WDC(A). ICES has taken a lead role in this project for the ICES region, which provides a focus for member states activities.

c) Report on procedures for processing and storage of shipborne ADCP data;
WG members whose countries use shipborne ADCPs should assess the data processing, quality control and archiving procedures in use and report on these. Cooperation and exchange of information with the Japan Oceanographic Data Centre (RNODC(ADCP)) and the University of Hawaii will continue.

d) Critically analyse data processing procedures for moored current meter data in ICES member countries;
There is a need to check standards and compare methods of data quality control for current meter data, for although WG members have their own procedures in their respective countries and/or laboratories, there is no international centre responsible for moored current meter data.

e) Quantitatively analyse SCOR WG 51 recommendations for processing CTD data;
The SCOR WG 51 recommendations for processing CTD data were published 6 years ago (UNESCO Technical Papers in Marine Science No. 54) and stem from discussions held several years earlier. The WG wish to assess these to ensure that they meet the needs of today, and suggest improvements if appropriate.

f) Assess the results of the intercomparison of quality assurance methods for station data;
Members of the WG will take part in an intercomparison, using a data set of 810 stations, to check that certain minimum standards are being met by the quality assessment procedures currently in place in ICES member countries.

g) Report on the development of an umbrella for gopher (on Internet);
This could be a useful way of advertising ICES and the data centres in the ICES community, and now that the technology is available the WG need to investigate gopher and exploit the system to the best advantage.

h) Report on the work of the IOC/IODE Group of Experts on the Technical Aspects of Data Exchange (GE/TADE);
GE/TADE meets during the year and will report on items of interest and value to the WG. Among the items to be discussed are new formats and data dictionaries.

ii) Time and place of next meeting

The WG expressed its wish that the next meeting should be held at the Irish Marine Data Centre, Dublin, between 1 - 3 May 1995. This follows on from the Oceanic Hydrography WG the previous week to be held in Oban, Scotland, allowing continued cooperation and interchange of ideas between the two working groups.

The Chairman closed the meeting by thanking the participants for their hard work, enthusiasm and valuable contributions. On behalf of the WG, she also thanked H. Loeng for an efficiently arranged and enjoyable meeting.

Annex 1

Agenda

C.Res. 1993/2:14

The Working Group on Marine Data Management (Chairman: Dr. L.J. Rickards, UK) will meet in Bergen, Norway, from 21-23 April 1994 to:

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 Member Countries;

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

Annex 2

Highlights from the reports of the Data Centres

ICES: The ICES data bank will reach its 100th anniversary in eight years' time and some thought needs to be given to products which could be developed to mark this occasion. There are still problems in obtaining historical data collected by Discovery, but progress has been made in obtaining data sets collected by IfM, Kiel. Recent data have been flowing in at quite a fast rate. ICES 'Punch Card' format is still in use although it is going through a stage of 'upwards evolution'; consideration will soon be given to a more flexible way of storing the data. Quality of salinity data received by ICES is now good, but the same cannot be said for nutrients. Cooperation is continuing with a variety of non-ICES projects, in particular those relating to the EC MAST programme and the IOC Global Ocean Observing System (GOOS).

Canada: In Canada the Marine Environmental Data Service (MEDS) is the depository of all physical and chemical data. MEDS gathers Cruise Summary Reports (ROSCOPs) or cruise reports from individual institutions and forwards them to WDC(A) and ICES annually. MEDS exchanges data with WDC(A). MEDS is involved with Data Archaeology and is collaborating with Syd Levitus at WDC(A). As part of a Green Plan contaminants, nutrients and other chemical data are to be collated into a database.

Denmark: Water bottle and CTD data from several sources, collected mainly between 1992 and 1993, are almost ready to be submitted to ICES. Time series of sea level and salinity/temperature are also collected. Sea level data from Danish sites are checked weekly and those from Greenland monthly (and also annually). From 1994 onwards a Scanfish (SeaSoar) and bottom mounted ADCPs will be operated by the Royal Danish Administration of Navigation and Hydrography (RDANH). Some data are transmitted back in near real-time and stored in a relational database.

