Oceanography Committee

ICES CM 1998/C:7

REPORT OF THE

WORKING GROUP ON MARINE DATA MANAGEMENT

Brest, France
20–23 April 1998

This report is not to be quoted without prior consultation with the General Secretary. The document is a report of an expert group under the auspices of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council.

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Palægade 2–4 DK-1261 Copenhagen K Denmark

Table of Contents

1. Opening of the meeting

The meeting was opened at 9:00am on 20 April 1998, hosted by the SISMER/IFREMER, Centre de Brest, France. Participants were welcomed to the meeting by the WG Chairman. M. Gerald Riou, Director of Computers, Network and Data Management Department (IDT), welcomed the Working Group to IFREMER and provided a comprehensive overview of IFREMER and the IDT department. Dr. Catherine Maillard, Head of SISMER, also welcomed the Working Group to SISMER and provided a presentation on the activities of SISMER. M. Fichaut also welcomed participants and explained the local arrangements.

Members of the Working Group present were: S. Almeida, Portugal, M. Fichaut, France, M.J. Garcia, Spain, R. Gelfeld, USA, J. Gagnon, Canada, D. Hartley, UK, A. Isenor, Canada, N. Kaaijk, the Netherlands, H. Loeng, Norway, F. Nast, Germany, O. Ni Cheileachair, Ireland, R. Olsonen, Finland, L. Rickards, UK (Chairman), H. Sagen, Norway and J. Szaron, Sweden. ICES was not represented due to budget restrictions. Apologies for absence were received from S. Feistel, Germany, K. Medler, UK, P.B. Nielsen, Denmark, G. Slesser, UK and H. Valdimarsson, Iceland. G. Riou, C. Maillard and M. Pitel, from IFREMER, attended parts of the meeting. A complete list of names and addresses and contact points of participants can be found in Annex 1.

2. Adoption of the Agenda

The agenda for the WG meeting was adopted as a resolution of the Annual Science Meeting in Baltimore, U.S.A. (C.Res. 1996/2:21, Annex 2).

3. Data Centre reports

The WGMDM 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 3 and the reports were distributed to WG members. Those reports received prior to the meeting were made available on the MDM Web pages; the remaining reports were added to the Web pages after the meeting. These can be found at:

http://www.pol.ac.uk/bodc/mdm/dcreports.html.

4. Assess the post-1990 oceanographic data sent to ICES by each member country, identify problems and suggest solutions

The Working Group has reviewed data flow to the ICES Oceanographic Data Bank annually over the last few years with a view to assessing the problems and improving the data submission. A brief report had been received from the ICES Oceanographer relating to the status of data submission. Over 55000 profiles had been received during last year (Annex 4) and for the first time more than 20000 profiles were held for 2 individual years (1988 and 1989). Recent data submissions had been received from Finland and France; these had not yet been added to the database. But low submissions were still a problem from Germany, Ireland, Spain, Portugal, Norway (nutrients) and the UK (NERC). However, the profile and surface data sets from the OMEX project, supplied on CD-ROM, have been merged into the ICES databank. In addition, the JGOFS parameter code table (available on the OMEX CD-ROM) has been used to expand the ICES format to cater for an increased number of parameters. The situation did seem to have improved somewhat over the past few years. The figures below indicate the number of profiles received at ICES, by year.

Year	Number of Profiles
1993/1994	14184
1994/1995	16000+

1995/1996	17627
1996/1997	51000+
1997/1998	55000+

L. Rickards reviewed the situation over the past five years since the WGMDM first investigated this problem. Various comments had been made including: 'North Atlantic data submission poor', 'Major gaps in the German data set', 'target of getting up to date by the ICES centenary (2002)', 'problems in obtaining JGOFS, WOCE and nutrient data', 'data policy working, but still some problem areas'. In 1993, L. Rickards submitted a paper to the ICES Annual Science Conference describing the status of Cruise Summary Report (ROSCOP) submission and data flow to the ICES Oceanographic Data Centre. After some discussion, the WG agreed that this should be updated and widely circulated. It would be included on the MDM Web pages, but would also be available for newsletters that the WG knew about. It was also suggested that WG members should use their Web pages to point to the maps available on the ICES Oceanography pages showing the geographic distribution of available data. This can be found at: http://www.ices.dk/ocean/maps/maps.htm.

C. Maillard felt that Cruise Summary Reports (CSRs) were important - they had been used for a long time in France, and were a valuable management tool for keeping track of 'who has been collecting what where'. J. Szaron agreed, giving some examples of where CSRs had been useful in tracking down data. R. Gelfeld also backed this up by noting that the CSRs had been useful for the WDC-A Ocean Climate Laboratory (OCL) for searching for nutrients and biological data. C. Maillard further noted that the SISMER Web statistics showed that the cruise information is the most frequently consulted, and that they are now starting to link this to the database.

Various countries (e.g. France, Germany) have their own CSR-like systems, which will dump out the information needed to send on to ICES. N. Kaaijk commented that the EU MAST EURONODIM project, in effect a follow-on from the MAST Data Committee, was intending to produce an on-line searchable system for CSRs; this was to be done by DOD.

O. Ni Cheileachair asked if data sent to ICES needed to be submitted in a particular format, or media, and was pleased to hear that data could be supplied on CD-ROM, in any properly documented ASCII format. M.J. Garcia wished to know if data submitted to ICES were public. The ICES data policy is that if data less than 10 years old are requested by an enquirer, the data originator is contacted to authorise release of the data. If data products (i.e. gridded data sets or statistics) are generated, then all available data are included. H. Loeng noted that he had agreed with the ICES Oceanographic Data Centre that all Norwegian data over 2 years old were public and available without restriction.

With regard to data submission, J. Gagnon noted that oceanographic data for the Northwest Atlantic were submitted to ICES as they are processed and updated into the MEDS archive, but that Cruise Summary Reports were not. H. Loeng said that the Norwegian nutrient data would be forwarded to ICES when his institute has accreditation for their quality assurance procedures. They are also working through the backlog of data. O. Ni Cheileachair said that temperature and salinity data would be sent to ICES once their new system was on-line. F. Nast commented that it takes time to increase the service available, but quite good progress was being made with the German scientists. M.J. Garcia felt that the situation was improving in Spain and S. Almeida promised some Portuguese data by the next MDM meeting. L. Rickards noted that the problem of the UK NERC data was almost entirely related to the lack of resources and other activities taking priority, however this may be remedied soon.

To summarise, the number of profiles submitted to ICES has increased over the last two years; this is a good sign, but there is still a large amount of data not being submitted to ICES. The WG agreed that the information in the 1993 paper should be updated and widely circulated. The ICES Oceanographer was requested to provide some input, in particular, about where things are going wrong. The WG felt that since the ICES Oceanographic Data Bank is such a valuable resource, the topic of data flow should be considered again in the coming year, with the emphasis on data collected in the last five years.

5. **Review progress in the implementation of IOC's Global Oceanographic Data Archaeology and Rescue (GODAR) Project in each member country, including consideration of biological oceanographic data types**

R. Gelfeld introduced this item by saying that the updated version of the World Ocean Atlas, known as the World Ocean Database 1998, produced by the Ocean Climate Laboratory (OCL) at WDC-A, was now available. It comprises almost 5.5 million profiles. Annex 5 shows the number of profiles for the different types of measurement (OSD, CTD, XBT, MBT, Tao buoys), and the geographic coverage. The GODAR project has led to the rescue of 190000 CTDs, 1.5 million bottle stations and 21000 profiles of biological data (zooplankton, phytoplankton, bacteria and some icthyoplankton). The biological data may include counts, biomass and volume. As this phase of the GODAR project is now coming to an end, an international GODAR conference is planned for October/November 1998 to discuss the direction the project should now take, and a steering committee has been set up to plan the meeting and decide who to invite.

During the course of GODAR, the WDC-A archive was compared with the ICES archive to remove duplicates. In addition, ICES has been a major force in getting GODAR off the ground. And ICES also acts as a backup for the World Ocean Database. There is a need for long term secure archives: ICES and WDC-A both perform this function.

Two years ago, the Marine Laboratory, Aberdeen, UK, shipped out to the WDC-A data books containing 18000 profiles which had not been digitised. In addition to temperature profiles, these also included meteorological data in the headers. Data from the MEDATLAS project have been received, but are not yet included in the World Ocean Database. R. Gelfeld asked about the joint Russian/Irish data collected to the west of Ireland - and agreed to work with the Irish Marine Data Centre to obtain these data. Most work on GODAR has been at an international level so far, rather than concentrating on data from the USA.

H. Loeng noticed that much of the Norwegian data included as station data (water bottles) were, in fact, CTD data supplied as reduced standard level data. N. Kaaijk asked what the status of Dutch data was. R. Gelfeld offered to send him an inventory of the cruises held, and after some discussion, agreed that it would be beneficial if all members of the WG received such a list, as this would enable them to check what data were missing and forward them to the WDC-A. Several members of the WG also requested summaries of their data held at ICES.

Funding for a follow-on to GODAR may come from climate change programmes, where data are needed for input to models, for prediction, and for sustained healthy coasts work. The more data recovered the better as far as the modellers were concerned. J. Gagnon backed this up, adding that data archaeology was one of the fundamental functions of data centres, where secondary users of the data are of prime importance. Data archaeology was especially useful to climate change work - for example, in the new Canadian Atlantic Zone Monitoring programme, historical data is required, which makes data archaeology a justifiable activity. It was also a necessary activity as the data had cost billions of dollars to collect in the first place, and would cost even more now.

C. Maillard noted that units and standardisation were a problem. Scientists, for example, may deliver data in a variety of units and not provide the extra information needed to convert between them. In addition, coastal and monitoring data may use widely differing protocols.

An exchange of data had taken place between MEDS, Canada, and the OCL at WDC-A to check that their archives agreed. C. Maillard commended this: a similar exercise had been carried out with SISMER, which revealed that OCL/WDC-A held French data not held at SISMER.

The emphasis is now moving towards nutrients, chlorophyll and biological parameters, although the best way of handling some of these data types has not yet been completely resolved. S. Almeida noted that for biological data, it is often difficult to identify exactly what is there. Header information, units and other qualifying information is needed more that ever. Mention was also made of contaminant data - these are useful for investigating trends.

The WG felt that this first five year phase of GODAR had been most important, and had uncovered a lot of non-digital (mainly temperature and salinity) data not previously available to the community. They commended the work of the OCL. The WG looked forward with interest to the outcome of the planned GODAR conference later in the year and wished to contribute to the next phase of the project. With this in mind, it was agreed that this should be considered at the next MDM meeting, when R. Gelfeld would update the WG on progress. Plans could then be developed for maximum contributions to the next phase of the project, which could well concentrate on biological data. The WGMDM would continue over the year to investigate and search out biological data sets.

6. Quantitatively analyse the minimum requirements for quality assurance of oceanographic data

Mr. Stig Carlberg, Chairman of ACME, had requested that the WGMDM consider one of the tasks for the Marine Chemistry Working Group (MCWG):

C.Res 1997 2:12p 'Advise on the need to standardise nutrient and oxygen units to μmol/kg.'

A. Isenor gave an outline of why the change had been made by the WOCE community from a volume to a mass unit. In summary, this is because with a volume measurement, one cannot compare deep ocean values with those made at the surface, as pressure influences volume. The difference is similar to that between temperature and potential temperature. In simple terms, with a volume measurement, one cannot tell how many molecules are being dealt with.

Some discussion followed. The basic view was that data centres are not in a position to dictate to a scientist what unit to use, but that the data centre needed to understand precisely what had been measured, and what had subsequently happened with the measurements. There was some agreement that one should always keep the 'measured' value, rather that those which have been calculated. Others thought that the chemists should decide what measurements (mass or volume) should be made, and that the data centres should store what they are sent.

L. Rickards noted that in the BODC database data are stored as volume. In fact, almost all of the data are received in this way. For those which are not, part of the dialogue with the data supplier is to find out how the conversion has been done, and then convert back to volume. A conversion factor is stored in the database for the convenience of those who wish to receive the data in mass units. The reason a factor is stored rather than a second set of parallel units is to keep the parameter coding under control. In principle any water column could be required in both units and therefore would need two codes. Finding the right code is enough of a problem with the present number of codes without doubling the problem.

She also described two problems that BODC had encountered. Firstly, BODC received some continuous underway nutrient measurements (4 channels, every 30 seconds). On one occasion the thermosalinograph stopped working, but the autoanalyser functioned correctly. So what should be done? Throw away 2 days of 30 second measurements of nutrients? Make a best guess of temperature and salinity and convert nutrients to per kilogram? Have some of the nutrients per litre and some per kilogram? Store the per litre data and have a conversion to per kilogram available so that users can either have the data per litre or can select any of the above if they wish? BODC chose the last option.

Secondly, some dissolved oxygen data were received by BODC with the units quoted as μmol per kilogram. Saturations calculated from these by BODC looked wrong and subsequent investigation revealed that the data were labelled as per kilogram because it was 'trendy', and that the data were in fact per litre at *in situ* temperature and salinity.

Converting using the value of 1.025 did not seem a sensible option. If the appropriate information is not available to perform the conversion accurately, then the scientist requesting the data should be informed exactly what is available, and can then make decisions about whether the data are useful to them based on this.

The WGMDM went through the conclusions reached by the MCWG and their comments on each of these is noted below.

After considerable discussion the MCWG agreed that:

- *It is essential that laboratories be allowed to report their data to the ICES Oceanographic Data Centre either on volume basis or mass depending on their normal practice and/or the requirements of special programmes (e.g. WOCE or JGOFS) they may be participating in,*

MDMWG agree with this, but stressed that the units should be clearly stated.

- *It is also essential that metadata (supporting information) is reported so that conversion from volume basis to mass basis is possible,*

Yes, the information required to convert from mass to volume and vice versa is required, or a conversion factor.

- *This reporting should be supported by the data reporting format (amended as might be needed),*

If this means that the data format description should accurately describe the format used, then this is fine, but format is not a word that we would recommend using if it can be avoided. So that when data are submitted they are accompanied by an accurate description of how the data are stored in the file, and all the relevant accompanying qualifying information is also submitted. The way in which the data are actually stored at the data centre should not be dictated by the data collecting scientists, but will be done to suit the data centre.

- *It is essential that data are stored in the data centre in their original form (either volume basis or mass basis) so that the integrity of the original data is not compromised,*

Ideally, this could be done, and all data stored as they are received (not on the original media, but maintaining the integrity of the original information). But this leads to 'holes' in the data if you cannot convert. It is also possible that when data are extracted for a secondary user, they will assume that the data supplied will all be in the same unit.

- *Any conversion of data is performed either by the data user or by the data centre on a direct and specific request by the user,*

The WGMDM had some problems in deciding what was really being said here. But we felt that the data centre should have all of the information to hand, and should be able to provide all parameters and conversions. Any conversion should be clearly documented, so that if a single conversion factor has been assumed, then it is obvious that this has been done. As much metadata as possible must be supplied with the data sets and these must also be maintained by the data centre.

- *While converting the data, the user should be responsible to ascertain that the original as well as the converted data have/will have the quality needed for the particular purpose for which the conversion is performed.*

The WGMDM were concerned to ensure that anything done to the data is documented, so that the user knows precisely what is being supplied.

J. Szaron noted that Mikael Krysall was to contact the WGMDM with regard to quality assurance for nutrients and oxygen, He volunteered to follow this up, and in the coming year the two WGs intend to collaborate over this.

7. **Report on the development of World Wide Web pages and links between them within member countries**

Several WG members demonstrated some of the developments at their Web sites. These included:

- IMR. Fixed station data, these are updates 2-4 times per month (http://www.imr.no/mil/fhs/coast/top.html)

- IMR. TASC pages for data management. This uses the US JGOFS/GLOBEC software. The data are available, but are in a secure area. (http://tasc.imr.no/tasc/datamanagement.html)

- REMSSBOT (Regional Environmental Management Support System Based On Telematics). This shares environmental information, not by building a central data warehouse, but by keeping the data at its original location. At present there is a demonstrator available for the Schelt river estuary (Netherlands). More details can be found at: http://www.hellas.eu.net/remssbot/.

