Intergovernmental
Oceanographic
Commission

Manuals and Guides 5

GUIDE FOR ESTABLISHING A NATIONAL OCEANOGRAPHIC DATA CENTRE (Revised version)

1997 UNESCO

TABLE OF CONTENTS

GUIDE FOR ESTABLISHING A NATIONAL OCEANOGRAPHIC DATA CENTRE

INTRODUCTION

This guide has been prepared to assist Member States of Intergovernmental Oceanographic Commission (IOC) in establishing and operating a national oceanographic data centre (NODC). Since the early 1960s, approximately 55 Member States of IOC have established oceanographic data centres or designated a national agency as responsible for international oceanographic data and information exchange. These data centres and designated national agencies (DNAs) support both national and international clients with oceanographic data and information services. The map in Figure 1 shows the distribution of these centres around the world.

This guide provides information on:

- a typical mission for a national oceanographic data centre (NODC);
- various organizational models for an NODC;
- steps in establishing the NODC;
- development of a client base;
- typical computer technologies useful to an NODC;
- data management planning;
- development of products and services;
- international data and information exchange.

The Manual is divided into two parts. Part I deals with the national aspects of an oceanographic data centre. Part II deals with the international system of the IOC, the International Council of Scientific Unions (ICSU), and the World Meteorological Organization (WMO). This separation has been made so as to clearly separate the national and international aspects of a data centre programme.

PART I. NATIONAL SERVICES

Section 1 discusses the development of a mission for a new NODC being created. It describes in general the types of functions a data centre will carry out. It is not intended as a complete list. Depending on clients and circumstances, the centre may do other things.

Section 2 discusses the options of organizing the centre as either a more traditional centralized location, or a distributed centre with the databases maintained in other organizations, probably those that collected the data. In practice, the organization of the centre often turns out to be a combination of centralized and distributed databases.

Section 3 lays out an administrative approach to establishing the centre. It does not present a rigorous process that follows a specific administrative model. The approach is rather one of identifying particular steps that will likely be part of any administrative model and which will provide a good base of information from which decisions on the role and operation of the centre can be made.

Section 4 provides a view of potential clients for a new data centre. Any particular NODC may not have all of these clients. The list is a composite of the clients to be found in many IOC Member States. It should be considered while finalizing the mission of the centre.

Section 5 provides some examples of the computer technologies that might be acquired for a new NODC and how they might be used. This Section is intended to demonstrate how the various technologies would be used in a data centre. It should not be used to replace a careful study of the technology requirements of the centre.

Section 6 discusses data management planning. It is important that data collection activities be examined and separated into similar types. Appropriate plans can then be developed and agreed. This Section gives an example to illustrate the elements of a data management plan.

Section 7 discusses data products and services that can be provided by an NODC. It makes the point that data centre products and services should be planned with specific clients and their needs in mind.

Section 8 discusses the very important subject of quality control of the data holdings. Quality control is one of the most important aspects of the work of a data centre. If data quality is not evaluated and documented, users will have little confidence in the data centre.

1. MISSION OF A NATIONAL OCEANOGRAPHIC DATA CENTRE

1.1 THE CONVENTIONAL ROLE OF THE NODC

The basic mission of an NODC as originally conceived was to acquire data, primarily from scientists who collected it for their own research projects. The quality of the data was assessed by various simple tests and procedures. The data were then archived for use by "secondary users".[1]

Generally, the data were not submitted to the data centre until the scientist had finished with it unless the centre were assisting him with the processing. When assisting with the processing, the centre offered not to distribute the data until the scientist was finished with it. Research projects were usually regional or local in scope.

An NODC would provide as a minimum:

- Copies of all or part of its data holdings in the format in which the data were received.
- Inventories of its holdings so that users could select data of interest.
- Provision of sufficient documentation about the data so the user can apply it to his problem with confidence of its value.
- Referral of the user to sources of additional information about the data or to other national and international sources for oceanographic data not stored in the NODC.
- Participating in international oceanographic data and information exchange.

1.2 THE CONTEMPORARY ROLE OF THE NODC

The present research picture has changed drastically. Research projects are now ocean basin and global in scope. Technological advances in data collection by moored and drifting instruments and satellites has increased the volume of data collected by orders of magnitude. More and more data are flowing in operational time-frames from *in situ* data collection systems and satellites. The data are being supplied to users ranging from marine weather forecasters to ship operators, from agricultural forecasters to insurance industry workers, and from numerical modellers to managers of the international science experiments.