Finland: A database for hydrographical, chemical and biological data has been built. The hydrographical and chemical data have been checked and passed to ICES. A separate CTD database is under construction using Paradox. 1991 and 1992 data have been added to the database and ICES have been of great help during the checking. ROSCOP forms were filled in until the late eighties. Now the computer system saves the information needed into numerous small files. The plan is to make a routine which automatically fills the ROSCOP during cruises. The Finnish Institute of Marine Research has collected chlorophyll-a data since 1990 on several ferry routes.

France: Work has progressed on recovering the historical database for CTD, water bottle and current meter data, leading to the publication of catalogues of the archived data. The TOGA/WOCE data have been migrated and a user interface has been developed for loading data. A user interface has also been developed to facilitate access to the data files archived by SISMER using interactive menus, control of access rights and data delivery. This has been designed for different types of terminals.

Iceland: Approximately 500 CTD stations were taken during 1993, including water samples for chemical analysis. Historical data have been loaded into an Oracle database and work has been done improving software for extracting and inserting data. CTD data are now being inserted in a similar manner. Data submissions to ICES began again during the year. Current meter measurements are being continued in the NW shelf and Denmark Strait. Use of satellite drifters will also continue.

Ireland: The Irish Marine Data Centre is now one year old and much of the work over the last year has been on establishing the data centre. Even so, the Irish EDMED has been upgraded and updated and a digital bathymetric data set for the waters around Ireland is being compiled. Other projects include compilation of a detailed inventory of all of the data held by the marine section of the Irish Geological Survey, compilation of a database of scientific publications, management of data from the Lough Beltra and the completion of the EROS 2000 database.

Netherlands (National Institute for Coastal and Marine Management (RIKZ)): 1992 data from the national joint monitoring group (JMG) programme have been sent to ICES. The DONAR relational database system for the storage of all waterbound measurements of Rijkwaterstaat became operational. Users (500) of the system have been instructed. Application software development is now underway. Progress has also been made in the Rijkwaterstaat Monitoring Network Infrastructure. Contributions have been made to several EC MAST projects.

Netherlands (MARIS): MARIS is currently working on 4 projects; a Dutch funded project for the clean up of Gdansk and Vistula Bay in Poland, compiling an inventory of data collected during the Netherlands Indian Ocean expedition, participating in the European Seabed Resource Geographical Information Service, and a feasibility study for the compilation of a CD-ROM containing North Sea Tidal Data.

Norway: Over the last 3 years an integrated database has been developed to link together all kinds of marine data. After testing in early 1993, the database was revised and this will be ready by the end of April 1994. Procedures for quality assurance are now being developed, programs are being developed for data input and applications. During the last year data flow to ICES was rather low, but almost 100% of the data from 1990 was transferred to ICES. Submission of ROSCOPs is working well.

Portugal: All of the information relating to bathythermograph data since 1957 was identified. This amounts to 15000 stations. More recent data, from 1991 to 1993, has been processed and published with the identification of the upper layer. Data collected between 1988 and 1990 have been processed and sent to ICES. The moored current meter inventory has been updated and the information supplied to BODC.

Sweden: Water bottle, compressed CTD and biological data from Swedish research vessels, coastguard and icebreakers have been submitted to ICES (1990) and Helsinki Commission (BMP stations 1992). Work has continued on the data collected during the SKAGEX and GULF OF BOTHNIA YEAR projects. The database management system has been upgraded and the relational database reconstructed. Swedish marine data collecting institutes are engaged in an ambitious national programme for quality assurance.

U.K.(BODC): Highlights over the year include the production of CD-ROMs for the GEBCO digital atlas and BOFS (the UK contribution to JGOFS). In addition, work on establishing a data bank of CTD profiles was continued with funding from the UK Hydrographic Office. The main BODC database and accompanying software has been re-established on Unix workstations. Work has continued on the EC funded EDMED directory which should be completed in the coming year. BODC has a contract with the EC to carry out the data management for the Ocean Margin Exchange (OMEX) project.