- SISMER pages, including those for the MATER project (http://www.ifremer.fr/sismer/)

- NODC/WDCA pages (http://www.nodc.noaa.gov/)

After the demonstrations, the WG reviewed the MDM pages (http://www.pol.ac.uk/bodc/mdmwg.html). L. Rickards noted that these are not official ICES pages, they are maintained by the WG, not the Secretariat, because the WG wishes to advertise its work, expertise and data holdings to as wide an audience as possible. The WGMDM pages include information about MDM the Terms of Reference, data centres within the ICES area and their data holdings, guidelines for handling various types of data and last year's WGMDM report. In addition, all of the data centre reports received before the WG meeting were made available on the Web pages. Those received at the meeting would be added later.

A. Isenor thought that the guidelines were very useful, but would be better if they all adhered to the same format as there were variations between them. L. Rickards agreed to look at them again and rationalise them. After some discussion, it was suggested that a 'What's New' or 'New Products' section would be very useful. All WG members were to contribute to this by sending appropriate URLs to L. Rickards, who was currently responsible for maintaining the MDM Web pages. A. Isenor and J. Gagnon offered some help in maintaining the Web pages if required. L. Rickards also requested updates for the data centres within ICES pages. WG members can check their current entries at http://www.pol.ac.uk/bodc/mdm/dcindex.html.

The visibility of the page also needs to be raised. The WG felt that some parts of it could link in to the ICES Oceanography pages (as well as the present link through the Committee and Working Group pages). WG members should ensure that they have a link from their own home pages and a link with IOC will be investigated. It is likely that this will be to the GE-TADE pages when they are available.

8. **Instigate an analysis of the parameter code list used for the IOC Cruise Summary Report, and produce an improved and updated set of codes**

At last year's WGMDM meeting, it was agreed that there were many problems with the parameter codes on the present Cruise Summary Report (CSR or ROSCOP) form. The most pressing problems are the lack of codes for underway data (with the exception of temperature and salinity) and the difficulties posed by codes such as 'cores' which occur in the geology section, although cores are also taken by biologists. In addition, it might be better to separate shipboard ADCP from moored ADCP measurements, and geophysical measurements made at the surface and at the sea floor. It is also necessary to include the difference between bottle samples taken for measuring dissolved oxygen and CTD oxygen measurements. Moreover, there are now many more chemical parameters being measured (e.g. CFCs, CCl_4, etc.) which need to be included. A further inconsistency is that, at present, nutrients are included separately, but freons are grouped together.

The WG decided that it would be most useful to combine the discussion on Cruise Summary Report codes together with the more general discussion on data dictionaries. This discussion can be found in Section 10.

9. Investigate the Data Services available from NODCs in member countries and suggest a scheme to improve cooperation between countries to provide an improved service to the community

L. Rickards introduced this item. She explained that she had sent out a questionnaire to the WGMDM a few weeks previously asking the following questions:

1. *How many requests for data, data products or information about data (i.e. inventories, catalogues) have you handled in 1997?*

2. *Summarise the sort of data/information requested (e.g. waves, currents, XBT, CTD, data sets on CD-ROM, catalogues of data holdings)*

3. *Where do the data requests come from?*
   Your own organisation?
   Other organisations in your country (Universities, government, commercial organisations)?
   Organisations abroad?

4. *Do you have standard products available (e.g. CD-ROMs, statistical or gridded products)? If, yes, what are these products?*

5. *Can you always respond positively to requests (i.e. do you have the data requested?) or are you asked for data you do not hold?*

6. *If you do not hold the data requested, what is your response? Can you refer the enquirer elsewhere? And if so, where do you usually refer them to?*

7. *How do you think your service could be improved?*

If you are someone who requests data from NODCs (or from ICES), it would be very useful to have your comments on how easy (or difficult) it is to obtain data from NODCs or ICES. Do they provide the service you would like? And how would you like to see the service improved? Please also add any other comments that you have.

The response to this had been very encouraging with 14 responses. These are summarised in Annex 6. The US NODC handles many more requests than any of the other centres, but most other data centres handle about 150 or more requests a year.

Most centres answer a wide range of requests from their own institute, country and abroad. These requests can generally be answered by the data centre, but some need to be referred elsewhere. So it is important to know the appropriate organisations to refer enquirers to. Standard products (e.g. CD-ROMs, gridded data sets) were thought to be useful, as was on-line access to data. O. Ni Cheileachair commented that compiling Web sites which point to others holding data was useful for referral. R. Gelfeld noted that 'networking' (i.e. contact between data centres staff) was very valuable, and increased individuals knowledge of what was available at other centres. F. Nast reminded the Group that in the first instance, an enquirer should go to their national oceanographic data centre - which should have the knowledge and expertise to obtain data for them, if the required data were not held by the centre. Data sets might also be acquired without charge by one data centre from another as part of international data exchange agreements.

There was some discussion about data products, in particular whether it was clear where the data had come from. The WG agreed that it was most important to acknowledge all data sources. It is also beneficial to request feedback and reporting of any errors in the product.

Information supplied by H. Dooley, relating to requests to the ICES Oceanographic Data Centre, indicated that 90 requests had been answered during 1997. Of these, approximately half were for data, most of the remainder were for statistics or gridded products, and a few were for plots of station locations, information about data availability or inventory type information. The requests originated from 10 different countries, with 10 or more requests received from the UK (22), Denmark (14), Germany (14) and Norway (10). Finland, France, Spain, the Netherlands, Sweden and the USA had all between 1 and 3 requests each. The origin of a few of the requests could not be determined from the information provided. The WG were pleased that the ICES Oceanographic Data Bank was being utilised, but were surprised that there were not more requests to ICES and felt that the existence of the Data Bank should be widely and vigorously promoted.

Several actions were agreed as a result of the questionnaire and the subsequent discussion. On the WGMDM Web page, there should be a data products section with links to the relevant Web pages. This will contribute to answering the question of who has what data where. When a centre/laboratory has a new product available, in addition to alerting members of the WG, they should e-mail L. Rickards with the URL for inclusion on the Web pages. A map would be put on the Web showing the data centres and linking to their home pages. Also the international moored current meter inventory had been found very useful in the past, and a new version should be put on the Web. L. Rickards agreed to contact WG members and others who had supplied information for updates to the current meter inventory.

10. Investigate and evaluate the data dictionaries available to the marine science community

O. Ni Cheileachair provided an overview of some of the data dictionaries available to the marine science community. This included ROSCOP (Cruise Summary Report), EDMED, JGOFS (OMEX), MATER, GF3 and the Irish Marine Data Centre (IMDC) systems. A summary of her presentation is given in Annex 7. The main issues to come out of this review were as follows:

1. Hierarchical system important (facilitates searching and retrieval)
2. Confusion exists between instruments and parameters (especially in ROSCOP and EDMED codes) which needs to be resolved
3. Is it necessary to base a parameter coding systems on 8 byte codes?
4. It is necessary to indicate method and place (surface, mid-water, bottom)
5. Units are a problem

Leading on from this were 2 questions, together with some possible answers or suggested ways forward:

- What is critical in moving towards a better and more standardised system?

    1. Consistency between data centres
    2. Easy searching for multidisciplinary parameters
    3. Remove instrument from measurement
    4. Formal way of letting people know what's being updated

- Where next?

    1. Standardise 'big bucket' headings
    2. Agree on hierarchical structure
    3. Define parameters distinct from gear/instrument
    4. Agree/adopt a system

The WG thanked O. Ni Cheileachair for her excellent overview, which was followed by some lively discussion. R. Gelfeld commented that he had seen many data dictionaries over the last 25 years, and he felt that what was needed was an authoritative list, rather than a code table. He felt that codes had been useful in the past, but not in today's world. A standardised, authoritative list, which defines the parameters and their units, is what is needed. Although there was general agreement that this was true, there was also a view that code tables also had their place.

Code tables or data dictionaries are used because it is useful to have an abbreviated version of the parameter name, particularly in relational databases. But the main consideration is really that we all need to know that we are talking about the same parameter (i.e. identification and comparability). Existing internationally agreed standardised data dictionaries could be of value to those setting up a new database, as it saves work and avoids 're-inventing the wheel'.

One problem associated with either an authoritative list or a data dictionary is the question of maintenance. Someone has to take responsibility for updating and adding new codes. Expertise is required in a wide range of disciplines, and quite a lot of work could be involved if many people are requesting new codes. Often systems fall down because this activity has been underestimated and insufficiently resourced.

All WG members agreed that standardisation was required, and in defining the way forward O. Ni Cheileachair suggested that the following questions needed answering:

1. What truly defines a parameter?
2. Do we want an abbreviated way of defining parameters?
3. Do we want to standardise at the category ('big bucket') level?
4. What is the easiest way of doing this?

An intersessional sub-group was set up to consider this further. In particular, to suggest the 'big bucket' headings and suggest the appropriate hierarchical structure. It should also consider other coding systems, for example the BUFR coding system, used by meteorologists, which now has oceanographic codes included. In addition, it would be useful to consider the different sorts of data flagging schemes in use and suggest which to standardise on. The sub-group will consist of O. Ni Cheileachair, M. Fichaut, L. Rickards, J. Gagnon (together with Bob Keeley from MEDS) and H. Dooley, led by O. Ni Cheileachair.

11. Consider the future work programme in relation to the remit of the Oceanography Committee and development of the ICES Five-Year Plan, including cooperation with other Working Groups

At the last ICES Annual Science Conference in Baltimore, USA (September 1997), the Hydrography Committee was dissolved and a new Oceanography Committee formed. The remit of this committee is as follows:

'The Committee's scientific area of responsibility should be physical, chemical and pelagic biological oceanography, especially with regard to processes relevant to living marine resources and environmental quality. This will include such issues as impacts of climate variability, physical, chemical and biological fluxes in coastal areas, shelf seas and the open ocean.'

H. Loeng introduced this topic and provided the WG with background information on the new structure of ICES and the mid-term meeting of the Consultative Committee. He noted that there was a Bureau WG on the strategic policy and that the work of the WGs should be related to the Five-Year plan. He had written to members of the Oceanography Committee soliciting their opinions. Responses had included climate variability and effects, GLOBEC, GOOS, pollution and data management. Those who had included data management in their responses commented on the following:

'...The principal of these is to establish a coherent scheme of ecosystem modelling for ICES regions and a coherent policy on ecological data management.......At the present, ICES is pretty good at handling hydrographic and nutrient data, but that's about it. The fish survey data are a bit patchy, benthos data are ok for major surveys, and plankton data are non-existent. I think we need to get some major commitment for member institutes to get a coordinated systematic monitoring plan for various aspects of the health of the ecosystem. Unless we do this, then we will be struggling with patchy, messy horrible data in 10 years time, let alone 5. We need to press for an international data centre to take on archiving for a wider range of ecosystem data than currently catered for by ICES, BODC or elsewhere......'

'....there are continuing issues relating to data standards and data exchange. As we get more interested in the movement of various chemical tracers of human activity though the marine environment, we need to ensure that measurements of these substances can be mapped through space and time.'

'.... I would think that subjects related to......and (3) data management and exchange systems, would be of interest to many countries.'

'...The relevant topics for ICES to be monitored and promoted by the Oceanography Committee are.....(4) environmental data banking.'

The WG first considered the ICES Oceanographic Data Centre and its usefulness and its resources. In an ideal world, where all scientists worked up their data, stored it in an easily accessible manner and made it available to others in a standardised way, then national and international data centres may not be necessary. But in the real world this does not happen, as there are many other pressures on scientists, so it necessary to have national, regional and international data centres. The WG agreed that there were many good reasons for maintaining and further developing the ICES Oceanographic Data Centre. These are noted below:

- The ICES Oceanographic Data Bank has data from over 1.5 million profiles going back to the beginning of the century. The data are all quality controlled to a high standard. The data set comprises a valuable resource for many purposes including climate change and operational oceanography.

- A long-term archive is needed for data. Scientists retire and regional or international centres often have more long term stability than national centres. For example, the French centre, BNDO, was closed down and then several years later SISMER was established, and had to go to international data centres to re-acquire their data.

- Initiatives from the ICES Oceanographic Data Centre have been adopted by IOC.

- It provides a forum for developing guidelines for handling data, agreeing quality assurance procedures, etc. These are not restricted to the parameters currently stored in the ICES Oceanographic Data Bank. Over the past 10 years guidelines have been developed for moored current meter data, CTD data, XBT data, shipboard ADCP and SeaSoar. Some of these have been endorsed by IOC.

- ICES acts the National Oceanographic Data Centre for Iceland and Denmark. It also holds a back-up copy of the data from the WDC-A Ocean Climate Laboratory.

- The ICES Oceanographic Data Centre Data Policy is effective. Scientists will submit their data knowing that it is safe and will not be released without their permission within a 10 year period. In addition, scientists from some countries will send their data to an international centre, rather than a national one, especially if their national centre is not well resourced or developed.

- ICES expertise has been very valuable to projects such as MEDATLAS, where the ICES Oceanographer acted as an independent data expert.

- The ICES Oceanographic Data Bank can adapt to change, adding in new parameters as appropriate, as has recently happened with the adoption of the JGOFS data dictionary to allow the inclusion new parameters.

- It provides a valuable forum for discussion. The problems of one data centre are often the problems of others. Many lessons can be learnt and time saved.

The WG then turned its attention to whether an MDM Working Group was necessary and, if so, how it could contribute to the Oceanography Committee remit. The WGMDM is not a scientific or advisory WG, but data management activities should form an important part of any scientific programme which involves data, whether it be data collection, compiling data sets quality assurance, data products or final archiving. Within the WG there is an existing infrastructure for data management. A pilot project could be developed, building

on this, using perhaps an operational/monitoring approach for data types that are common to the data centres. This would provide a focus for activities. Whereas the WGMDM should not define the scientific programmes of the Oceanography Committee, it should ensure that data management is part of any programme. For example, MDM can contribute expertise in the areas of data exchange, formats, quality control, data products, data dissemination, and data archiving. The expertise of the WG is not confined to a particular data type; several WG members are involved in data management for multidisciplinary projects, which include many different parameters (e.g. physical, chemical, biological, fisheries, meteorology, geology/geophysics).
The WG decided that it would be valuable to have a general statement outlining its own function - distinct from the Terms of Reference which change from year to year. A. Isenor has put together a first draft of this. There are four parts to the Mission Statement as outlined below. Under each part he has some words that could be used to describe the Working Group, and from these words, the remit has been constructed. The order of the four parts is flexible. He prefers to have the purpose up-front, as this makes a stronger statement.

1. whose needs are we addressing (who we are)
    - ICES WGMDM
2. our uniqueness (what makes us unique)
    - we serve the ICES oceanographic community
    - we serve various ICES committees
3. our purpose (what we hope to achieve, or our outcome)
    - increase data and information exchange within ICES membership
    - advise ICES members and Committees , as appropriate, on data management issues
4. our function (how we will achieve this)
    - by monitoring data exchange/flow
    - by improving data exchange/flow
    - by being knowledgeable on current data management practices

Draft Remit for WGMDM:

'The ICES Working Group on Marine Data Management will maintain and develop expertise in oceanographic data management and will monitor, co-ordinate and improve data and information exchange within the ICES oceanographic community.'

The WGMDM links could be developed further with other WGs. There are quite good links with the Oceanic Hydrography WG and these two WGs had collaborated over the development of the guidelines for ADCP and SeaSoar data and have held several joint meetings. There is a need to build more links to the WGs on Zooplankton Ecology and Phytoplankton Ecology, especially with the increase in interest in biological data. It is also likely that links will be further developed with the Marine Chemistry WG, as the WGMDM collaborates with them in the development of quality control guidelines for nutrients and oxygen.

The WGMDM members are contributing data management expertise to a number of projects, nationally and internationally. At present these include the following: Global Temperature and Salinity Profile Project (GTSPP), WOCE, JGOFS, TASC, MEDATLAS, World Ocean Database 98, GODAR and a variety of EU MAST projects. Other projects where contributions are just beginning include GOOS, EuroGOOS and CLIVAR.