Thus, the present day NODC may find itself doing many of the following functions:

- Receiving data from researchers in the traditional manner, performing quality control, archiving and disseminating it on request.
- Receiving data from buoys, ships and satellites via electronic networks on a daily basis, processing the data immediately, and providing outputs to various research and engineering users, forecasters, experiment managers, or to other centres participating in the data management plan for the data in question.
- Reporting the results of quality control directly to data collectors as part of the quality assurance module for the system.

[0] The "secondary user" was considered to be a scientist or engineer working on another project to which the data could be applied.

Figure 1

- Participating in the development of data management plans and establishing systems to support major experiments, monitoring systems, fisheries advisory systems.

- Processing and publishing data on the Internet and on CD-ROMs.

- Operating Responsible National Oceanographic Data Centres (RNODCs) or World Data Centres (WDCs) on behalf of the international science community (see Part II).

- Operating as a data assembly and quality control centre for part of an international science experiment (see Part II).

- Developing standards for the data documentation and processing.

- Publishing statistical studies and atlases of oceanographic variables.

- Participating in international oceanographic data and information exchange.

- Assisting in the development of data exchange standards and procedures.

- Assisting with data management aspects of global or regional programmes or IODE pilot projects.

1.3 MAKING THE CHOICE

The previous few paragraphs describe a wide range of potential activities for an NODC. Before an NODC is established it is vital to clearly determine its mission and the level of activity to be achieved. The mission will lie somewhere between the two views above. It must meet the requirements of the Member State and fit in with the organization of oceanography between the various scientific and government labs in the country.

An NODC must provide a good service as a national centre before it can provide one as an international centre. If a centre does not provide a good quality national service, it will not receive the co-operation of its data collectors and users and will fail. An NODC should not be established to serve a perceived international need alone. The centre should generally not take on responsibilities for international datasets that are of no interest or are not collected nationally. This point should be considered when developing the mission for the NODC.

1.4 NODCs AND INFORMATION

The term “information” as used here covers a wide variety of forms ranging from data products such as statistical analyses and maps of the distribution of variables over the ocean, to the published scientific literature. Oceanographic data centres tend to produce the first types of information; for example, the analyses and maps. Libraries are usually the providers of books and scientific journals. Thus the term “information” as used in reference to data centres in this guide will usually refer to analyses, maps, tables, etc. Usually, oceanographic data centres that produce such information are more successful than centres that only provide digital datasets.

There is a growing benefit in closer co-operation between oceanographic data centres and libraries that deal in the ocean science and engineering fields. With the advent of the Internet, CD-ROMs, and the World Wide Web (WWW), information is increasingly computer based as opposed to paper based. Electronic information can be published at a lower cost and circulated quickly. Both the data centres and the libraries can benefit by sharing expertise in the preparation and distribution of electronic information and data. The data centre generates the products and provides computer and communications expertise. The library provides expertise in distribution, well established reference systems for its holdings, and an existing broad base of clients.

2. ORGANIZATIONAL CONSIDERATIONS FOR AN NODC

When considering the establishment of an NODC, it is beneficial to look at the options of a centralized data centre offering a referral service to data not held in the centre versus a distributed centre.

2.1 A CENTRALIZED DATA CENTRE

Before the advent of high speed data communications and database management systems that were designed to operate over networks the only choice for most Member States of IOC was to establish a centralized data centre. The centre would acquire copies of data from the data collectors and build databases, processing and quality control systems, and data retrieval systems to provide services to users directly from the centre.

In most Member States there would be some data sets that were not appropriate for duplication in the data centre. For example, some research datasets are highly specialized and may consist of the data from a single experiment. It would not be productive to establish databases and build software systems in the NODC for such data that might be of interest only to one or two other researchers. Instead, the data centre would maintain an inventory of such datasets. The data centre could search this inventory and "refer" interested users to the organization or person holding the data.

This type of data centre is simpler in concept, and can operate with more autonomy that a distributed centre that must depend on the co-operation of other organizations. This type of data centre does not have to acquire and operate the complex communications networks an database management systems required for distributed databases. For a new centre it may well be appropriate to begin with a centralized type of operation and then gradually include some distributed databases if the situation warrants it.

It is worth remembering that the complexity of the data centre operation and the sophistication of its data management systems are not a measure of the success of the centre. If a data centre is acquiring national ocean data and saving it from being lost, it is performing a valuable service and has taken a significant step towards being a successful NODC. As in the development of any new service, it is usually better to start off doing the most basic and important operations well. Once the basic systems are in place the centre can build on their success to improve its services to users.