U.K.(HO): The main task each year is to process BT data from the UK Royal Navy, scientists and ships of opportunity. Releasable data will soon be passed to BODC. During the year a CD-ROM has been produced containing all Navy MBT and XBT data collected in the North and South Atlantic between 1947 and 1990. This comprises approximately 115000 profiles and is available for bona fide research. In the coming year data from the remaining oceans will be checked out and a second edition of the CD-ROM will be produced.

U.K.(MAFF): MAFF's fieldwork over the year has concentrated on a study of nutrient fluxes in three UK estuaries; the data should be with ICES by the late summer. Samples were also collected and analysed for nutrients, supported by CTD profiles, as part of the WOCE Antarctic Deep Outflow Experiment (ADOX). ROSCOP submission to ICES is good and an effort is made to submit all physical and chemical data to ICES.

U.S.A.: The National Oceanographic Data Centre (NODC) and the co-located World Data Centre A for Oceanography (WDC(A)) have spent the last calender year (1993) actively pursuing new activities in data management, global data archaeology and rescue, and in producing new products and services. Over the year computer activities have been moved from VAX to Unix. Transfer of all physical oceanographic profile data to a Sybase data base management system is under development. In the summer new CD-ROMs containing data obtained through GODAR and a new Levitus climatology will be released. GODAR workshops have been organised and run in Russia and China. US ROSCOPs have been sent to ICES and updated versions of the NODC/WDC(A) archives have been sent to ICES.

Annex 3(a)

Report on the Global Oceanographic Data Archaeology and Rescue (GODAR) Project

A critical requirement for climate and global change research is the availability of digital oceanographic data covering long time spans. It has been estimated, however, that perhaps two-thirds of all historical oceanographic data exist only in manuscript form and have not been submitted to a national data centre. These data remain effectively unavailable to researchers.

To address this problem, the U.S. National Oceanographic Data Centre (NODC) and the co-located World Data Centre A for Oceanography (WDC(A)) have begun a project known as Global Oceanographic Data Archaeology and Rescue (GODAR). This project is being organized and conducted in cooperation with the Committee on International Oceanographic Data Exchange (IODE), a subsidiary body of UNESCO's Intergovernmental Oceanographic Commission (IOC).

History of NODAR and GODAR

The NODC Oceanographic Data Archaeology and Rescue (NODAR) project and IOC/IODE GODAR project had their origins at an international meeting held at NODC/WDC(A), Washington, D.C., in September 1990. As a result of the meeting various national and now international projects that are known generically as "Oceanographic Data Archaeology and Rescue" projects were initiated.

In 1992 a follow-up meeting known as the "Workshop on Ocean Climate Data" was held in Greenbelt, Maryland, USA. (Churgin, 1993). The progress of various national data archaeology and rescue projects prompted the workshop to recommend the expansion of these projects to global scope under the umbrella of an existing international organization. A proposal for a GODAR project was submitted to the Fourteenth Session of the IODE held in Paris, France, in December 1992. The IODE in turn recommended to the IOC that this project be adopted as an IOC project. At the March 1993 meeting of the IOC Assembly the proposal for a GODAR project was adopted.

The first GODAR workshop was held in Obninsk, Russia, in May 1993. This meeting focused on data sets and activities in eastern and northern Europe, a region where economic conditions place the preservation of many oceanographic data sets at risk. The magnitude of the problem is indicated by the amounts of data that exist in manuscript form. The Russian delegation reported the existence of approximately 450,000 mechanical bathythermograph (MBT) profiles and 800,000 oceanographic station (OSD) profiles in manuscript form. Ukraine has data from at least 100 cruises.

The second GODAR workshop was held in Tianjin, China, in March 1994. This meeting focused on data sets and activities in western Pacific rim countries. Like the first GODAR workshop, the purpose was to bring scientists, administrators and data managers from nations of a specific geographical region to focus on problems of historical oceanographic data preservation and access. The workshop helped lay the groundwork for a major upgrading and modernization of the ocean data management system for the entire region.