In addition, the WG has much expertise in designing and using database systems. Relational databases are in use at, for example, SISMER, IMR, RIKZ, BODC, IMDC, NODC/WDCA, and SMHI. This expertise is available to ICES.

12. Comment on the 1997 ACME statement (Agenda Item 21.3) concerning the development of GOOS initiatives in ICES

H. Loeng introduced this item and provided some background information. At the last ICES Annual Science Conference, an ICES Steering Group on the Global Ocean Observing System (GOOS) was established. Its term of reference is to: 'Prepare an action plan for how ICES should take an active and leading role in the further development of GOOS at a North Atlantic regional level with special emphasis on fisheries oceanography.' The chairmen of the Working Groups under the Oceanography, Marine Habitat and Living Resources Committees make up the Steering Group . Prior to coming up with an action plan, it was necessary to define more precisely the degree of ICES involvement in GOOS. The WG is asked to comment on the following four alternatives:

Alternative A: ICES is formally represented in all appropriate GOOS fora, such as the new GOOS Steering Committee, I-GOOS, the relevant GOOS Module Panels as well as in EuroGOOS. All the operational activities are organised by the member countries themselves and there is no regional GOOS system within the ICES area. This alternative is only slightly above the present involvement and may be characterised as 'Business as usual'.

Alternative B: An official GOOS Pilot Project has been established within the ICES area (e.g. North-east Atlantic, North Sea, the Baltic) by other bodies. In addition to what is mentioned under Alternative A, ICES have a role as an advisory and service agency for the regional GOOS component. Types of services could be:

* Databases and data management
* Quality assurance - methods, manuals, guidelines, inter-calibration exercises
* To support the Living Marine Resources Module, in particular concerning phytoplankton, zooplankton and benthos

Alternative C: ICES take the responsibility to run a centre for operational fisheries oceanography on non-meteorological time scales (i.e. more than two weeks) or on the time scale of fish stock assessment (some months) for the whole North Atlantic or parts thereof, i.e. the North Sea. The centre coordinate national and international data collection, the rapid transmission of data to computerised data assembly centres for processing through numerical and statistical models to produce regular:

* Climatic prediction (time scale season to some years)
* Regular environmental status reports
* Time series for identifying trends or changes

Alternative D: In addition to the tasks mentioned under Alternative C, we could also include processes of meteorological time scales, i.e. ICES establish a Centre for operational fisheries oceanography on time scales from days to years.

H. Loeng asked how the WGMDM could be involved and suggested it could have a role in real time data exchange, quality assurance, common data formats and products. The WGMDM also noted that there were many national committees for GOOS, and that R. Keeley (MEDS) and B. Searle (AODC) had written a paper showing how GOOS could use the existing IODE system for managing GOOS data, in particular the scheme used for the Global Temperature and Salinity Profile Project (GTSPP).

The initial reaction of the WG was that for options other than Alternative A, funding and personnel would be needed, which could be a problem. However, leaving that consideration aside, Alternative C was favoured by most members of the WG. J. Szaron noted that in Sweden, some of this type of work was already being carried out. Similarly, in Canada, work was just starting in this area, and J. Gagnon felt this was opportune. He recommended a pilot project first. He also noted that Canada was most interested in the western Atlantic.

O. Ni Cheileachair noted that EuroGOOS was accelerating, and setting up a data management scheme - how would an ICES GOOS regional project link with EuroGOOS? She also felt that there were definite advantages of regional data sets, where data have been pulled together over a large area.

As a data management group, it is not for us to suggest the scientific elements of the scheme, but data management should form a part of any project, and input could be provided on databases, quality assurance, assembling regional data sets, presentation of data (via the Web, for example) and production of products.

In summary, the WGMDM view is as follows:

1. To the question should ICES be involved, one response was – Can ICES afford to be left out? GOOS is a major project, and an ICES regional project would be very likely to contribute much in terms of knowledge, expertise and data.
2. Alternative C was the favoured option.
3. The WG could contribute a wide range of expertise in data management.

13. Election of chairman

L. Rickards reported that she had now been chairman for six years and that the WGMDM needed to nominate a new Chairman. At the last Annual Science Conference, it had been agreed that Chairmen of Working Groups should be appointed for three years. L. Rickards proposed that R. Gelfeld should be put forward to the Oceanography Committee as the next Chairman; H. Loeng seconded this, and this was agreed unanimously by the WG. L. Rickards thanked the WG for their support over the last six years and wished R. Gelfeld every success in chairing the WG in the future.

14. Any other business

(i) Shipboard Ocean Data Information (ODIN)
A. Isenor gave a presentation on ODIN, an oceanographic data collection and management system that he has been developing at the Bedford Institute of Oceanography, Canada. A more detailed account of this can be found in Annex 8. Several members of the WG were very interested in this software and A. Isenor agreed to make copies available. He also noted that it could be adapted to, for example, produce Cruise Summary Report forms at the end of a cruise. He provided a demonstration of the system. The WGMDM looked forward to hearing of further developments with the system in the future.

(ii) Taxonomic codes
R. Gelfeld provided a brief update on the taxonomic codes issue. NODC have frozen its Taxonomic Code system with Version 8.0 (on CD-ROM) and have switched to the Integrated Taxonomic Information System (ITIS). The initial on-line version of the ITIS database contains information from the NODC Taxonomic Code Version 8.0. The ITIS system is available on-line (http://www.itis.usda.gov/itis/) and will provide the Serial Number for the species requested. If it does not exist in the system, then a code will be allocated. NODC is participating in this project. The ITIS system is currently available for beta-testing. Users are encouraged to use the system, but to be aware that names may change status or position in the taxonomic hierarchy as groups are reviewed and modified. During this transition period the database is being actively updated with data that meet the quality criteria. The Web pages contain more details of progress.

(iii) Ocean Data Symposium
The Ocean Data Symposium was held in Dublin, Ireland, in October 1997, and was jointly organised by IOC, NOAA, EU MAST and the Irish Marine Institute. It followed on from the Climate Data Workshop held at the Goddard Space Center in 1992 (organised by CEC, ICES, ICSU, IOC and WMO). The objectives of the Ocean Data Symposium were to bring scientists, data managers and industry to a forum similar to the Climate Data Workshop; to assess the data management requirements of end users (scientists, data managers and industry); to deal with all aspects of marine data collection, methodologies, instrumentation and analysis techniques as well as data archaeology, dissemination, storage, retrieval, exchange and management; and to investigate the application of technological advances in order to increase the efficiency and effectiveness of present data management methods. There were four main themes: the data and metadata requirements of scientists in order to support ocean research; the benefits of statistical techniques and numerical modelling for analysis and prediction; development of advanced technology for data collection, analysis and exchange; and advances in information and data management tools for policy and decision makers.

Several members of the WGMDM had attended the Symposium and had presented papers and posters. It was felt to be a useful and successful meeting, and the WG were pleased to hear that another similar meeting would be held in 2-3 years time. In addition, the proceedings of the meeting would be published by IOC very soon.

(iv) IOC Group of Experts on the Technical Aspects of Data Exchange (GE-TADE)
L. Rickards reported that she had attended this meeting which had followed on from the Ocean Data Symposium. The main priorities for the GE-TADE meeting were: metadata, formats, OceanPC, data documentation and procedures, and raising the profile of IODE and GE-TADE.

It was agreed that it is unlikely that agreement will ever be reached on formats, and that perhaps data dictionaries are the way forward. RNODC(Formats) already holds information about ship codes, country codes and GF3 parameter codes on a Web page. GE-TADE members will look at the formats actually used to exchange data to try and work out which are the most commonly used formats, review them and come up with an 'approved' list. Guidelines may be better than formats - and to some extent these already exist, as GE-TADE produced a set of guidelines a few years ago (available from the RNODC(Formats) Web page). It was felt that the most commonly used formats are those required for major packages (e.g. SURFER, netCDF, ATLAST) and also comma separated values (.csv).

It was noted that there was a pilot project for MEDI, which gave the IODE/GE-TADE the opportunity to lead in the field of marine-related metadata. The Australian Blue Pages and EDMED were reviewed. A comparison of the fields in the two directories has been made and suggestions made as to which fields (there are about 15) should be used in MEDI. A MEDI pilot project product is needed to demonstrate to IODE at their next meeting in 2000.

The present status of OceanPC was reviewed and it was agreed what is really needed is a more integrated set of tools. OceanPC should include the following: it should be freely available; it should deal with coastal data, not just deep ocean - and a wider range of data types (e.g. time series, remote sensing); it should include a data dictionary; it should have tools for manipulating formats and it should be able to use commercial software, for example ACCESS and EXCEL. (It already links to SURFER).

(v) MAST Data Committee
F. Nast reviewed the activities of the MAST Data Committee, which is nearing the end of its life. However the EURONODIM project, which has been accepted by MAST, will largely replace the Committee. The MAST Data Committee looks at the data collected on MAST projects and draws up guidelines and policy for data management within MAST. EURONODIM will continue work with EDMED, Cruise Summary Reports etc.

J. Gagnon asked how MAST contributes to monitoring programmes. F. Nast replied that it is project driven and data collected are confidential until the end of the project. The data sets are often multidisciplinary and data centres are funded for the data management of specific projects. J. Gagnon felt that it would be beneficial to promote rapid release of some types of data - to go into the GTSPP, for example - as this would help climate modellers. A. Isenor added that although reduced profiles are sent for GTSPP in near-real time, the PIs have a proprietary period of two years over the full resolution data. He felt that MAST should encourage this early release of data. F. Nast agreed to pass on this to the MAST Data Committee at its next meeting.

15. Date and location of 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 last five years data (1994-1998) sent to ICES by each member country, identify problems and suggest solutions;

Although the data received by ICES over the last two years has been encouraging, there is still a large amount of data outstanding especially nutrient data and data from global projects. This item should act as encouragement to Member Countries to supply the ICES Oceanographic Data Centre with data in a timely manner.

b) Review progress in the implementation of IOC's Global Oceanographic Data Archaeology and Rescue (GODAR) Project in each member country, including consideration of biological oceanographic data types;

Much data have been recovered by the five year GODAR project, 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; investigations suggest that much biological data is available within ICES Member Countries. This item serves to help quantify the data and associated documentation available, and their status.

c) Quantitatively analyse the minimum requirements for quality assurance of oceanographic data;

There is a need for simple guidelines for those collecting, processing and quality assuring data. Having reviewed those guidelines and manuals presently available, and produced a set of guidelines for moored current meter, CTD, shipborne ADCP and SeaSoar/Batfish data, other data types will now be considered (e.g. moored ADCP, drifting buoys, XBT and sea level) and guidelines developed and updated.

d) Develop guidelines for the quality assurance and data management of nutrient and oxygen data in cooperation with the MCWG;

The MCWG have been reviewing quality assessment procedures for nutrient and oxygen data. Following on from this, the MCWG and WGMDM will jointly develop guidelines. The existence of written guidelines has distinct advantages. It shows laboratories reporting to the ICES data bank how important it is to apply quality control procedures on the data, and it will provide ICES with data sets which are easier to handle and which have a properly documented QC history behind them.

e) Report on the development of World Wide Web pages and links between them within member countries;

This is an opportunity to exploit developments within the Internet and raise the profile of the data centres within in the ICES community. In particular, collaboration on data products will be investigated and the WGMDM pages will be further developed.

f) Investigate and evaluate the data dictionaries available to the marine science community, including an analysis of the parameter code list used for the IOC Cruise Summary Report, and produce an improved and updated set of codes.

A number of Data Dictionaries, each covering a wide range of parameters, have been developed by the oceanographic community. Last year, these were critically reviewed by the WGMDM. An inter-sessional sub-group will continue this and suggest the appropriate hierarchical structure and standardistation at the category level. Data flagging schemes will also be addressed.

ii) Time and place of next meeting

The WG expressed its wish that the next meeting should be held at the Marine Environmental Data Service (MEDS), Ottawa, Canada, between 3 and 6 May 1999.

The Chairman closed the meeting by thanking the participants for their hard work, enthusiasm and valuable contributions. On behalf of the WG she thanked M. Fichaut for the excellent arrangements made for the meeting.

Annex 1

ICES Working Group on Marine Data Management, IFREMER, Centre de Brest, France.
20-23 April 1998

List of participants

Dr. Lesley Rickards,
BODC,
Proudman Oceanographic Laboratory,
Bidston Observatory,
Birkenhead,
Merseyside L43 7RA, U.K.
Tel: +44 151 653 8633
Fax: +44 151 652 3950
e-mail: ljr@pol.ac.uk
http://www.pol.ac.uk/bodc/bodcmain.html

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

Mr. Dave Hartley
Physical Oceanography,
UK Hydrographic Office,
Admiralty Way, Taunton,
Somerset TA1 2DN, U.K.
Tel: +44 1823 337900
Fax: +44 1823 284077
e-mail: dhartley@ocean.hydro.gov.uk
http://www.hydro.gov.uk/

Mr. Harald Loeng,
Institute of Marine Research,
P.O. Box 1870 Nordnes,
N - 5024 Bergen, Norway
Tel: +47 55 23 85 00
Fax: +47 55 23 85 84
e-mail: harald.loeng@imr.no
http://www.imr.no

Dr. Michele Fichaut,
IFREMER, Centre de Brest,
B.P. 70, 29280 Plouzane,
France
Tel: +33 2 98 22 46 43
Fax: +33 2 98 22 46 44
e-mail: mfichaut@ifremer.fr
http://www.ifremer.fr/sismer/

Ms. Maria Jesus Garcia,
Instituto Espanol de Oceanografia,
Corazon de Maria 8,
28002 Madrid, Spain
Tel: +34 1 3473612
Fax: +34 1 4135597
e-mail: mjesus.garcia@md.ieo.es

Mr. Jan Szaron,
Swedish Met. and Hydr. Institute,
Oceanographic Laboratory,
Byggnad 31 Nya Varvet,
426 71 Vastra Frolunda,
Sweden
Tel: +46 (0)31 751 8971
Fax: +46 (0)31 751 8980
e-mail: jszaron@smhi.se
http://www.smhi.se

Mr. Nico Kaaijk,
National Institute for Coastal
Marine Management/RIKZ,
P.O. Box 20907,
2500 EX Den Haag,
The Netherlands
Tel: +31 70 3114258
Fax: + 31 70 3114321
e-mail: n.m.kaaijk@rikz.rws.minvenw.nl

Mr. Jean Gagnon,
Marine Environmental Data Service (MEDS),
Department of Fisheries and Oceans,
200 Kent Street, Ottawa,
Ontario K1A OE6,
Canada
Tel: +1 613 990-0257
Fax: +1 613 993-4658
e-mail: gagnon@ottmed.meds.dfo.ca
http://www.meds.ca/

Mr. Friedrich Nast,
DOD/BSH,
Bernhard Nocht Strasse 78,
Postfach 30 12 20,
D-20305 Hamburg,
Germany
Tel: +49 40 3190 3530
Fax: +49 40 3190 5000
e-mail: nast@m5.hamburg.bsh.d400.de
http://www.bsh.de/

Ms. Riitta Olsonen,
Institute of Marine Research,
P.O. Box 33,
00931 Helsinki,
Finland
Tel: +358 0 613971
Fax: +358 0 61394494
e-mail: riitta@fimr.fi

Ms. Sara Almeida,
Instituto Hidrografico,
Oceanography Department,
Rua das Trinas, 49,
1296 Lisboa Codex,
Portugal
Tel: +351 1 3955119
Fax: +351 1 3960515
http://www.insthidrografico.pt

Ms. Orla Ni Cheileachair,
Irish Marine Data Centre,
80 Harcourt Street,
Dublin 2,
Ireland
Tel: +353 1 4757100
Fax: +353 1 4757104
e-mail: orla.ni@marine.ie
http://www.marine.ie

Mr. Helge Sagen
Institute of Marine Research,
P.O. Box 1870 Nordnes,
N - 5024 Bergen, Norway
Tel: +47 55 23 85 00
Fax: +47 55 23 85 84
e-mail: helge.sagen@imr.no
http://www.imr.no