2.2 A DISTRIBUTED DATA CENTRE

Whether a data centre has opted for a distributed model from the beginning or not, there are a number of advantages of distributed data management that are worth recording. As the centre develops closer working relationships with its data collectors and suppliers, it may be advantageous to implement distributed data management for some data types.

The data communications technology that is now available permits the creation and maintenance of distributed databases. This in turn allows an organization flexibility in how a data centre is organized. It is no longer necessary to bring all the data and expertise that the centre needs to a single location. Various parts of the database and the provision of services from the databases can be distributed.

The choice of whether to choose a distributed model will be based on a variety of considerations. One major considerations is the distribution of oceanographic research and expertise in the Member State or even the region. A number of Member States could choose to establish a centre to serve a region. Another consideration is the data types that will be collected by the centre. Some data types are more difficult to interpret than others and may require that the databases be located in the scientific centres that are collecting them.

As discussed above, the distributed model has several potential advantages. The first advantage is that certain database tables can be maintained in the location where the expertise exists to maintain them. For example, if the NODC is to maintain a database table of taxonomic protocols, this table could be maintained in a location where the national expertise is located. The second advantage is that the work required to maintain a distributed database can be distributed with benefits both to the data centre and the partner organization.

Benefits of a distributed data centre will include savings in operating costs and in salary costs by not having to duplicate specific scientific expertise. For example, if a data centre has mainly physical oceanographic expertise, then through partnerships with other organizations it can acquire the necessary expertise and other staff support to manage chemical or biological data. Of course this sort of arrangement is only possible if the partner organization realizes tangible benefits from the arrangement.

The tangible benefits that a partner organization can receive from a data centre would include such items as computer software, data management expertise, improved standardization, and a direct connection to other national and international oceanographic data. The data centre should be in a position to assist the organization in making data management plans for national and international experiments, in developing and implementing data quality control systems, in setting standards for data documentation (meta data), archival, and exchange, and in distributing data to the relevant national and international systems.

Thus as the development of an NODC proceeds over the years, it is important to consider the new centre as an integral part of the national or regional system. The organization and functions should be planned to take advantage of expertise and the work already being done. It is important to consider all opportunities for mutually beneficial partnerships with existing organizations collecting or using oceanographic data. Missions of existing organizations in the oceanographic field should be examined. The international responsibilities for exchange of data and information should be reviewed. Then a plan can be developed that will achieve benefits for all, while providing a broader and better range of services to users. In this manner the benefits and services that a data centre provides need not be limited by the expertise and resources available within the centre itself.

Many NODCs now have World Wide Web pages that describe their activities. Some have included a mission statement. This information may be useful to someone developing a mission statement for a new NODC. These national Web sites can be reached through the IODE Homepage at the address given in Annex B.

3. STEPS IN ESTABLISHING AN NODC

Two of the steps in establishing an NODC have already been discussed. Section 1 has described the development of a mission for the centre. Section 2 has discussed how a centre can be integrated into the oceanographic programme in a Member State or region through partnerships and distributed responsibilities. Of course these two steps are not taken sequentially. The mission of the centre will depend on whether these mutually beneficial partnerships can be established. The question is how to start off the process.

The most important requirement in establishing an NODC is to secure the support and co-operation of the oceanographic and in some cases meteorological organizations that collect and use oceanographic data. Without this cooperation the new centre will have great difficulties in acquiring data for its databases and will lose its first and most natural group of clients, i.e. the ones who supply the data. A centre that is designed with the collaboration of the collectors of the data will be off to a much better start.

There is always a danger that the new centre will be seen as an additional burden on the data collectors. They may not see benefits for themselves while having to do additional work to supply the centre with data. There is no universal solution to this problem. There are, however, some initiatives that can be taken to show mutual benefits:

- The centre should plan and demonstrate that it will have data management and other related expertise that is more specialized and focussed than that available to the data collectors. Thus, the data collectors can expect to receive benefits from working with the centre. The centre must then establish this expertise by recruitment or training and maintain it.

- The centre must also relieve the collecting organizations of some existing responsibilities. For example, if the organizations are having to deal with the international systems as well as local requests for data, the centre can take on that work for them.

- The centre must demonstrate that it can provide information on standards and on hardware and software technologies that will be useful to the collectors in improving their own data management.