Future workshops are planned for October 1994 for the Indian Ocean region and May 1995 for the Mediterranean.

GODAR activities

The national archaeology and rescue projects at various centres have been coordinated to avoid duplication of effort and to maximize the use of scarce resources. Joint activities include the exchange of data, data distribution plots and catalogue information about data holdings. Exchange visits between scientists and data managers from different centres have also taken place.

To provide information about its own holdings of physical - chemical data, the NODC compiled an inventory publication containing data distribution plots and tables of the number of profiles by year for each major measurement type (Levitus and Gelfeld, 1992). This work showed the distribution of NODC holdings for all countries. NODC/WDC(A) has also prepared similar plots on a country-by-country basis and distributed these summaries to data centres, scientists and institutions in approximately 20 countries (as of December 1993). These summaries have generated much interest and resulted in the exchange of more information and data.

A new report on the initial results of GODAR has been published. Titled "Results of the NODC and IOC Oceanographic Data Archaeology and Rescue Projects: Report 1", this publication provides station location plots and tables of counts of observations by year for the 1.2 million ocean temperature profiles received during the last two years.

GODAR results

Numerous institutions worldwide are now participating in the GODAR project. The navies of several countries have been declassifying oceanographic data and making

these data available internationally. For example, the Russian navy is participating in the GODAR project by making manuscript data available for digitization and distribution. The U.S. Navy has recently declassified approximately 100,000 XBT profiles that are now part of the NODC/WDC(A) databases. In 1994 we expect to receive approximately 115,000 bathythermograph profiles that will be released for international distribution by the British Navy. Some of these data were made as early as 1945.

As a result of GODAR, more than 1.2 million profiles have been received by NODC/WDC(A) in the past two years (see accompanying table). In addition, the first GODAR workshop identified about one million additional profiles in manuscript form that remain to be submitted.