Mr. Anthony Isenor,
Bedford Institute of Oceanography
P.O. Box 1006,
Dartmouth,
Nova Scotia,
Canada B2Y 4A2
Tel: 902 426 4960
Fax: 902 426 7827
e-mail: isenora@mar.dfo-mpo.gc.ca

M. Gerald Riou,
IFREMER, Centre de Brest,
B.P. 70, 29280 Plouzane,
France
Tel: +33 2 98 22
Fax: +33 2 98 22 46 44
e-mail: griou@ifremer.fr
http://www.ifremer.fr/

Dr. Catherine Maillard,
IFREMER, Centre de Brest,
B.P. 70, 29280 Plouzane,
France
Tel: +33 2 98 22 4? ??
Fax: +33 2 98 22 46 44
e-mail: cmaillard@ifremer.fr
http://www.ifremer.fr/sismer/

Ms. Mathilde Pitel,
SISMER,
IFREMER, Centre de Brest,
B.P. 70, 29280 Plouzane,
France
Tel: +33 (0)2 98 22 41 91
Fax: +33 (0)2 98 22 46 44
e-mail: mpitel@ifremer.fr
http://www.ifremer.fr/sismer/

Annex 2 Terms of Reference

The Working Group on Marine Data Management [WGMDM] (Chairman: Dr. L.J. Rickards, UK) will meet in Brest, France from 20 - 23 April 1998 to:

a) assess the post-1990 oceanographic data sent to ICES by each member country, and identify problems and suggest solutions;

b) review progress in the implementation of IOC's Global Oceanographic Data Archaeology and Rescue (GODAR) Project in each member country, including consideration of biological oceanographic data types;

c) quantitatively analyse the minimum requirements for quality assurance of oceanographic data;

d) report on the development of World Wide Web pages and links between them within member countries;

e) instigate an analysis of the parameter code list used for the IOC Cruise Summary Report, and produce an improved and updated set of codes;

f) investigate the Data Services available from NODCs in member countries and suggest a scheme to improve cooperation between countries to provide an improved service to the community;

g) investigate and evaluate the data dictionaries available to the marine science community;

h) consider the future work programme in relation to the remit of the Oceanography Committee and development of the ICES Five-Year Plan, including cooperation with other Working Groups;

i) comment on the 1997 ACME statement (Agenda Item 21.3) concerning the development of GOOS initiatives in ICES.

WGMDM will report to the Oceanography Committee at the 1998 Annual Science Conference.

Justification:

a) Although the data received by ICES post-1990 over the last year has been encouraging, there is still a large amount of data outstanding, especially nutrient data and data from global projects. This item should act as encouragement to member countries to supply the ICES Oceanographic Data Centre with data in a timely manner.

b) 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 the lead in this project for the ICES area, which provides a focus for member state activities. Initial investigations suggest that much biological data is available within member countries. This item serves to help quantify the data and associated documentation available, and their status.

c) There is a need for simple guidelines for those collecting, processing and quality assuring data. Having reviewed those guidelines and manuals presently available, and produced a set of guidelines for moored current meter data, CTD and nutrient data, other data types will now be considered (e.g. ADCP, SeaSoar/Batfish, XBT and sea level) and guidelines developed and updated.

d) This is an opportunity to exploit new developments within the Internet and raise the profile of the data centres within the ICES community. In particular, WGMDM pages will be further developed.

e) The results of the intersessional sub-group work in mapping the existing Cruise Summary Report codes to the JGOFS data dictionary codes will be critically reviewed.

f) Collaboration will lead to increased data exchange and efficiency, and better collaboration between the NODCs.

g) A data dictionary covering a wide range of parameters has been developed for JGOFS. This and other data dictionaries known to the WG will be examined to determine the most appropriate system to use.

Annex 3 Highlights from the reports of the Data Centres

The reports submitted to the WGMDM meeting can be found on the WGMDM Web pages at:

http://www.pol.ac.uk/bodc/mdm/dcreports.html

The highlights below provide a brief summary of the reports.

ICES: During 1997, 55055 profiles were added to the database for the years subsequent to 1980. The distribution of the number of profiles by year is given in Annex 4. For the first time there are more than 20000 profiles in any one year (i.e. in 1988 and 1990). A number of submissions have been received in the weeks leading up to the MDM meeting, notably from Finland and France, and these have yet to be processed. Concern persists around the very low submissions from a number of countries, especially with regard to UK (NERC), Germany, Ireland, Spain, Portugal, and Norway (for nutrients). One country has requested the withdrawal of all of its data for the period 1989-1991 because of suspected quality problems.

A special archive of MAST ROSCOPs is maintained, and these are listed on a special MAST part of our web site. In spite of the compulsory requirement to provide ROSCOPs for MAST projects it remains difficult to receive them. Activities in connection with the MAST Projects ESOP and VEINS are well underway.

All the OMEX profile and surface data sets have been merged into the ICES databank. This task has served as a test for expanding the ICES format to encompass any number of data types, and also to see how ROSCOP may be adapted to reflect the expansion in parameters. The nucleus of both of these developments has been the BODC/JGOFS data dictionary.

Software systems have continued to be developed to facilitate the data management activities. No proprietary software is in use apart from producing final graphical products. The software is both Windows-based (data management), and Unix-based (for the preparation of gridded products which now represents more that 50% of the requests received).

Canada: The Marine Environmental Data Service (MEDS) is a branch of the Fisheries and Oceans Science Directorate of Canada's Department of Fisheries and Oceans (DFO). Its mandate is to manage and archive physical and chemical oceanographic data collected by DFO Regional Institutes or acquired through various arrangements from Canadian researchers in government, university and industry, and international research conducted in the Great Lakes, the St. Lawrence River and the major ocean areas adjacent to Canada.

MEDS continued its ongoing programs relevant to the acquisition, processing, quality control, dissemination and archival of physical oceanographic data received in both operational and delayed mode. These included over 75K wave spectra, 3K days of tidal hourly height data, 1.5M drifting buoy messages and 60K temperature-salinity profiles, all of which were reporting in real-time.

Two major initiatives within DFO Science were initiated to co-ordinate ocean science in Canada, and in particular its data management. The formation of a National Data Management Working Group, chaired by MEDS, to coordinate physical, biological and fisheries data within DFO Science, and an Atlantic Zonal Monitoring Program for which MEDS will be the focal point for the safekeeping and dissemination of data and information through a centralised World Wide Web server.

Denmark: The Oceanographic Department (OD) in the Royal Danish Administration of Navigation and Hydrography (RDANH) has continued the operation of (i) the network of tide gauges in Danish waters, (ii) the network of oceanographic stations (equipped with current meters and C/T-chains) in Danish waters, and (iii) the network of tide gauges in Greenlandic waters. A hydrographic cruise covering six hydrographic east-west sections along the Greenlandic westcoast was carried out in June 1997. The data will be submitted to ICES soon.

The main task of OD is to maintain a network of stations that collect data for dissemination in real time to the users. Also numerical models are important tools in operational oceanography. They provide forecasts for the oceanographic parameters sea level, current, salinity and temperature. RDANH has been invited by the Swedish Meteorological and Hydrological Institute (SMHI) to take part in a working group for the development of an oceanographic model for the Baltic-North Sea area called HIROMB (High Resolution

Ocean Model for the Baltic). The main input from OD will be real time data from our network of stations; validation of the model output will be another topic. The HIROMB model will be operational spring this year and forcasted fields will be available on the Internet.

Finland: In 1997 R/V Aranda made 15 cruises. Along with the routine monitoring cruises there were several cruises connected to international and national research projects. Information is provided in the institute's web pages. Hydrographic and chemical data from 8 cruises are in the central data bank. The data consist of 189 stations, among them four helicopter stations from the Bothnian Bay.

The CTD data from 1990-1996 have been thoroughly checked and the corrected data from standard depths are brought to the central data bank. These data have also been sent to ICES, as well as the whole hydrographical and chemical data from 1995-1996. An inventory of the data sets in the institute was made. The directory is according to the EDMED format. This is really useful for many purposes. The institute achieved accreditation for biological methods in 1997.

A remarkable collection of biological data exists that have been measured in 6 - 7 coastal stations in 1962 - 1997. This data consists of chlorophyll a, phytoplankton, zooplankton and primary production data, even few hydrographical and nutrient data recordings are included. The data are dispersed in papers, diskettes and magnetic tapes.

France: SISMER has carried out quality checks of all the new hydrological data and part of the historical database. Now, all the CTD data set has been controlled (13154 CTD casts); and 30% of the bottle data are controlled too (9374 bottle casts among 32459). In terms of exchange with international Data Centers, SISMER have recently sent all its new CTD and bottle data sets (from 1987 to 1997) to ICES and to NODC/WDCA: 1873 CTD stations from 30 cruises and 95 bottle stations from 2 cruises. The ROSCOPs of the 1996 cruises (153 cruises) have been sent to ICES in February of 1998. These ROSCOPs files are available on the Web.

MEDATLAS: 1997 was the year of achievement of the project and the MEDATLAS CONSORTIUM (IFREMER/SISMER, NCMR/HNODC, IEO, SHOM) for which SISMER was the coordinator, have produced a Mediterranean hydrological atlas on a set of 3 CD-ROMs. The final data set contains 50695 temperature and salinity profiles (bottle and CTD casts) and 154911 temperature profiles (XBT and MBT) with quality flags for each measured value. A selection software (SELMED, written by IFREMER/DITI/IDT/ISI) allows easy extraction of data from the CD-ROM following several criteria: data types (CTD, Bottle, XBT, MBT, Thermistors), measured parameters (temperature, salinity, chemicals), quality flag, period, geographical location, ship, source country, cruise name or identifier. The observed data or data interpolated to standards levels can be extracted from the CD-ROM. The CD-ROMs contain also the gridded climatological statistics computed at 28 horizontal levels. For temperature, statistics are monthly from the surface down to 300 meters depth, seasonal between 400 and 800 meters depth and annual below. For the salinity, climatology is seasonal from the surface down to 800 meters depth and annual for deeper levels. Finally the CD-ROMS contain a selection of climatological maps at the Postscript and the Gif format.

SISMER is now turning its attention to the MTP II - MATER (Mass Transfer and Ecosystem Response) project. This project, which includes 55 research groups from 16 different countries, is a Mediterranean Targeted Project.

Germany: Routine water bottle and biota data submissions have continued. 1997 was a good year for CTD data -data were received in 25 formats, reformatted and forwarded to ICES. 220 Cruise Summary Reports were sent to ICES, over 100 from 1997. Cruise Summary Reports from visiting ships are also collated. Inventories and Cruise Summary Report information has been put on the Web, and is updated monthly. Inventories can be queried by ship and year. Future cruise schedules and details of monitoring programmes are also available. This has a high visibility for funding agencies. The Web page development is continuing.

Digital request forms are available over the Web, and this has led to an increase in the requests dealt with via the Web. It also means that the requestor needs to be precise in defining their request.

DOD was audited during the year. The outcome of the review was that a central archive is needed, but that little project-oriented work is carried out. However, it was deemed important that in formulating plans for data collection, the location of the final archive place should be included. It was also suggested that climate change simulations should be archived. At present there are no archives for geophysical data, ice cores, current meter data and ship's cruise data. DOD will do the last two of these, but the others are undecided at present.

German scientists would like data and cruise reports on-line. Ideally they would like to be able to click on data and retrieve it. DOD will be working on this over the next two years.

Ireland: In 1997, the Irish Marine Data Centre developed a five year strategic plan and focussed on consolidating its core data management activities, integrating with its parent organisation, the Irish Marine Institute and formulating its requirements in terms of human resources, IT strategies and infrastructure. Certain *ad hoc* core activities have been suspended until key core programmes and infrastructure are secured.

In association with the Fisheries Research Centre, Aquaculture and Environment Section, the Data Centre have been developing the Environman Database to facilitate A&E reporting to ICES for water quality, sediment and biota. This modified system is currently being populated and tested with the data collected by the Fisheries Research Centre. Data management requirements for environmental data have also been outlined in the Marine Institute strategy for environmental R&D which is currently awaiting Government approval.

EDAP During 1997, the Irish Marine Data Centre completed the MAST II Supporting Initiative on Electronic Data Publishing (EDAP). This was a significant milestone in terms of the Data Centre's development and the work is underpinning national marine data management activities in 1998.

The Irish Marine Data Centre hosted the Ocean Data Symposium in October 1997 with over 150 delegates from 22 countries. Report on proceedings will be published as IOC/IODE Technical Series and is expected to be available in June 1998.

The Irish Marine Data Centre is currently responsible for the data management of the following MAST III projects: CANIGO (in conjunction with the Spanish Data Centre) BENGAL, ENAM II and COLORS.

Norway: During 1997 the Institute of Marine Research, IMR, deployed 6 moorings with total of 22 current meters and completed 3414 hydrographic stations for fisheries and environmental projects. CTD profiles was performed by Johan Hjort 1167 stations, G.O.Sars 1254 stations, Michael Sars 874 stations, Jan Mayen 119 stations, total of 3414 stations. Data from 1997 are quality controlled, and data from the first 6 months have been sent to ICES. Data from the last 6 months were converted to 5 dbars intervals (instead of 1 dbar) due to low quality on CTD instruments. All calibration data are ready. Water samples were also taken on many stations leading to nutrients, chlorophyll data (about 25000 samples (x 6)) and phytoplankton data.

The work with Norwegian Standards on moored current meter data and measurements of temperature and salinity have been continued. The current meter standard is close to being finished.

The MAST III project TASC (Transatlantic Study of Calanus finmarchicus) has its data management homepage at IMR on http://tasc.imr.no/tasc/datamanagement.html/. IMR is responsible for getting the data sampled by partners available to partners and stored the data in a database. The final banking of data will be at ICES. 1997 was the main year of data sampling. Environmental data have been sent to IMR, but zooplankton data are delayed. The data are presented on the web using the US GLOBEC JGOFS software to view data in a web browser. Work is being done to get the IMR database model to communicate with web browsers. This is being done using Open Ingres/ICE relational database system.

Portugal: The Oceanographic Department of IH presently comprises the areas of Physical Oceanography and Marine Geology. The SEFOS (Shelf Edge Fisheries and Oceanography Studies) project was concluded in November 1996 and the final report delivered in May 1997. Current measurements over the upper slope initiated with that project, near latitude 40 degrees N at depths 50m, 100m, 300m and 600m, have been maintained operative since then, aiming to build up a long time series.

Data acquisition and processing from directional and non-directional waveriders continued, and statistics presented in internal reports.

The tide tables 1998 for harbours of Portugal and Portuguese speaking countries were published. IH continues the quality control of tidal data from the national tide gauges network so it can be sent to the UH Sea Level Center.

The inventories of current meter, thermistor chain moorings and meteorological data were updated to 1998.

A conceptual model to support the relational database of oceanographic data was created in Oracle. The metadata concerning the Portuguese oceanographic cruises and moorings of current meters was prepared, to feed that system.

Spain: During 1997, the IEO has carried out 64 cruises for fisheries, ecology, physical and geophysical projects. Also the 12 stations of the tide gauge network along the Spanish coast is still operative.

The data collected in the IEO RADIALES project (Studies on time series of oceanographic data on several transects along the Spanish coast) are being managed with the Oracle Relational Data Base. Up to now, the data of the Santander transect for the period 1994-1997 have been banked in that database, and the data for the remaining transects are beginning to be processed this year. New modules for quality control have to be implemented in this database. The controls for physical data are more or less resolved. However, for chemical and biological data the best approach to a data quality control is to give as much information as possible in the metadata.

The inventory of the data collected during the MAST3 CANIGO project is being maintained in the IEO Data Center and also the IEO is responsible together with ISMARE for banking all the data in order to share the data between the project partners during the course of the project, and prepare some data product for distribution at the end of the project. In this case the data that will come to the Data Center will be already qualified.