- The centre must demonstrate that it will not be a roadblock in the data flow, but will improve efficiency.

An important step in the process is to build confidence that the centre will be established to serve, that it will have the expertise to fill a leading role in national oceanographic data management, and that it will work in a co-operative manner to ensure that the organizations that supply data will receive commensurate benefits for the work involved in supplying data to the centre.

Given the above ideas and principles the following steps are suggested as an outline on proceeding to establish an NODC.

- **Step 1**

  Recruit a team of interested parties (including potential clients and partners) to propose a mission and organizational model for the centre. This team should examine the missions of the organizations in the Member State or region that collect or use oceanographic data. Some of these organizations may be examined as a class (e.g., coastal engineering firms designing shore-based facilities of various sorts). This examination will also provide information on the potential client base. At this stage it is important to ensure that effective consultation occurs with all potential stakeholders including all levels of government (local, state and federal), industry, the science community, community interest groups, etc. Workshops may provide a suitable forum for the generation of ideas and directions.

- **Step 2**

  From this study construct a draft mission for the centre and review the opportunities for useful partnerships with existing organizations. Also review the needs of the centre for specific types of expertise and make a first attempt at defining a client base. This document should basically be a preliminary proposal for a national centre including a draft cost proposal. If possible look for some sort of governmental approval in principle pending a final formal proposal for the centre so that meaningful negotiations can be taken under Step 3 below.

- **Step 3**

  Conduct negotiations with the potential partners as to the possibilities for mutually beneficial partnerships or at least support for the establishment of the centre. During these negotiations the mission can be revised to reflect the partnerships. The expertise proposed for the centre can be modified as well.

- **Step 4**

  Do a study of the computer and communications hardware required and a final plan for the staff needed. Prepare a draft administrative organization. It will probably be necessary to have assistance from appropriate specialists for this step. Contracting for some of the experts is an option.

- **Step 5**

  Having a more final version of the mission and information on partnerships, prepare a final proposal for the centre for final approval. This proposal should include a high-level national data management plan for ocean data. Recruit the appropriate team to prepare it. Include a more detailed budget.

- **Step 6**

  Develop in consultation with the partners clear written agreements on the responsibilities of both sides. This will ensure that expectations are met in the future. If expectations are not met due to misunderstandings there will be difficulties in maintaining productive partnerships.

These steps and principles are suggestions for guidance only. Depending on the circumstances and administrative arrangements in the organization establishing the NODC, modifications may have to be made, or a completely different approach taken. Director's of NODCs in other countries could help with this process based on their experiences. Member States should not hesitate to contact NODCs in other countries or the IOC Secretariat for advice in developing their NODC.

After the centre is approved and staffing is complete there will be an appreciable time before routine operation is achieved. There should be two main thrusts. The first will be to acquire and implement the computer and communications hardware and software. The second will be to make the centre visible to the client community.

In implementing the systems it is important to choose some goals that can be achieved earlier rather than later. There will be significant expectations within a broad community. It is not a good idea to implement a new centre and then not see anything produced for two years or more. It is important that visible progress becomes evident quickly. Even small steps that nevertheless show distinct progress will keep up interest in the centre. It is important not to promise more than can be delivered. This will maintain confidence in the client community and among partners.

One method of achieving quick results with a moderate development effort is the establishment of a meta data system or data directory. The development of a meta data directory has a number of benefits:

- It involves the marine community in the process of establishing the NODC in the early stages and gives a feeling of ownership of the activity.

- It creates a directory for the marine science community and becomes a useful reference source.

- It provides the new data centre with an indication of quantities and types of data that exists at the national level.

- It is relatively easy to set up.

4. DEVELOPMENT OF A CLIENT BASE

This section provides some guidance on potential clients for a national oceanographic data centre. Some of the material here can also be useful in examining the need for an NODC in the initial stages of deciding whether or not to establish the centre. The sections below mention some economic considerations that will influence the decision. If a nation has a large investment in oceanographic data collection it may be appropriate to ask whether it can afford not to have an NODC.

There are several groups of potential clients that can need support from a national oceanographic data centre. While defining and developing data centre services contacts should be established and requirements should be reviewed through a consultative process. Knowledge of these requirements will be valuable in deciding priorities for choosing data to be included in the centres data banks, in defining services and products to be provided, and in choosing the types of expertise data centre staff will need. The following sub-sections provide a brief overview of the types of clients a data centre might serve.