Future work

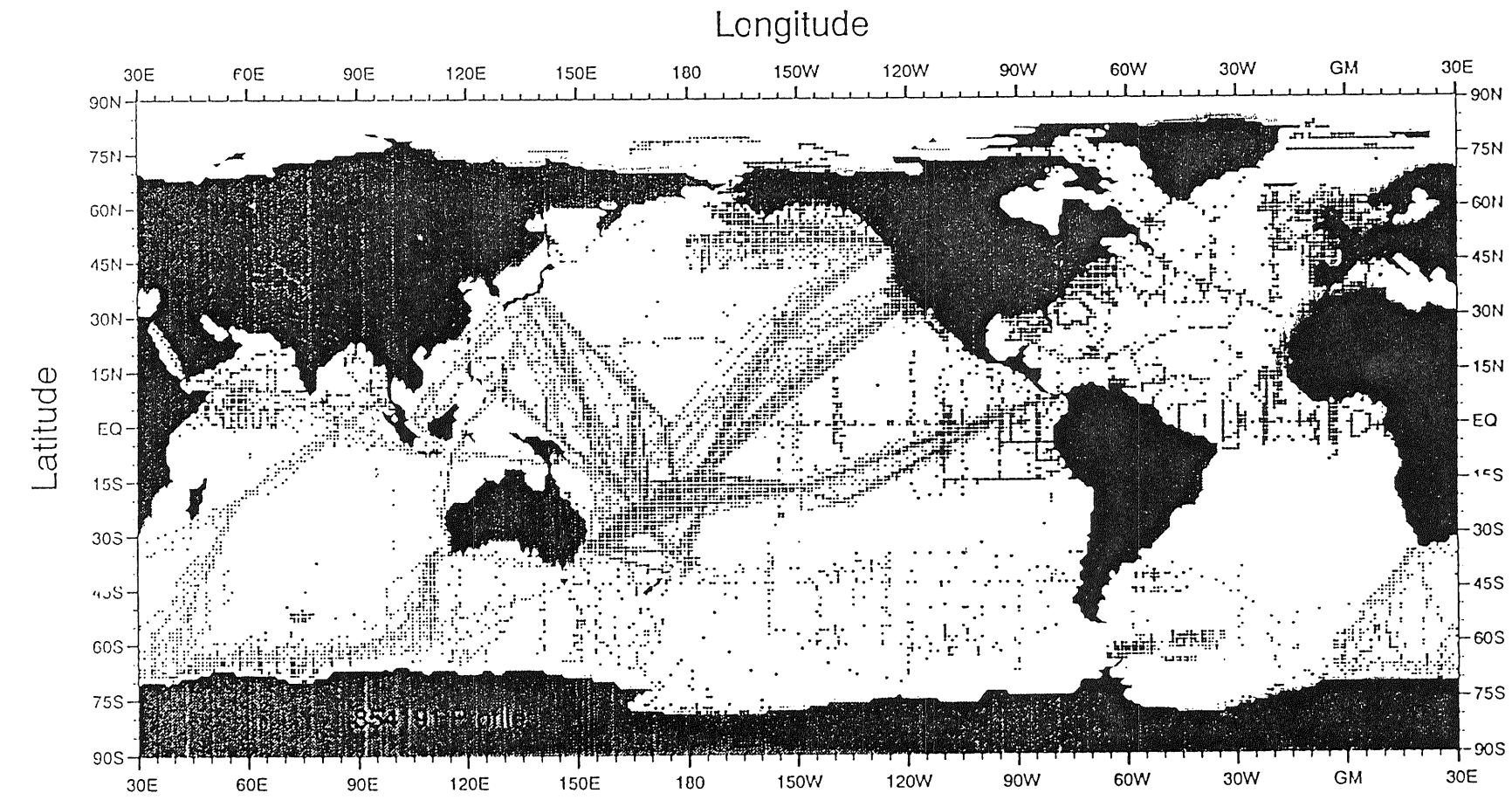
The data sets received so far are being processed at NODC/WDC(A) and will be distributed on CD-ROMs and magnetic media. While we expect the data and data analyses to be made available over an electronic network, CD-ROM technology best serves the international distribution of these data since some countries do not yet have the required network capability.

In addition, a set of objective analyses has been prepared from the NODC's augmented oceanographic database. Similar to analyses presented in the "Climatological Atlas of the World Ocean", published by S. Levitus in 1982, the new analyses will also be distributed both in digital form and as a multivolume "World Ocean Atlas, 1994."

We hope that other investigators, research groups and data centres from all countries with marine research and operational programs, will participate in and benefit from the Global Oceanographic Data Archaeology and Rescue project.

Table 1: Data sets received through NODC and IOC Data Archaeology and Rescue projects

COUNTRY INST	DATA TYPE	NO. OF PROFILES	PERIOD OF OBS.	REMARKS
AUSTRALIA				
CSIRO	OSD	22,190	1929-1990	
CANADA				
MEDS	XBT	46,658	1968-1988	
	DBT	11,563	1982-1981	
	MBT	145,286	1943-1988	Approximately 26,000 of these are "new" to the NODC/WDC(A) archive. The rest are replacing existing profiles because the "new" ones are digitized at observed levels rather than at 5 m intervals.
CHINA				
WDC(D)	OSD	8,053	1958-1990	
DENMARK	OSD		1970-1992	
FRANCE	MBT	2,791	1964-1972	MBT+Surface Salinity
GERMANY	OSD	20,521	1925-1929	Surface Salinity
ICELAND	OSD		1938-1988	
ICES	OSD		1971-1974	CINCECA
	OSD		1948-1990	OWS L
	OSD		1983-1990	OWS C
	OSD		1948-1988	OWS M
	OSD		1925-1930	ATLANTIC SLOPE
INDIA				
INODC	OSD	650	1976-1988	T, S, nutrients
JAPAN				
JODC	OSD	254,846	1965-1990	JAPAN FISH. AG.
	DBT	23,452	1979-1986	JAPAN FISH. AG.
	MBT	60,764	1965-1985	JAPAN FISH. AG.
	XBT	1,774	1979-1985	JAPAN FISH. AG.
	CURM	28,487	1964-1985	JAPAN FISH. AG.
RUSSIA				
POI	OSD	5,636	1947-1988	S. CHINA SEA
	CTD	4,249	1981-1989	S. CHINA SEA
WDC(B)	MBT	232,793	1940-1990	
	OSD	12,931	1969-1991	N.& S. Atlantic
SOUTH KOREA				
KODC	OSD	30,222	1961-1992	
UNITED STATES				
SIO	MBT	73,987	1943-1970	North Pacific
	OSD	162	1972-1973	TASADAY Legs I-IV, Southtow, Climax
US COAST GUARD	OSD	634	1961	Surface T&S
WHOI	MBT	242,264	1941-1961	U.S. Navy, Coast Guard and research ships
WDC(A)	OSD	1,169	1958-1978	Digitization of manuscript data held at WDC(A)
UNITED KINGDOM				
IOS	OSD	105	1973	R.R.S. Shackleton, Cruise 6 ICES Overflow 73 Expedition. These data were digitized at NODC/WDC(A) after receipt of manuscript tables from Mr. Jim Crease.