In Spain there is a project called RAYO (Alert and Observation Network) leaded by PE (Puertos del Estado) that has 8 buoys deployed along the Spanish coast that transmits the data to the PE Center. At present, the PE and the IEO are preparing an agreement for installing marine sensors (CTD, chlorophyll) and current meters on those buoys. The agreement will also contemplate, for the RIMA project (Integrated Spanish Tide Gauge), the data assimilation to give the tides prediction daily, including the astronomical tides and meteorological effects.

Sweden: SMHI acts as "national data host for physical and chemical oceanographical data" from national and regional marine monitoring programmes. 993 series from the national monitoring programme for 1997 and 632 series from the regional programmes were added to the Swedish National DataBank (SHARK). SHARK has also been expanded with historical data from other Baltic countries.

Water bottle (including nutrients and chlorophyll) - and compressed CTD - data from the R/V ARGOS for 1995 have been submitted to ICES. Water bottle (including nutrients and chlorophyll) - and compressed CTD - data from ARGOS for the IBTS-cruise in Jan-Feb 1997 have also been submitted to ICES. A complete set of ROSCOP files from ARGOS for 1997 have been submitted to ICES. Historical marine biological data have been submitted to the Stockholm University, Department of Systems Ecology, who acts as "national data host for marine biological data"

SMHI has managed to maintain the high number of cruises so that the main stations in Skagerrak, Kattegat, The Sound and Baltic proper were visited almost/more than once a month. SMHI also continued to perform monthly monitoring in the near coastal zone in four counties in the west, south and southeast of Sweden. SMHI also took an active part (together with institutes in Germany and Poland) in marine data collection and management in connection with the flood-disaster in Oder and Vistula in July-August 1997

UK (BODC): BODC has been the project data centre for the EU-MAST Ocean Margin Exchange (OMEX I) programme. In November 1997, a two CD-ROM set was published containing the data from 47 research cruises undertaken by ships from nine countries operating on the European Continental Shelf Break between Portugal and Norway from April 1993 to November 1995. Over 95% of the 600 data sets collected during the field programme are assembled on the CD-ROMs.

BODC operates a WOCE Sea Level Data Assembly Centre (DAC), and has been doing so since early 1991. Over 3000 site years of data are currently held; 1600 series were quality controlled and added to the data bank during the year. A master CD-ROM was created in March 1998 including this sea level data set; also included on the CD-ROM are the 'fast delivery' Sea Level DAC data holdings, tidal harmonic constants from the WOCE Sea Level Data, the PSMSL data holdings (and other contents of their public access directory), an updated version of the GLOSS Handbook (Version 4.0) and two IOC sea level measuring manuals. The CD-

ROM forms part of a set of 13 WOCE CD-ROMs which will be available to participants at the WOCE Scientific Conference 'Ocean Circulation and Climate' to be held in Halifax, Canada (May 1998).

The UK Inter-Agency Committee on Marine Science and Technology (IACMST) has established a Marine Environmental Data (MED) coordinator and Advisory Group to facilitate communication on a regular basis among MED data managers and sources. The post is hosted at BODC and the remit includes: maintaining an inventory of UK sources of MED and their holdings, providing advice on the management and quality control of data, advising potential users of MED on their availability, serving as a UK focus of international MED issues and MED exchange, acting as the focal point for the UK distributed network of MED and convening the UK MED Advisory Group.

UK (CEFAS): The UK Directorate of Fisheries Research, which comprised the Fisheries Laboratory at Lowestoft and three smaller laboratories in the UK, was a division of the Ministry of Agriculture, Fisheries and Food. As of April 1997 the Directorate became an executive agency of the Ministry and was renamed CEFAS, the Centre for Environment, Fisheries and Aquaculture Science. The new title describes the three principal areas of scientific activity and agency status gives CEFAS more autonomy to run its affairs.

Scientific work is conducted by three groups, which again identify the areas of investigation: (1) Fisheries science and management (e.g. stock management and population dynamics, fish behaviour and physiology), (2) Environment (e.g. regulatory monitoring and assessments, pollution effects in the marine environment, radiological monitoring, assessments and services, physical and biogeochemical processes), and (3) Aquacutlure and Health (e.g. fish cultivation, shellfish cultivation, pathology inspectorate).

The Environment Group includes the physical processes and biogeochemical teams who work closely together. The JONUS2 (Joint Nutrient Study 2) field program, a joint exercise with other UK laboratories and a successor to JONUS1, was completed during 1997 and aims to quantify the flux of nutrients through estuaries and assess the potential impact upon UK coastal waters.

Oceanographic studies include a program to understand the circulation and transport around the Irish Sea and the North Channel. This used data from current meter moorings, satellite tracked drifting buoys and a CTD mounted on a SCANFISH undulating towed body. The study is planned to continue into the Celtic Sea during 1998.

UK (Fisheries Research Services): During 1997, the Marine Laboratory deployed 16 instrument moorings and completed 661 hydrographic stations for fisheries and environmental projects being undertaken by the laboratory. The instruments deployed were 17 current meters, 6 water level recorders and 3 ADCP. Of the 661 hydrographic stations 441 of these included CTD profiles.

All valid data recovered from the instruments deployed have been sent to BODC with the relevant documentation except for two moorings that will be recovered in April. The reversing bottle data for 1997 is in the process of being finalised and will be sent on to ICES in the forthcoming weeks. The International Young Fish Survey data for 1998 shall also be sent. The CTD data has been sent to BODC and will be sent to ICES when the water bottle data is completed. The 1997 Cruise Summary Reports are in the process of being compiled and will be sent to both BODC and ICES in the coming weeks.

The FRV Scotia which has served the Marine Laboratory for the past 26 years has been replaced by a new vessel bearing the same name. This multi-disciplined research vessel was launched at Ferguson's ship yard, Port Glasgow on 4th July 1997 and was followed by fitting out work and sea trials. The ship was accepted from the builders in March, 1998 and will undergo a short period of familiarisation prior to a busy first year of research trips.

UK (HO): The UKHO continues to maintain and populate its major global oceans observations database which principally contains observations of the physical parameters of temperature, salinity and sound speed. Data are received from a variety of sources including the (UK) Royal Navy (RN), foreign navies, ships of opportunity and civil institutions both in UK and overseas. The data processing task includes quality assurance (QA) checks utilising both software tools and the experience of staff in order to maintain the integrity of the database for UKHO-processed data. Data from the RN are received in raw form and undergo rigorous examination and, where required, editing before they are incorporated into the database. Other data are received in processed form from other institutions, requiring a different approach to the QA task, often involving some degree of software support to reformat the data.

The UKHO is the national data centre for BT observations. RN vessels routinely take XBT observations on synoptic hours whilst on passage using Sippican - and Sparton - manufactured T7 and T5 probes. Other XBT data are received from scientific cruises. In the year to March 1998, over 6000 XBT raw observations were processed of which 80% were accepted for inclusion into the observations database after QA checks and validation (noting that not all observations are conducted in ideal operating conditions).

The development of the recently established non-acoustic biological database continues with effort being presently concentrated on the sourcing, population and storage of both quantitative and qualitive data. ICES has, on request, undertaken to supply certain fish statistics and new sources of data would be most appreciated by the project manager, Dr Robin Hensley.

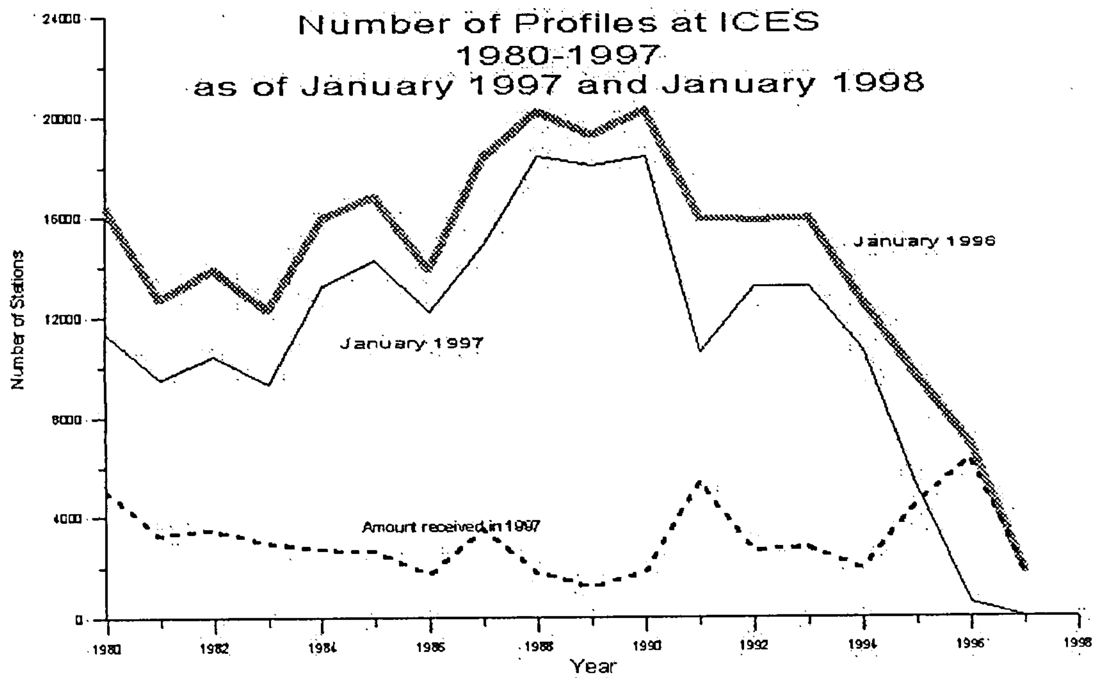
USA: During the year, 1996-97 ROSCOPs were sent to ICES to add to ROSCOP Database. The total U.S. ROSCOPs are now 6500.

The World Ocean Database 1998 (WOD98) was published. This provides additional data and has expanded the WOA94 to include additional variables such as chlorophyll, nitrite, alkalinity, pH, and plankton. Two million temperature profiles have been added to the historical archives of oceanographic data as well as 600000 plankton observations and 140000 profiles of chlorophyll. More than 5.4 million temperature profiles are available in WOD98, making it the most complete digital oceanographic database available to the international research community.

NODC Coastal Ocean Data Resources and Activities: Three Coastal Ocean Data Working Groups (acquisition, data and information products, and quality assurance in Silver Spring on October 28-31, 1997 and March 17-19, 1998) were hosted. These working groups were established in response to recommendations made by stakeholders at the NOAA Coastal Ocean Data Workshop held in March. NODC, NGDC, and NCDC staff presented background information on the history and current status of the Data Centers. Working group members reviewed their terms of reference, decided on a strategy for the next two years, and made a number of initial recommendations.

NOAA Virtual Data System (NVDS): This is a unified, seamless data access and delivery system which enables the entire NESDIS data system to work and integrate more effectively in a timely manner. It will offer an internet site, customer account and ordering system, data visualization and fusion tools.

Annex 4 ROSCOP and data flow into the ICES Oceanographic Data Centre



ICES ROSCOP Submissions 1970-1998 (as of 22/06/98)

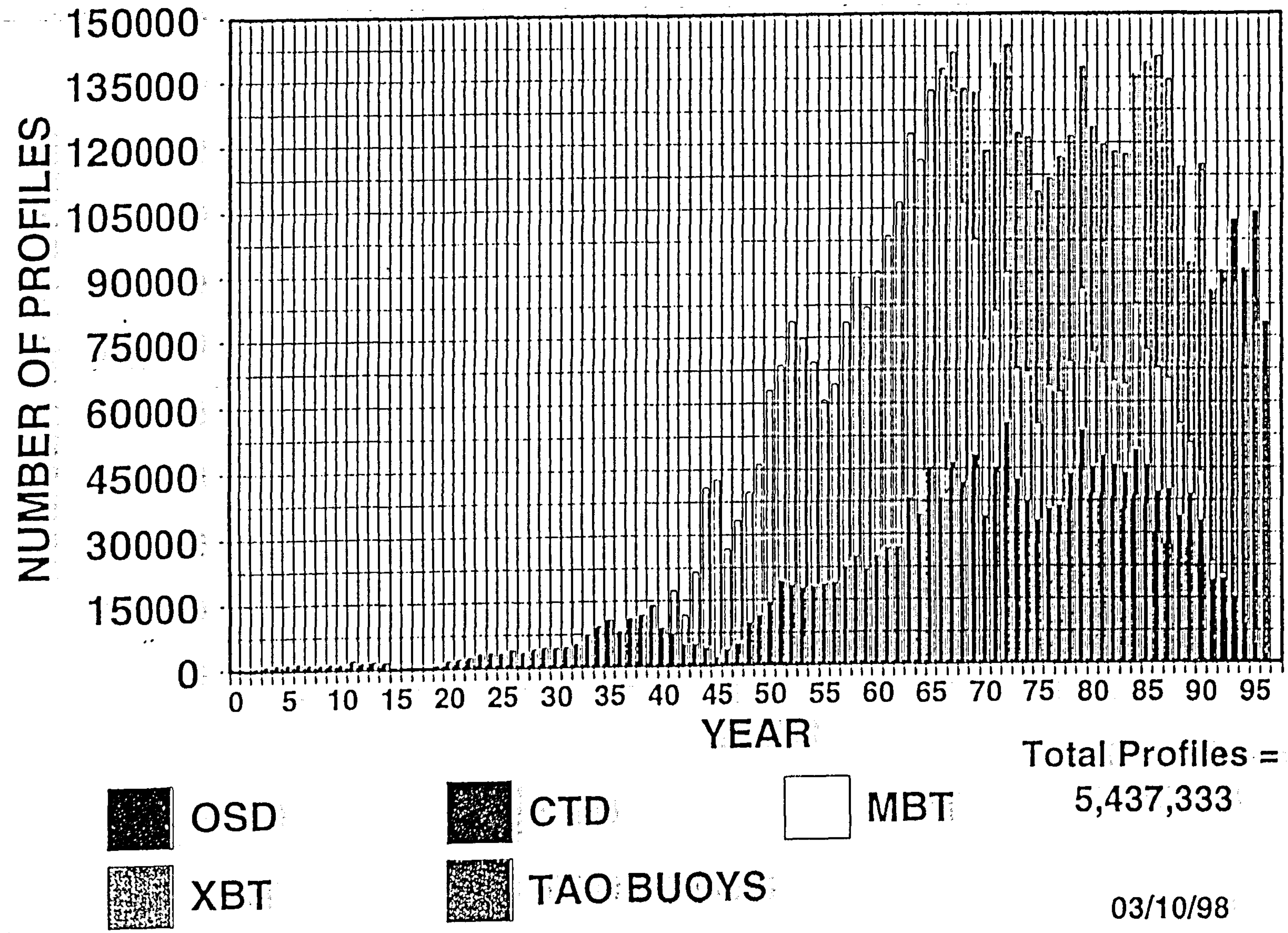
Country/Year	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	Total*
Belgium	1	6	7	1	1	1	3	1	1	1	1	1	1	1	0	0	0	0	0	21	26	24	25	30	31	26	30	27	0	267
Canada	12	5	14	4	1	13	9	17	18	20	14	21	16	13	0	25	9	41	52	13	15	17	27	41	27	1	25	0	0	470
Denmark	5	8	16	8	21	24	40	32	16	16	18	9	19	31	0	38	42	39	39	22	19	24	20	20	26	22	21	42	6	643
Estonia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Finland	4	3	6	7	7	8	0	11	13	11	13	11	10	4	0	12	1	0	0	0	0	0	0	0	0	0	0	1	3	125
France	38	57	60	52	76	68	55	62	64	76	88	73	68	76	0	52	53	34	55	97	127	108	82	86	134	99	155	2	1	1998
FRG	47	96	64	83	48	59	68	64	65	52	70	92	92	96	0	130	132	118	141	173	167	164	140	190	200	192	152	108	2	3005
GDR	1	2	7	6	7	9	8	12	10	11	9	13	10	11	0	10	10	8	11	8	15	0	0	0	0	0	0	0	0	178
Iceland	11	20	19	9	10	11	10	16	17	16	18	12	12	14	0	12	9	19	8	6	9	11	8	16	16	13	4	27	5	358
Ireland	1	1	2	13	1	5	3	5	5	5	0	0	0	0	0	0	0	16	4	0	9	2	10	11	3	2	0	0	0	98
Latvia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	0	0	2	0	7
Norway	27	31	25	26	26	38	31	28	28	35	35	40	41	48	0	48	38	47	46	49	40	42	71	56	63	70	84	100	20	1233
Netherlands	13	23	27	24	38	33	26	28	57	69	70	63	83	74	0	76	74	26	0	83	82	81	13	10	17	23	14	6	1	1134
Poland	14	10	32	25	15	9	14	12	11	5	15	11	14	13	0	16	15	11	12	11	14	15	7	4	10	3	13	4	0	335
Portugal	0	0	0	1	0	0	1	2	0	1	1	2	1	0	2	1	1	2	2	1	1	0	0	2	5	4	0	0	0	30
Spain	5	4	4	5	4	8	4	2	2	5	1	0	0	0	0	2	1	5	4	1	4	25	34	28	27	27	22	18	0	242
Sweden	10	12	8	9	9	13	24	19	9	8	9	11	10	14	24	23	23	24	9	3	4	14	0	20	19	19	19	19	9	394
UK	105	216	173	178	187	163	183	197	181	164	150	181	155	135	115	114	108	119	114	117	132	135	120	134	81	82	64	7	0	3810
USA	6	11	14	63	394	650	731	707	654	377	311	292	127	228	190	192	126	167	152	132	120	143	112	115	81	44	33	24	0	6196
USSR	10	18	13	6	5	3	16	15	3	3	1	1	4	7	2	3	3	0	0	0	4	1	0	0	0	0	0	0	0	118
Other	1	1	0	0	0	0	0	0	0	4	4	6	3	11	0	15	10	10	13	13	1	9	117	110	110	48	62	1	0	549
Total	311	524	491	520	850	1115	1226	1230	1154	879	828	839	666	776	333	769	655	686	662	750	789	815	786	876	852	675	698	388	47	21190