(i) Research Science

The research community has long been one of the most important clients for data and services from an oceanographic data centre. As the experiments and programmes carried out by scientists increase in complexity and size, the need can only increase. Multi-disciplinary research is becoming the norm rather than the exception. More and more the datasets that are needed are assembled from data collected by others in order to meet the requirements for geographic and temporal coverage, and to include suites of parameters describing physical, chemical and biological properties of the oceans and atmosphere. It is in this area of data integration that data centres can play a key role.

There is a more urgent need to have these datasets assembled in a shorter time-frame to provide inputs to numerical and other models that are trying to understand and predict effects on the environment, climate change, the fishing industry, etc. The usefulness of a data centre that can efficiently assemble these datasets and provide them to present and future users is clear. Oceanography is moving from mostly research to application of what has been learned to real problems. The problems being addressed have significant economic aspects. It is important that the best and most recent data be available for these applications.

There are also economic benefits in that a good data management plan can reduce the costs of handling data and producing the necessary information for clients. *If the data are processed and documented properly once, then the clients can get on with the science rather than having to spend a significant part of their time being data managers as well.* This is an important point. Good data management can enhance and improve the effectiveness of a countries ocean science programmes.

(ii) Marine Engineering Interests - Ship and Offshore Facilities Design

When building and operating marine facilities it is necessary that data be available to ensure a good design. If adequate data are not available facilities (and ships) may be over engineered to cope with the unknown. This increases costs. If the facility is under engineered serious consequences will likely result. The consequences can be of several types. Chemical spills can endanger the environment and human health. Human safety can be at risk.

The economic consequences of inadequate design can also be serious. Poorly designed facilities can be more expensive to operate and can be extremely expensive to repair or rebuild after a failure. Cleanup of toxic spills has huge costs. Damage to food supply results in large economic losses. If an expensive marine facility suffers significant down time because of inadequate design, it make be impossible to recoup the original investment.

Clients for marine data include designers and operators of ships and offshore platforms, including navies. They are interested in waves, water levels, ocean temperature, currents and some chemical properties. Developers of ferry services have similar interests both for the design of the ferry boats and the length and slopes of ferry ramps. Statistics of waves, water-levels, and currents are important for strategic planning to ensure operation in expected conditions a large percentage of the time.

Developers of shore-based plants of all kinds are interested in water-levels and waves for design of water intakes and outlets. Physical, chemical and biological water properties are needed in design, materials selection and for anticipating the requirements for operational maintenance. Storm surge occurrences have to be considered in the design as well.

The development of harbours and their maintenance through dredging is another activity that needs to be done based on all available data for the area. Waves, water levels and storm surges will determine the percentage of the time the facility can be used safely. Waves and currents will determine the amount of dredging that must be done to keep the facilities operational. Waves and water-levels will determine the size and types of ships that can use the facilities.

In short, there are a large variety of clients for ocean data in the marine engineering and ship building industries. It is important for reasons of environmental and human safety that the best datasets be available to them for the design of these facilities. Since designs are generally for facilities that will be in use for decades, the longer the time-series the better. Data centres should be there to archive data and keep it for these future uses.

(iii) Regulatory Agencies

Most countries have regulatory agencies that oversee various types of marine activities. This includes such activities as commercial and recreational fishing, hydrocarbon exploration and production, passenger ferry operation and ocean dumping. To regulate these activities requires both strategic and operational data. Strategic data is required in order that the regulator can approve the design of ships, platforms and associated facilities. In fact, strategic data is also required so that the regulator can approve the operational plans for the activity. An example of this would be resupply of an oil rig. If a rig requires resupply of food and fuel on a regular basis, then the regulator would need data to ensure that the resupply operation could proceed as necessary during stormy winter months.

The activity being regulated would generally be required to collect certain types of data when it was operating. The purposes of the data would be several. In the first instance the rig or ship would need the data to make operational decisions about activities on board. The data would be needed for engineering use when unanticipated problems occur. The operational data would be archived and periodically be used to look for changes in "average conditions" over time.

The regulator would require the data to ensure that the operator was conforming to conditions of his license. For example, if an activity was to be closed down under certain wave conditions the regulator could monitor conformance.

In the case of the fishing industry the operational data could consist of temperature and salinity profiles which could be used in determining the preferences of species of fish. Wave conditions could be used in determining the performance of various types of gear in prevailing conditions.