Distribution of Chlorophyll Data in the Ocean Climate Database

Annex 3(b)

Status of GODAR in ICES Member Countries

Canada: In Canada, individual Department of Fisheries and Oceans laboratories have set up data management committees, which meet up nationally on an annual basis. They have agreed to standardise on header information and keep the same information in each file. MEDS has copies of all of the data from the Bedford Institute of Oceanography, St. John's and the Institute of Ocean Sciences. The laboratory at Winnipeg has a lot of CTD which they are trying to get into the system at MEDS. There has been good collaboration between MEDS and WDC(A) especially over the last 6 months; MEDS has passed on the data it holds to WDC(A). There has been a problem with nutrient data, but there is now a programme underway to recover data. Now everyone is interested in climate they want access to old data.

Denmark: There is no data archaeology programme at the Royal Danish Administration for Navigation and Hydrography (RDANH), but the Danish data set at ICES is virtually complete. At present, the availability of 'old' short records of sea level data from locations in Greenland is being addressed. These were gathered during hydrographic surveys in order to establish chart datum. They may be of value in relation to the tide tables for Greenland, for which RDANH has responsibility. There is a large body of current meter data from the Belt project in the 1970s which will be incorporated in the RDANH data base.

Finland: Ice data, collected in the 1720s, are the best, oldest data in Finland. These have not been sent to any international archive. The hydrographic/nutrient data are in quite good order. Old hydrographical data (1898-1963) have been collected from publications and added to the database. ICES has provided a large amount of data from the beginning of the century. Comparing the ICES data set with that held by the Finnish Institute of Marine Research (FIMR) shows that FIMR has considerably less data than ICES, especially before 1970. FIMR has biological data mainly since 1969, macro fauna data from Tvarminne station since 1920. Sea level and wind generating wave measurements are in good order, but flow data and sediment data need collection and checking.

France: There used to be an active data centre in France, which then declined, but this was reestablished in 1991. The data have been reorganised into a new database; only some current meter data have not yet been put into the system. Data collected from 1986 onwards are scattered in the data collecting laboratories. There is cooperation with WDC(A), and the first data for some time have been forwarded to WDC(A).

Germany: Data from 75000 stations are being keyed in from the North Sea and the Baltic. This may also cover some Polish data. These data are not yet in the database, and it may be some time before this happens. The data can be made available, but the format may be inconvenient at present. MBTs have been loaded to the database.

Iceland: All hydrographic data collected by Iceland are at ICES/WDC(A), except for data collected in the last few years, and possibly some data collected close to the coast.

Ireland: The infrastructure for the Irish Marine Data Centre is now in place. One project so far has been compiling 20 years of data from the geological survey. Temperature and salinity data will be put together later. Data have also been collected by joint Russian-Irish cruises.