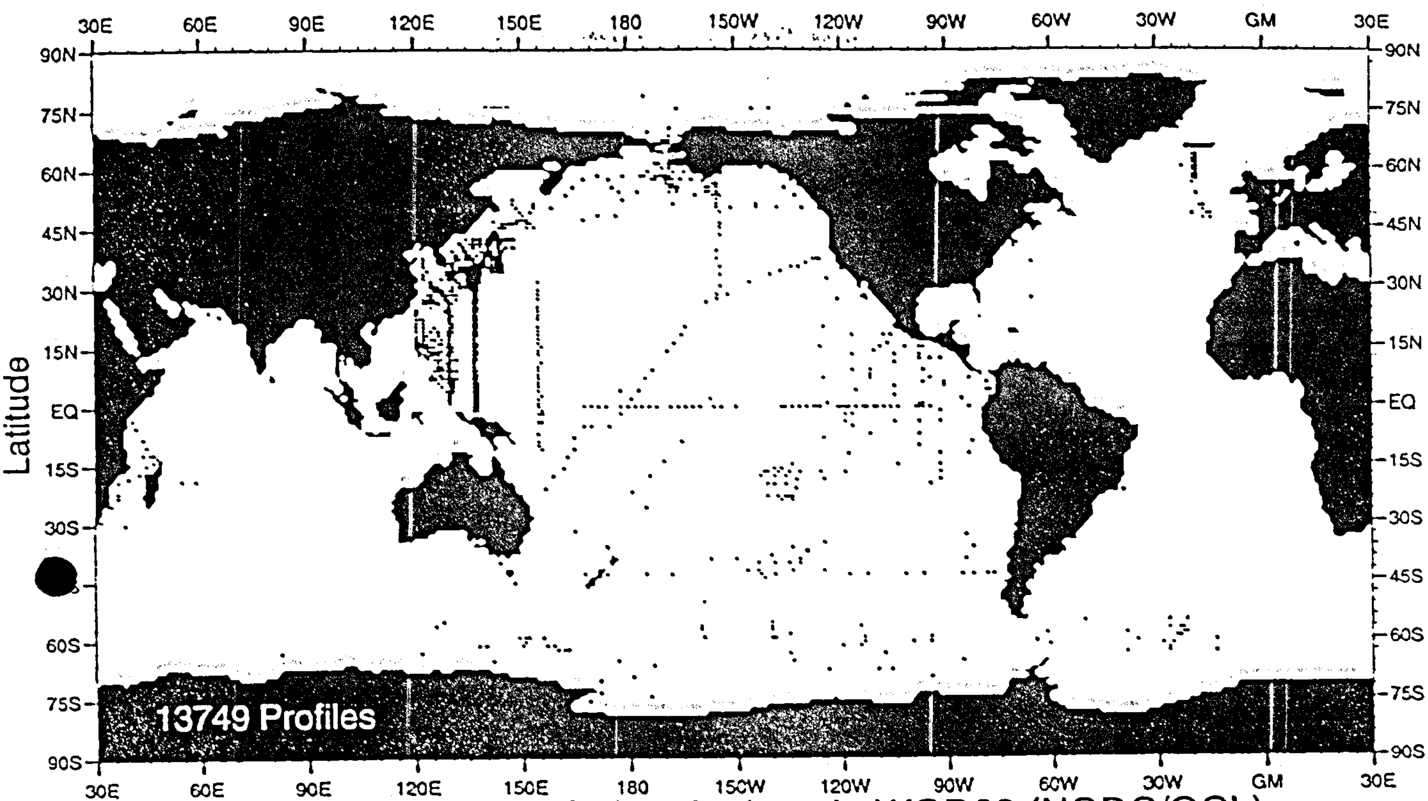
* Totals include only ROSCOP submissions only, ie excludes those forms created by ICES which total 4113 forms)

Number of Cruises where data held at ICES (as of 22/06/98 - Source: ROSCOP)

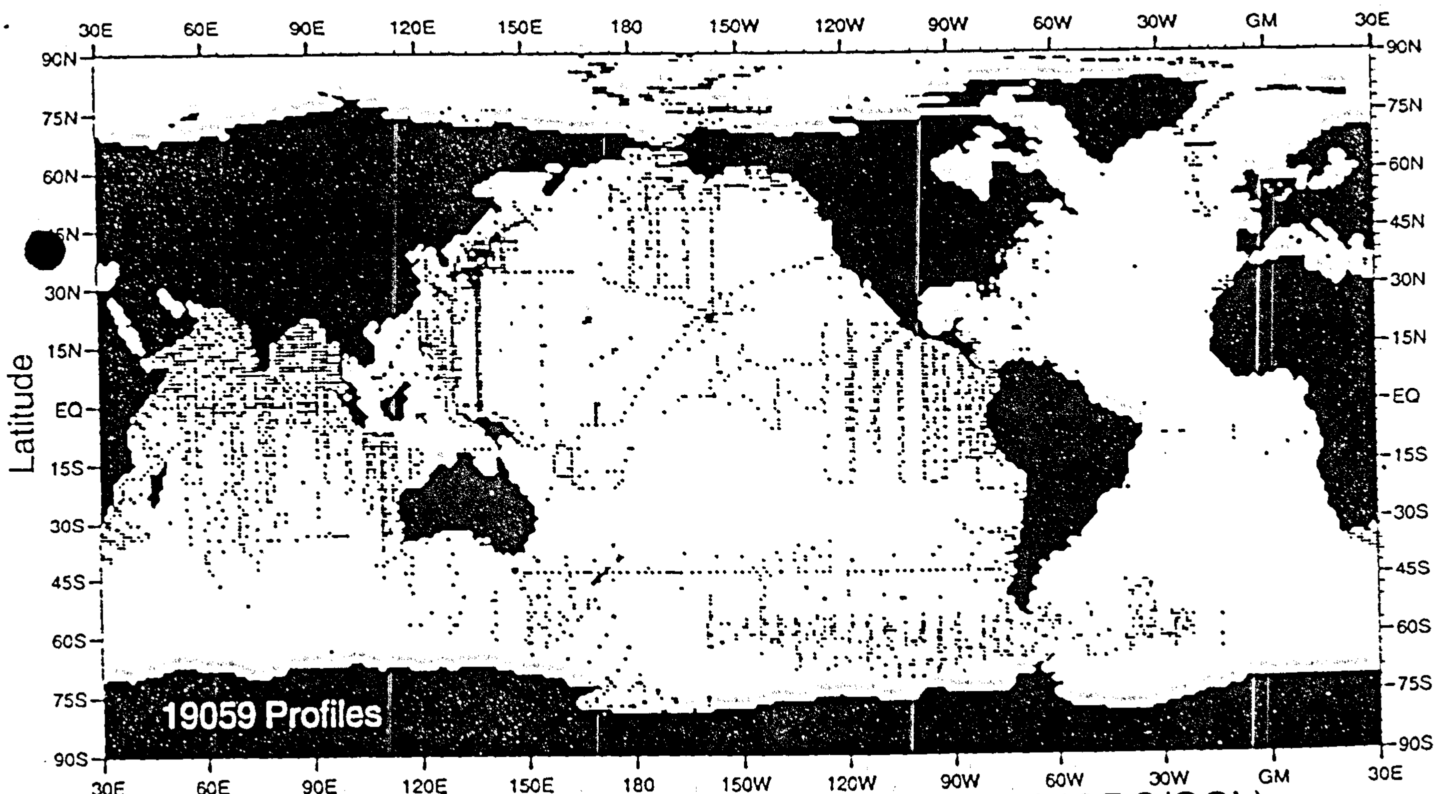
Country/Year	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	Total
Belgium	0	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1	1	2	1	9	8	7	6	9	10	10	6	0	0	87
Canada	0	0	0	1	0	0	1	0	0	39	43	52	48	44	41	45	33	24	28	26	17	15	8	3	2	3	1	0	0	474
Denmark	7	6	13	7	17	15	31	20	13	12	13	8	13	16	22	20	22	21	28	19	31	18	19	19	26	24	7	0	0	467
Estonia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Finland	7	4	6	7	7	8	0	15	13	15	15	17	11	4	14	19	16	20	11	4	7	16	9	11	14	0	0	1	0	271
France	7	13	16	24	19	15	12	12	9	10	10	8	16	21	16	6	5	7	12	10	12	25	15	17	18	3	1	2	1	342
FRG	32	21	18	18	21	20	24	21	21	16	44	48	49	41	51	53	51	44	46	49	65	38	47	52	16	22	9	4	0	941
GDR	5	9	6	5	3	3	6	5	5	5	6	5	5	7	7	6	8	7	6	7	6	0	0	0	0	0	0	0	0	122
Iceland	6	13	11	9	10	11	10	16	17	16	18	12	12	14	10	12	9	19	9	6	9	10	8	16	15	0	1	0	0	299
Ireland	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	6	1	0	0	0	0	11
Latvia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	8	6	7	0	23
Norway	36	17	23	19	13	14	29	37	30	27	46	45	45	46	61	54	42	48	57	59	75	68	77	68	80	64	64	67	0	1311
Netherlands	23	14	17	12	17	12	10	5	8	7	8	6	8	9	2	2	4	6	2	6	6	15	23	35	34	28	25	1	0	345
Poland	14	8	10	13	15	9	2	6	0	3	10	11	9	7	7	4	5	22	12	9	17	19	11	15	14	18	0	1	0	271
Portugal	7	17	6	9	2	0	0	0	0	0	0	0	1	1	0	1	0	0	2	1	2	0	0	2	4	4	0	0	0	59
Spain	1	2	1	4	2	1	3	0	0	0	0	0	0	0	0	1	0	2	1	0	1	10	16	16	24	31	11	13	0	140
Sweden	25	32	36	19	18	39	30	32	23	16	18	19	16	14	25	24	26	28	32	31	44	15	14	22	22	19	1	1	0	641
UK	21	38	33	42	21	24	33	35	27	27	24	31	22	17	15	26	25	33	40	61	54	30	38	43	35	30	20	8	1	854
USA	0	0	1	1	1	3	7	0	2	0	0	1	0	1	2	0	0	9	5	11	4	1	0	0	0	0	0	0	0	49
USSR	15	20	27	30	20	14	26	47	62	66	10	2	1	0	16	33	19	41	51	19	25	9	1	0	0	0	0	0	0	554
Other	0	0	0	4	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	1	1	1	3	1	1	1	0	0	0	17
Total	206	215	226	227	187	189	225	253	232	261	266	267	257	243	290	308	267	333	344	328	384	297	296	335	318	265	152	105	2	7278

World Ocean Database 1998
Profile Count (NODC/WDC-A/GODAR)

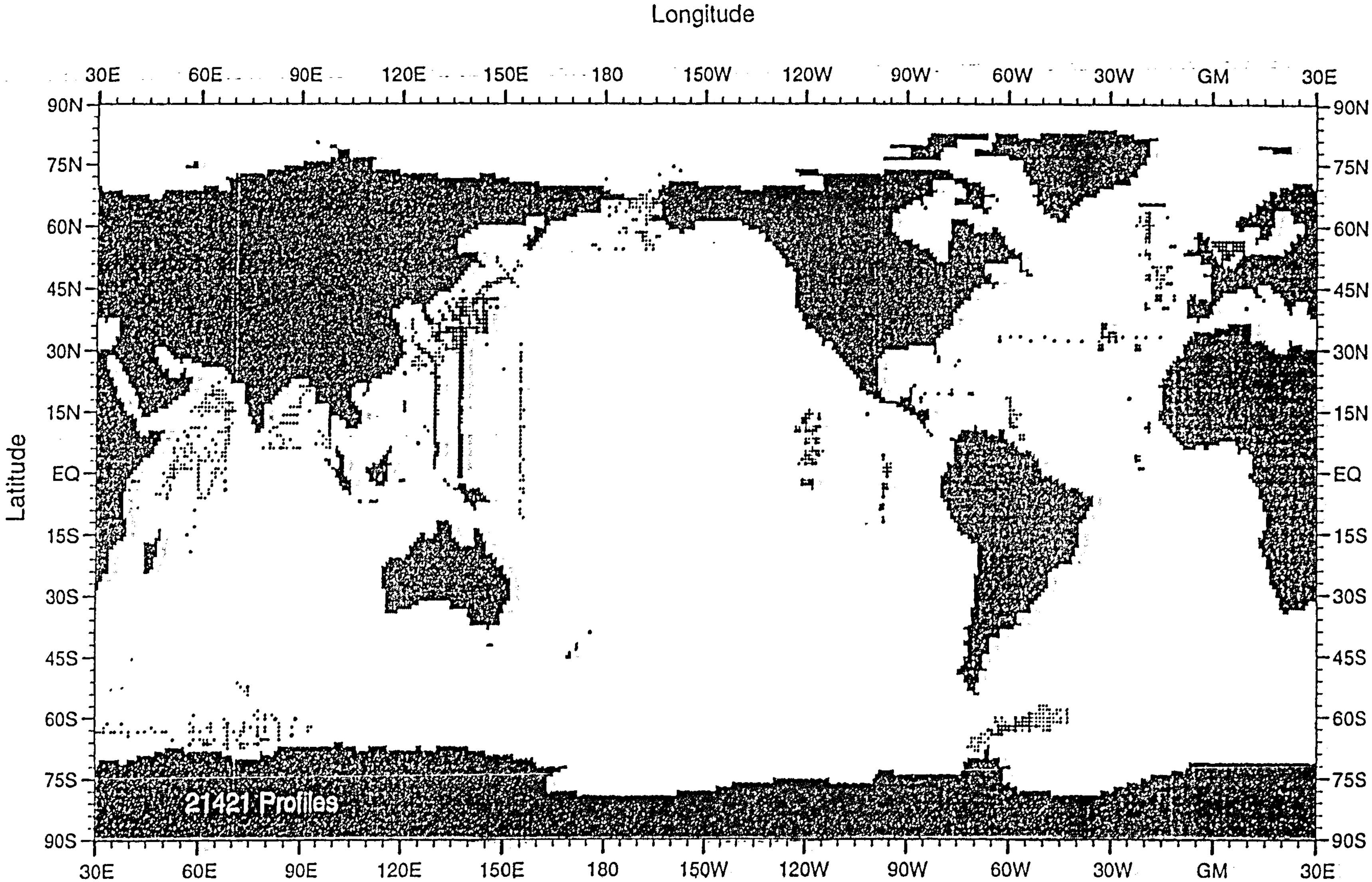




Distribution of Phytplankton in WOD98 (NODC/OCL)