Netherlands: MARIS maintains an inventory of Dutch data, and Rijkwaterstaat are aware of the data held by their organisation. ICES is currently reviewing data held by WDC(A) and ICES and when this is completed a summary of the data held will be available. It is known that a lot of Indonesian data should be available. The Netherlands Institute for Sea Research (NIOZ) also hold much data, but only one scientist passes his data onto the appropriate data centres.

Norway: Very little work has been done over the last year, but some manuscript MBT and thermograph data is available. A list of missing data from the Institute of Marine Research has been obtained from ICES. Data are not submitted to ICES from laboratories other than the Institute of Marine Research in Norway.

Poland: Some tapes are held, the contents of which are unknown. There are also some old German data.

Portugal: The Instituto Hidrografico (IH) has compiled an inventory of water bottle data and holds MBT/XBT data from IH and the fisheries institute. The Instituto Hidrografico have no biological data. Portuguese data pre-1968 is fairly complete. Data from the University of Lisbon is not available at present.

Sweden: Quite a lot of work has been done so far. Data collected from 14 light vessels between the 1880s and 1970s are being digitised. So far, data from 1923 onwards have been digitised and will be quality controlled soon. Historical data from ships conducting coastal surveys have been identified and will be digitised. The Navy hold temperature and salinity data,

but the positional information is not included, the data are only known by the square in which they were collected. Other data of relevance to data archaeology include photographs of MBT data, which are not very accurate (only to ~0.5°C). Some contact has been made with the universities and some disk and manuscript data has been received, copied and stored. Commercial ships collecting sea surface temperature data send this to the Swedish Meteorological and Hydrographic Institute to produce maps.

U.K. (BODC): The main problem has been a lack of resources for this work (i.e. no funding). However the contract with the UK Hydrographic Office will ensure that most CTD data collected over the last 30 years will not be lost. Rescuing these data might in turn uncover other older data sets. The Institute of Oceanographic Sciences at Wormley is due to move to Southampton in the near future and, before this takes place, checks would be made to ensure any historical data which is not safely archived is not lost.

U.K. (HO): The HO has been checking that data have been forwarded to the appropriate centres. In addition, previously unreleased data have been made available to the HO. Whether these can be released to the scientific community is under discussion. The HO supplies scientists with XBT probes; if they do not submit the data to the HO, they are unlikely to receive further supplies of instruments.

U.K. (MAFF): MAFF hold some light vessel data and merchant ship/ferry data from the 1960s onwards which are now being quality controlled. This work is about 80% complete. They are also updating their files with water sample data from research vessel cruises in collaboration with BODC. This apart there is little emphasis on data archaeology at MAFF; scientists are more interested in recent data. However during the mid-eighties there was a need to find all of the nutrient data that had been accumulated by MAFF since 1960. A considerable amount of data was located and subsequently forwarded to ICES.

U.S.A.: Data have been received from the US Navy from Arctic Icebreakers (25000 CTD and bottle stations). Submarine XBTs should be made available in the future. Work is currently underway in the major universities of Duke and Miami. Chlorophyll and nutrient data are being chased up. In the near future CTD data from 25-50 cruises carried out by Woods Hole Oceanographic Institution will be acquired. 250000 MBT profiles, where the ships have been marked as unknown, have been checked and codes added in.

Annex 4

Proposed Guidelines for the Management of Shipborne ADCP data
(compiled by MDM WG 1992, amended 1993)

It is felt premature to store all ADCP data in data centres. Each institute should store their own data and send an inventory of these data with an estimate of their quality to the appropriate data centre. Data can then be submitted to the data centre on request.

All parameters collected should be stored (i.e. two horizontal components of velocity, vertical velocity, error velocity, per cent good and automatic gain control (AGC)). Navigation should also be submitted, either merged with the data or as a separate file which can be linked to the ADCP data file using date and time.

The data should be fully checked for quality and pre-edited or flagged for erroneous values such as spikes, gaps, etc. An explicit statement should be made of the checks and edits made to the data.

Sufficient self-explanatory series header 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 their original collection, processing and quality control. These are described in more detail below.

Data can be exchanged in the IOC standard format GF3, but it is equally acceptable to exchange data as ASCII files provided a format description accompanies the data.

1. Series header information should include the following:

1.1 Project, platform (i.e. ship), cruise identifier

1.2 Country, organisation

1.3 Date, time, latitude, longitude, (error of GPS), water depth for each profile

2. Accompanying information

2.1 Details of the instrument (eg manufacturer, model, instrument configuration (depth, location on hull), and any modifications carried out)

2.2 Data collection: description of operational procedures including time interval over which ensemble averages are performed, bin size, number of bins, bottom tracking on/off, pitch and roll on/off, methods of position fixing

2.3 Data calibration, quality and processing: brief description of procedures including

a) whether horizontal components of velocity are N and E components or are relative to the ship

b) criteria used for flagging or rejecting data (eg threshold values of error velocity and/or per cent good)

c) method used for correction of profiles for ship speed (i.e. bottom tracking, navigation or reference level of no motion)

d) calibrations carried out to correct for transducer misalignment

e) problems of contamination of the data due to bubbles in rough weather, high ship speed (propeller noise), change in ship direction or ship speed zero and how dealt with

f) whether corrections have been made to account for variances in PC clocks

g) whether sound speed corrections have been made, to improve the data, since the instrument is calibrated to standard values of speed of sound in the water column, and what these correction are

h) estimate of the final accuracy of the data

2.4 Applied units in which the data are expressed, should be clearly stated.

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

Proposed Guidelines for the Management of SeaSoar (batfish) Data

(compiled by MDM WG 1992)

Data should be stored as time series (of 1 second averages) and should be merged with navigation data. If the navigation has not been merged, they should be submitted as a separate file which could be linked to the SeaSoar data using date and time. If data are not available in this form, then data split into 'pseudo-CTD' casts are acceptable.

All relevant calibrations should be applied to the data including instrument calibrations and field corrections. The data should be fully checked for quality and pre-edited or 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 series header 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 their original collection, processing and quality control. These are described in more detail below.

Data can be exchanged in the IOC standard format GF3, but it is equally acceptable to exchange data as ASCII files provided a format description accompanies the data.

1. Series header information should include the following:

1.1 Project, platform (i.e. ship), cruise identifier

1.2 Country, organisation

1.3 Date and time of the start and end of the SeaSoar run

1.4 For data supplied as 'pseudo-CTD' casts date, time, latitude, longitude, and an up/down cast indicator for each cast

2. Accompanying information

2.1 Details of the instrument (eg manufacturer, model number and any modifications carried out)

2.2 Data collection: description of operational procedures including sampling rate, sensor resolutions, undulation rate, methods of position fixing

2.3 Data calibration and quality

a) types of sensors

b) laboratory calibrations (eg whether carried out in accordance with SCOR Working Group 54 recommendations)

c) *in situ* calibrations (eg lowering the CTD before and after a SeaSoar as a conventional CTD, use of thermosalinograph, or water samples taken from the non-toxic supply)

d) report on corrections made to data especially for offsets in salinity due to fouling of the conductivity cell

e) estimate of final uncertainty in the data

2.4 Data processing: brief description of procedures including

a) filtering/de-spiking/smoothing methods

b) editing/quality control methods

c) time lag correction scheme

d) adjustments made due to variations in calibration

2.5 Data should be expressed in oceanographic units, which should be clearly stated.

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

Annex 5

Recommendations

Proposed Agenda for next year's meeting

a) Assess the post-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 (GODAR) Project in each ICES member state;

c) Report on procedures for processing and storage of shipborne ADCP data;

d) Critically analyse data processing procedures for moored current meter data in ICES member countries;

e) Quantitatively analyse SCOR WG 51 recommendations for processing CTD data;

f) Assess the results of the intercomparison of quality assurance methods for station data;

g) Report on the development of an umbrella for gopher (on Internet);

h) Report on the work of the IOC/IODE Group of Experts on the Technical Aspects of Data Exchange (GE/TADE).