Distribution of Zooplankton in WOD98 (NODC/OCL)



WOD98 BIO Profile Distribution

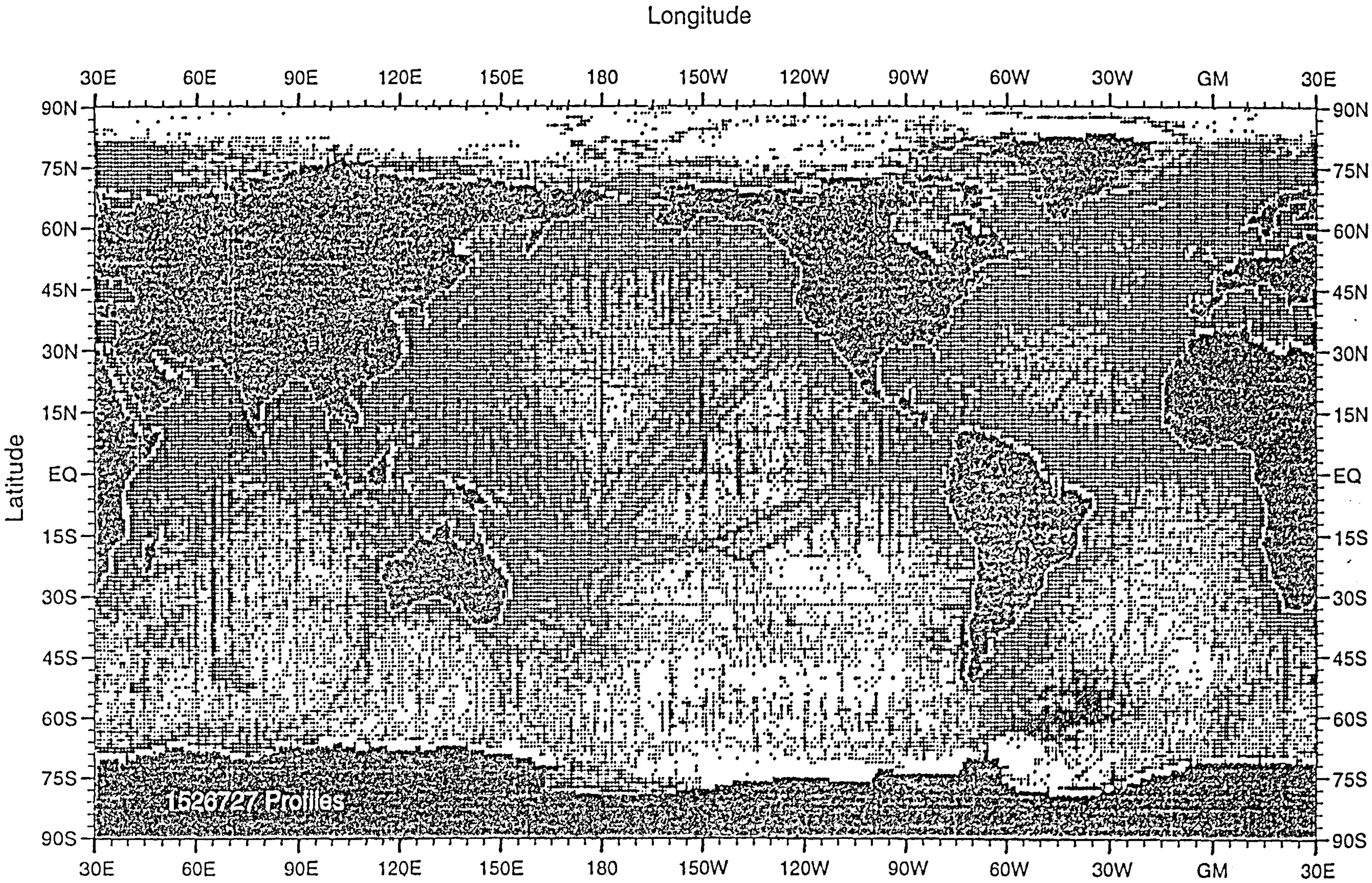


Fig. 5 Distribution of all profiles in the Ocean Station Data file

	How many requests for data, data products or information about data (.e. inventories, catalogues) have you handled in 1997?	*Summarise the sort of data/ information requested (e.g. waves, currents, XBT, CTD, data sets on CD-ROM, catalogues of data holdings)*	*Where do the data requests come from?*	*Do you have standard products available (e.g. CD-ROMs, statistical or gridded products)? If, yes, what are these products?*
ICES	approx. 90	Statistics or raw data	Universities, government, commercial, abroad	mainly individual processing
NODC/WDCA	Non-digital (catalogues):20065 Digital (CDs, etc.): 3338 Web pages: 152615	CD-ROMs, also customised data sets	general public, NOAA, other government, academia, private business, foreign	See list of CD-ROMs below
BODC	2694 (77 standard products, 1190 self service access, 527 ad hoc	Sea level, CTD, water bottle, surface hydrography, bathymetry, meteorology, geology, sediments	173 organisations for ad hoc requests, GEBCO - 150 organisations in 44 countries, OMEX 75 organisations in 15 countries.	GEBCO, North Sea Project, BOFS, OMEX, GLOSS Station Handbook, WOCE Sea Level CD-ROMs; UKDMAP, CMI, EDMED.
MEDS	330	tide/water levels, waves, drifting buoys, XBT, CTD, water bottle, GTSPP, National Energy Board Environmental data sets, etc.	Many organisations	GTSPP CD-ROM, NEB CD-ROM, No gridded products
DOD	137, but this is only requests which took more than 0.5 day to answer. Very different workloads (2 days to 2 months)	Over half request temperature and salinity	10% DOD, 45% Universities, 5% Government, 6% Commercial, 23% Foreign	From yearly Government Bulletin to the State of the Sea
SISMER	149	Geophysics 57; Physical/chemical 56; TOGA/WOCE XBT 36	IFREMER, public organisations, private sector, foreign	MEDATLAS CD-ROM; gridded data sets for IFREMER intranet and internet subscribers
IMDC	62	Digital bathymetry, waves, temperature and salinity, sea level, currents, long term nutrients	Universities (26%), state sponsored body (20%), research institution (17%), private companies (16%), government department (7%), private individual (5%), interest group (2%)	PIRATE prototype CD-ROMs (PROFILE and BENTHOS), Guideline Document on Electronic Data Publishing and gridded bathymetry
SMHI	150 from e-mail/letters; internal requests 5-25 per day	CTD and water bottle (O_2, nutrients, chlorophyll); inventories, time series, budget calculations, statistics, custom designed, cruise reports.	most internal and from government counties and universities; a growing number from abroad.	Statistical and gridded products; Cruise reports; data files suitable for export to commercial packages

	How many requests for data, data products or information about data (.e. inventories, catalogues) have you handled in 1997?	*Summarise the sort of data/ information requested (e.g. waves, currents, XBT, CTD, data sets on CD-ROM, catalogues of data holdings)*	*Where do the data requests come from?*	*Do you have standard products available (e.g. CD-ROMs, statistical or gridded products)? If, yes, what are these products?*
IMR	~1 per week; inside IMR there is direct access	mainly CTD, some currents	Own organisation; also universities, government, commercial and abroad	No, but can produce gridded data sets. Fixed station data on the Web
IEO	~50	15 sea level, 40 MEDATLAS, also temperature and salinity near coast	Own organisation, universities, other public bodies, commercial	MEDATLAS CD-ROM; sea level annual report
IH	Approx. 20	Wave data, currents, tides, CTD	Own organisation (currents and CTD), commercial organisations (currents and tides), abroad (Netherlands 'Teamwork Technology' for wave data and their analysis)	For wave data - statistical products. In general, catalogue of oceanographic data with spatial and temporal distribution for oceanographic cruises (water bottles, CTD, XBT and MBT, current meters, network of tide gauges, waves, meteorological and thermistor chains). We are working on a catalogue of geological data.
RDANH	Between 40 and 50	Sea level, currents, temperature and salinity. Data exchange with Danish Met. Institute and Danish Hydraulic Institute in near real time.	Own organisation (hydrgraphic department), universities, government and commercial organisations. Institutes in Germany and Sweden mainly.	No, data are extracted from database. Quarterly/annual reports include time series plots, which may be sufficient in some cases.
FRS	Not logged, but ~20-30	Inventories of hydrographic and current meter data; CTD data, current meter data; hydrographic, chemistry and productivity data	Own organisation, BODC, ICES, commercial organisations, other government bodies/agencies	Annual Cycles Working Paper with floppy disk (Program plus data)
CEFAS	from 2-3/month to 5/6 month	Temperature data from coastal programme for biologists, or North Sea/Irish Sea; current meter data	mostly CEFAS colleagues, but some universities; occasional requests for data reports from other CEFAS Labs.	'In-house' format current meter data and temperature, salinity and nutrients

	Can you always respond positively to requests (i.e. do you have the data requested?) or are you asked for data you do not hold?	*If you do not hold the data requested, what is your response? Can you refer the enquirer elsewhere? And if so, where do you usually refer them to?*	*How do you think your service could be improved?*
ICES			
NODC/WDCA	Try to refer enquirer elsewhere	Refer to proper person or institute - contact point plus Web address	Taking advantage of state of the art technology - placing data on-line, hot linking to data sites. Education of upper management
BODC	No, we are asked for data not held	Refer to ICES/WDCA or elsewhere either in the UK or abroad, as appropriate	Knowing who else has what quickly
MEDS	Act as a referral service for data not held	Refer to regional institutes and other relevant government departments	More data products and services on the Web. Central Web sites for Canada's marine environment. Contribute to ICES and use as referral.
DOD	Sometimes, requests passed to other NODCs and to ICES	Other NODCs and ICES, but enquirer may go there directly	German scientists want data on-line. DOD aim to respond to requests in 14 days, usually response is quicker. Digital data requested by Home Page on the Web.
SISMER	No, sometimes asked for data not held	Try to send to right place, i.e. IFREMER, SHOM, International Current Meter Inventory, other data centres, Web addresses. Also use World Ocean Atlas CD-ROMs	Inventories like the current meter inventory useful; access to data via the Web; links with other data centres Home Pages.
IMDC	Cannot always respond to requests because we may not hold the data or it is restricted, other commitments take priority we need to charge for time and are limited by resources	Clients are referred to appropriate sources to facilitate their needs	Needs to be made part of the Core activity; surveys have been carried out to consolidate user requirements
SMHI	Yes to 98% of requests	Offer to help customer - contact relevant institutes and present problem together with customer.	More use of internet; by working closely with customers
IMR	Can help almost everyone who needs data in Norwegian waters	Sometimes to ICES and other NODCs.	More products on the Web
IEO	Not always, but most of the time	For sea level data refer to PSMSL or other Spanish Institutes; Hydrographic data from WDCA	Compiling more data from institute; prepare catalogues for distribution; develop software to manage data more easily; preparing products for electronic media distribution

	Can you always respond positively to requests (i.e. do you have the data requested?) or are you asked for data you do not hold?	*If you do not hold the data requested, what is your response? Can you refer the enquirer elsewhere? And if so, where do you usually refer them to?*	*How do you think your service could be improved?*
IH	Sometimes requests are made for current meter data where we know moorings were deployed but we do not have the data	If we do not have what is requested and the enquirer is from within IH, we will suggest where to go.	It could be improved, if we can work and establish links outside IH.
RDANH	Data may be requested for areas where no data are held. Gaps in the data are a problem. The answer is 'no, not always'.	Give a best estimate from a neighbouring station. Danish Met. Institute may have some relevant data. Swedish (SMHI) and German (BSH) colleagues also may have relevant data.	Missing data (transmission failure, sensor breakdown, etc.) are a problem.
FRS	if we have data, yes	If data not held, refer to BODC and/ or ICES, depending on requestors requirements	All requests usually met within 48 hours, unless Data Manager is at sea, then delay could be 3 weeks. A deputy would improve service, but with the small number of requests this is unlikely
CEFAS	Yes, can usually offer something	Contact BODC (especially for current meter data) and ICES for water sample observations	Difficult to keep track of all the data available. Useful to have a way of identifying what data are available from BODC, ICES, etc. Current meter inventory was very useful.

Annex 7 Outline of presentation on parameter code tables/data dictionaries

Parameters - A discussion

Orla Ni Cheileachair
The Irish Marine Data Centre

MARINE INSTITUTE

Today's Talk

Quick Overview of :

- IOC ROSCOP (CSR) parameters
- BODC OMEX (JGOPS) Parameters
- BODC EDMED parameters
- IFREMER MATER Parameters
- GF3 Parameters
- IMDC PIRATE Parameters
- IMDC parameters for the MAST III projects, CANIGO, ENAM & BENGAL

MARINE INSTITUTE

Today's Talk

Why ?

- To highlight some issues we have encountered in the weird and wonderful world of 'parameters'.
- To prevent re-inventing of the wheel
- To look at possibilities of increasing comparability of parameters across Data Centres

MARINE INSTITUTE

IOC ROSCOP (CSR) Parameters

- 2 level hierarchy:
- Multidisciplinary
- Big bucket parameter group headings
- Specific Parameters listed per group
- Category assigned a letter and parameter a number = 3 byte code
- Physical Oceanography = H, H09 = Water bottle stations
- Equipment included as a 'parameter' in addition to naming what was actual measured, e.g. CTD, floats

MARINE INSTITUTE

BODC EDMED Parameters

- EDMED Parameters:
  - Multidisciplinary
  - Big bucket parameter group headings
  - Specific Parameters listed per group
  - Specific parameter code ?
  - 2 level hierarchy useful for searching
  - Equipment included as a 'parameter' in addition to naming what was actual measured

MARINE INSTITUTE

BODC OMEX Parameters (JGOFS)

- Data Dictionary:
  - 2 level hierarchy
  - 8 byte parameter code which identifies method and state
  - Units per parameter
  - Instrument included as a 'parameter'

MARINE INSTITUTE

BODC OMEX Parameters (JGOFS)

- Parameter Hierarchy:
  - Parameter Category, .e.g. *Curr* = Currents
  - 8 byte parameter code, e.g.
    - LCDA - 4 bytes = Horizontal Current Direction
    - EL - current direction (Eulerian Method)
    - 02 - Channel 2
  - *LCDAEL02 = Full 8 byte parameter code belongs to the parameter group 'Currents'*

BODC OMEX Parameters (JGOFS)

- **Parameter Code:**
- Parameter Category, .e.g. CNPS = C, N, P, Si data including nutrients
- 8 byte parameter code, e.g. PHOSAAD1
  - bytes 1-4 indicate the parameter name
  - bytes 5-8 indicate method plus state or sampling 'compartment'
  - PHOS - 4 bytes = Phosphate
  - AAD1 - bytes 5-8 = dissolved phosphate …easured using colorometric …toanalysis (GFF filtered)

BODC OMEX Parameters (JGOFS)

- Parameter Units:
  - Category = CNPS
  - Parameter = PHOS
  - Abbreviated name = PO4
  - Parameter name = Phosphate
  - Units UPOX = Micromoles/litre

MATER Data Manual - Parameter Inventory

- MTP II/MATER parameters:
  - 3 level hierarchy for 12 main data sets
  - 1) In-situ Physics; 2) Nutrients; 3) Metals; 4) Chlorofluoro-Carbons, 5) Radio-isotopes; 6) Biogenic Major Elements; 7) Biogenic Minor Elements; 8) Pigments, sugars, amino acids; 9) Primary production; 10) Microbiology; 11) Meso & Macro fauna; 12) Remote Sensing.

MATER Data Manual - Parameter Inventory

- Each of the 12 data sets (excluding Remote Sensing) are split into sub groups comprising of specific parameters
  - Eg. Level 1 = In-situ Physics
    - Level 2 = CTD Profiles
      - Level 3 = Pressure
        - Temperature
        - PSAL

MATER Parameter Code

- 4 byte Code unique to the element + unit
- Uses IOC/GF3 where available or
- Own SISMER code, or in other cases:
  - First 2 letters indicate main element *(if only 1 letter like C for Carbon, the letter is duplicated, eg. CC)*
  - W, P, F, S, O indicate State, e.g. P = particles in the water column, e.g. ICCP - particulate inorganic carbon
  - Last of first letter is related to <total> or …cies like <A, B> where relevant

MATER Parameters

- 4 byte parameter code (International or SISMER defined)
- Parameter name - unique & reflects sampling compartment (e.g. water, particles, sediment, biota, etc); 30 bytes
- Parameter unit = SI units (if another unit is used, it is considered another parameter); 30 bytes
  - TCCF = Settling Particulate Total Carbon Flux per 10-6 KG M-2 DAY-1

GF3 Parameters

- Grouped under 10 headings labelled from 7a-7j:
  - 7a) General Purpose; 7b) Date & Time within Day; 7c) Time & Frequency; 7d) Position & Navigation, 7e) Physical Oceanography; 7f) Waves; 7g) Meteorology; 7h) Geophysics; 7i) Chemistry; 7j) Special Purpose.

GF3 Parameter Codes

- 8 byte parameter code = PPPPKMMS
- PPPP = parameter identifier
- K = indicates whether the parameter, method & unit is standard or user defined and varies from 2, 4-7
- MM = method used to measure parameter (set to XX when unspecified)
- S = compartment in which parameter was measured (varies from a-j, n, x)

GF3 Parameters

- PPPP KMMS Name Units Ref
- CPHL 7XXD Chlorophyll-a mg/m^{-3} 7i-chem

- CPHL = Chlorophyll-a
- 7 = the parameter, method and unit are standard
- XX = method is unspecified
- D = sample measured in the hydrosphere

- Units = Micrograms of chlorophyll-a per cubic decimetre of water at 20 deg C.
- Ref = Parameter belongs to Chemistry group

IMDC Parameters

- IMDC Activities:
  - PIRATE system (RDBMS)
  - Data Tracking System (RDBMS)
    - Data model of relationship between cruise based data (gears and parameters)

IMDC Parameter Codes

- PIRATE system:
  - Multidisciplinary parameters including ROSCOP parameters
  - Specifically adapted to the MAST II projects BENTHOS and PROFILE
  - Internal coding system for parameters
  - 3 Level Hierarchical structure
  - Equipment included as a 'parameter' in addition to naming what was actual measured

PROFILE PIRATE system

- Parameter Category Code

1 Physical Oceanography
2 Chemical Oceanography
3 Biology
4 Geology and Geophysics
5 Contamination
6 Meteorology
7 Modelling (Physical Oceanography)
8 Biological Oceanography

NUMERIC CODING System

Category	Code	Description
1	100100	CTD Profiles
1	100101	–Temperature
1	100102	–Salinity
1	100103	–Conductivity
1	100104	–Density
1	100105	–Sigma-t

- *CTD profiles are parameter number 001 in the Physical Oceanography category (1) with sub-parameters 01-05 of temperature, salinity, conductivity, etc, etc*

BENTHIC PIRATE PARAMETER CODING System

ParameterCategory	Code		Mega	Macro	Meio	Micro	Nanofauna
3	300200	Community Ecology	Yes	Yes	Yes	Yes	No
3	300201	— Biomass	Yes	Yes	Yes	Yes	No
3	300202	— Abundance/Density	Yes	Yes	Yes	Yes	No
3	300203	— No. of Species	Yes	Yes	Yes	Yes	No
3	300204	— Diversity	Yes	Yes	Yes	Yes	No
3	300205	— Species Richness	Yes	Yes	Yes	Yes	No
3	300206	— Species Equitability	Yes	Yes	Yes	Yes	No
3	300207	— Other Indices	Yes	Yes	Yes	Yes	No
3	300208	— Size Spectra	Yes	Yes	Yes	Yes	No
3	300209	— Trophic Groups	Yes	Yes	Yes	Yes	No

In addition to the codes, the size class was also flagged for each parameter

IMDC Data Tracking System

- WWW System for viewing the parameters measured per station on a cruise and the institute responsible for each parameter
- Data (summary cruise information, positional information, gears, parameters and participating institutes) stored in a relational database.
- Observed parameters are related to the sampling instrument as defined in the cruise report (1 instrument to many parameters relationship)

Cruise Data Tracking System

- Contains parameters specifically for the CANIGO, ENAM II and BENGAL MAST III projects plus ROSCOP parameters
- No coding system has been defined yet
- Hierarchical approach planned
  - Big buckets such as EDMED parameters (e.g. Physical Oceanograpy, Chemical Oceanography, etc) for easier searching
  - Parameter code which allows another level of parameters grouping similar to PIRATE which would again facilitate easier searching ?

Main Parameter Issues

- Hierarchical Approach - facilitates searching but needs:
  - standard 'big bucket' categories
  - standard levels of grouping (as opposed to having Fatty Acids & Geophysics at same level)
- Confusion between instruments & parameters
  - Need to define what IS a 'parameter'
  - Require definition of relationship of parameter to measurement method , e.g. relationship of parameter to sampling instrument in DTS
  - he definition of this relationship varies epending on the stage of data generation

Main Parameter Issues

- Parameter Codes
  - 8 byte coding systems generally in place ?
  - 4 byte parameter identifier or name 'common' but not standard across Data Centres
  - Given multidisciplinary nature, only certain parameters have standard codes
- Generally, code contains identifier to the measurement method and state or 'compartment' in which the parameter was sampled but:
  - 'Compartment' identifiers are not standardised
  - Measurement methods will always be in a state of flux ?

MARINE INSTITUTE

Main Parameter Issues

- Parameter Units
  - Only certain parameters have standard units
  - Units will vary according to the science and measurement method
  - SI units /international classification systems are often not adopted by the scientists submitting data

MARINE INSTITUTE

What is critical ?

- Standardisation of parameters/data dictionaries between Data Centres to ensure easy & comparable retrieval of data ?
- Hierarchical structure for easy searching of multidisciplinary parameters ?
- Proper definition of parameters to remove confusion with instrument and measurement
- Some means of updating new parameters to ensure standardisation across Data Centres

MARINE INSTITUTE

What to do next ?

- Standardise on big bucket groupings which will be compatible with EDMED type activity?
- Agree on the hierarchical level of parameter groupings?
- Define a parameter only by what is being measured - distinct from instrument/gear
- Agree/adopt a system for coding

MARINE INSTITUTE

Parameter Codes ?

- Standardise on parameter identifiers (4 byte name, e.g. phos)
- Standardise on 'compartment' identifiers
- Methods - keep in but with the option of 'not specifying' - Needs more thought ?
- Include grouping levels within code as in PIRATE ? (Code could then be > than 8 bytes)

MARINE INSTITUTE

Parameter Units ?

- Units - standardise where possible
- decide approach if parameter unit differs

MARINE INSTITUTE

Annex 8 Shipboard Ocean Data Information - ODIN

Presented at the Annual Meeting of the
ICES Working Group on Marine Data Management
April 21, 1998

Anthony W. Isenor
Bedford Institute of Oceanography
P.O. Box 1006 Dartmouth, Nova Scotia
Canada, B2Y 4A2
isenora@mar.dfo-mpo.gc.ca

INTRODUCTION

Over the past decade, oceanographic institutions, such as the Bedford Institute of Oceanography (BIO), have become involved in data collection and research efforts that contribute to global research programs. The delivery of these data to the international community has brought attention to collection and management problems dealing with metadata details, sensor tracking, etc. To produce datasets useful to the global community, the datasets require: 1) at-source data entry, 2) established links between the data and metadata and, 3) management of the data as opposed to simple storage.

The at-source data entry is important to capture details of the scientific activity. Often, these details are either not recorded or are recorded in personal notes to which access is limited. By providing a means to capture this information, we acquire a more accurate and complete representation of the events.

The established links between the data and metadata help provide integrity for all collected information. Such links ensure that the entered data comply with known rules and are consistent with previously recorded data within the system.

Finally, management of the data is crucial for the distribution of the most complete and accurate set of information to users. Such management goes beyond the simple storage of the data, to include the automatic directing of data and metadata to logical and interrelated storage locations within a single database.

The following paper deals with a data management application for use onboard research ships. The system capabilities and initial system testing are reviewed.

APPLICATION

The application development was based on the specifications defined in a functional model. The details of window-window and window-database interactions, as well as window appearance were included within the specifications.

The application was developed in Powerbuilder Desktop® Version 5. Powerbuilder is a 4GL application development tool. The application was developed for Windows 3.1 but has recently been upgraded to Windows 95. Powerbuilder is database independent, thus providing the developer with a wide choice of possible databases.

Powerbuilder uses object oriented programming techniques which allow easier updates to the application code. A modular application component design also organizes the objects into logical groups, again making maintenance easier. For a review of object oriented terminology, see Hendee (1994).

Hardware requirements for the installation of ODIN include 20 Mbytes of free space, Windows 95 operating system on a 486 or better, with at least 8 Mbytes of RAM.

There are also functional requirements for the use of ODIN. These requirements represent rules of conducting the scientific activities. The two functional requirements of ODIN are unique numeric identifiers for the cruise and each individual water sample.

It is common for an oceanographic lab to reference individual cruises with numeric identifiers. Referencing individual water samples with unique numeric identifiers is also common, however, the methods used to determine the identifiers are varied. BIO references water samples using sequential 6 digit identifiers. Other labs commonly combine the station number and rosette position to produce a unique identifier for the water sample. Either method is acceptable to ODIN. However, ODIN does default to sequential identifiers, automatically computing the identifiers for the next planned operation.

COMPONENTS

ODIN has been structured into five main components that include cruise planning, personnel management, instrument configuration, scientific operations and water sample management.

The cruise planning component allows the user to construct a specific cruise dataset using the unique cruise number. The user then assigns personnel to the cruise and individuals to duties and watch periods. The user can specify the sampling order for water samples and can plan the scientific operations (for example CTD casts, XBT drops, moorings, floats, drifters, etc.). The details of the water sampling can be defined including the number and depth of bottle trips at each station, and the type and number of samples to be drawn from each rosette bottle. The detailed planning of scientific operations is optional to allow the system the flexibility to incorporate spontaneous operations.

The instrument configuration component allows the creation of instrument packages. The package is a term used to refer to any grouping of physical objects. Typically, the package represents a logical group of instruments that support a particular science related activity. Having defined the package, the user may then use this definition within the system to identify the instruments involved in a particular data collection operation.

The scientific operations component represents that part of the system used to track the details of the individual shipboard operations. These operations include and extend beyond the full suite of WOCE cast types to include non-science related operations. The user has complete control over the detail of the tracked operations and may include such things as steaming time and navigation logging. Examples of more traditional oceanographic operations include CTD casts (with or without Lowered ADCP), XBT drops, moorings, floats, drifters, etc.

All operations are given a unique operation number and thereafter are identified by this number. Operations are tracked backwards, if possible, to information stored within the planning component. When available, planning component information is used during the completion of the operation.

Identifying a particular stage of an operation (for example the beginning or end of a rosette cast) begins a "wizard" series of screens that lead the user through the required information for the particular operation type.

For operations involving the CTD, ODIN is capable of displaying the real-time Seabird CTD pressure. The data stream from the Seabird deck unit to one PC running ODIN, provides the necessary input to compute instantaneous pressure value. The computed value is then placed in the database and made available to other PCs running ODIN. The pressure display is useful to the winch operator for locating pressure surfaces for bottle trips and also for staff waiting for the package to arrive on deck.

ODIN also has the ability to read and decode NMEA navigation strings. Using a NMEA serial port feed into a PC running ODIN, the system can decode and display the navigation. The navigation may be accessed automatically through the wizard screen. Alternatively, a unique scientific operation called navigation logging may be initiated to store navigation data at a predefined interval to the database.

The water sample component tracks each collected water sample and associated attributes of the sample. Water samples may be collected during any assigned operation including rosette casts, pumping system, etc. For rosette casts, the start sampling time for each parameter is stored to allow an estimate of the time on deck before sampling. The user has the ability to assign values to parameters directly within the application, or import values using standard text files. Missing samples can be automatically identified and quality flags may be assigned to any sample. All assigned quality flags are stored with time stamps to provide a history of the quality control. Calibration criteria for individual parameters may also be included.

Throughout the system the user has the flexibility to include notes dealing with the cruise, individual operations, instruments, water samples, rosette bottles, or thermometers. The notes are all time and personnel stamped.

Users seeking information may browse all aspects of the system. Users may review the planning component to develop sampling strategies, review instrument notes to determine reasons for sensor changes, or rosette bottle attributes to identify leakers. The data may be exported via WOCE station summary reports, and in the future, WOCE SEA files. Alternatively, users may wish to copy and work directly with the MicroSoft Access® database tables.

OPERATIONAL USE

The system has been field tested on a 1997 cruise to the Labrador Sea. The test resulted in numerous modifications to the application although none were required to the database structure. Some expansion of ODINs functional capabilities were noted, but will not be implemented for the next field test. A solid software foundation must be established before adding features. ODIN will undergo a second field test in June 1998.

REFERENCES

Hendee, James C. 1994. Object-oriented Database Management Systems and Their Application to Oceanography, Earth System Monitor, Vol. 4, No. 4, June 1994.

Annex 9 Recommendations

Proposed Agenda for next year's meeting:

The Working Group on Marine Data Management will meet in Ottawa, Canada from 3 - 6 May 1999 to:

a) Assess the last five years data (1994-1998) sent to ICES by each member country, identify problems and suggest solutions;

b) Review progress in the implementation of IOC's Global Oceanographic Data Archaeology and Rescue (GODAR) Project in each member country, including consideration of biological oceanographic data types;

c) Quantitatively analyse the minimum requirements for quality assurance of oceanographic data;

d) Develop guidelines for the quality assurance and data management of nutrient and oxygen data in cooperation with the MCWG;

e) Report on the development of World Wide Web pages and links between them within member countries;

f) Investigate and evaluate the data dictionaries available to the marine science community, including an analysis of the parameter code list used for the IOC Cruise Summary Report, and produce an improved and updated set of codes.

Justifications:

a) Although the data received by ICES over the last two years has been encouraging, there is still a large amount of data outstanding especially nutrient data and data from global projects. This item should act as encouragement to Member Countries to supply the ICES Oceanographic Data Centre with data in a timely manner.

b) Much data have been recovered by the five year GODAR project, 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; investigations suggest that much biological data is available within ICES Member Countries. This item serves to help quantify the data and associated documentation available, and their status.

c) There is a need for simple guidelines for those collecting, processing and quality assuring data. Having reviewed those guidelines and manuals presently available, and produced a set of guidelines for moored current meter, CTD, shipborne ADCP and SeaSoar/Batfish data, other data types will now be considered (e.g. moored ADCP, drifting buoys, XBT and sea level) and guidelines developed and updated.

d) The MCWG have been reviewing quality assessment procedures for nutrient and oxygen data. Following on from this, the MCWG and WGMDM will jointly develop guidelines. The existence of written guidelines has distinct advantages. It shows laboratories reporting to the ICES data bank how important it is to apply quality control procedures on the data, and it will provide ICES with data sets which are easier to handle and which have a properly documented QC history behind them.

e) This is an opportunity to exploit developments within the Internet and raise the profile of the data centres within in the ICES community. In particular, collaboration on data products will be investigated and the WGMDM pages will be further developed.

f) A number of Data Dictionaries, each covering a wide range of parameters, have been developed by the oceanographic community. Last year, these were critically reviewed by the WGMDM. An inter-sessional sub-group will continue this and suggest the appropriate hierarchical structure and standardistation at the category level. Data flagging schemes will also be addressed.