Another factor is the increasing use of strategies for sustainable development of renewable resources. This requires regulatory agencies to access the best available and most comprehensive datasets to ensure the decisions they make are valid.

In summary, regulatory agencies and the industries they regulate are both an important user of oceanographic and marine meteorological data and a producer of data. A new data centre should be aware of these requirements and establish the necessary contacts to provide appropriate services and obtain data that are collected to be archived for future users.

(iv) Forecasters

Client agencies who issue forecasts of marine weather and other conditions such as water levels, storm surges, severe waves and high winds are users of marine data and information. The data that they need come from a variety of sources. Ships at sea report oceanographic and meteorological conditions. Satellites report winds, waves, sea surface temperature and are beginning to report ocean fronts and other features. Drifting and moored buoys report ocean currents, and surface and subsurface oceanographic variables as well as some surface meteorological conditions.

Historically, NODCs have not usually been part of these data flows. However, more and more data are being reported in real-time. For a significant and growing fraction of these data the real-time report is becoming the only way these data are available. Consequently it is becoming in the interest of NODCs to capture the real-time data for archival and use by its clients. Once an NODC begins receiving the real-time data there is an opportunity for the NODC to provide processing, quality control and synthesis of specific datasets for clients. NODCs in some IOC Member States are doing exactly that. Potential clients include not only those issuing forecasts and warnings but also regulators, the shipping industry, offshore operators of various sorts, and the fishing industry.

It should be noted however, that capturing and processing the real-time data flows will impose a significant workload on a data centre. There are a lot of real-time data on the networks. A new NODC would probably not collect the real-time data unless it had clients that were directly interested in using it in operational time-frames. If the new centre is only interested in acquiring the data for its historical databases, the real-time data is available at least annually from RNODCs, WDCs and other data centres in the IOC and WMO systems.

(v) Insurance Industry

Although not as yet a large user of oceanographic data, the insurance industry has shown an increasing interest in change of various sorts. Data that can be used to indicate change in frequency of storms, floods, or the success of agricultural crops is useful in anticipating changes in the pattern of insured losses.

(vi) Education

Educational institutions at the primary, secondary and university levels often show an interest in oceanography and in oceanographic data. At the primary and secondary levels the interest usually consists of requests for general information about oceanography. By providing a lecturer who is armed with general information and summary data concerning phenomena such as tides, wind waves, tsunamis, ocean circulation, global warming, potential pollution problems, etc. the NODC can assist the schools in providing a more interesting and realistic curriculum for their students.

At the university level, the data centre can provide a vital support to important research projects through the provision of national and international data relevant to the problem at hand. There is also the opportunity to foster interest in research topics of the NODC or to the organization to which the NODC is attached.

Working closely with schools and universities can also provide opportunities to interest students in careers in oceanography. It also provides an excellent opportunity to recruit capable students for employment during summer breaks or between university semesters.

(vii) Fishing Industry, Naval Operations, and Agricultural Interests

The fishing industry is a major client for oceanographic data and information services. Those monitoring fish stocks for management purposes, and researchers modelling population dynamics and examining processes that affect recruitment of the various stocks must include environmental factors in their studies. Fishermen of certain species find it useful to know temperature and salinity conditions to help locate the fish they are seeking. Designers of fishing equipment and fishing boats need various physical data while developing designs.

Navies in IOC member states are also important clients for oceanographic data and information services. Both the historical distribution of water properties and knowledge of present conditions in ocean areas of interest are important for naval operations.

Another source of clients for an NODC are agricultural organizations that plan planting and marketing strategies. Although these users are at an early stage in the development of uses for the data an NODC might hold, some are now using information such as the occurrence of El Niño events in the course of their work.

(viii) General Public

The volume of requests from the general public in a Member State will generally not be large. The requests will usually be of a general nature driven by curiosity. Occasionally there will be requests for more specific information as someone examines the possibility of starting a business related to the ocean. Some requests may come from recreational fishermen. As is the case of providing service to educational institutions, there is an opportunity to promote oceanography in general and the services of the data centre in particular. It is also important to remember that this member of the general public is probably supporting the operation of the centre by paying taxes.

5. SELECTION OF COMPUTER TECHNOLOGIES FOR THE NODC

The selection of computer hardware and software for the NODC can be daunting because of the number of choices available and the rate of development in the industry. It is important that a careful study be carried out by staff or contractors whose expertise and experience is up to date in order to do proper evaluations and make well documented and supported recommendations. For this task discussions with other NODCs can be useful.

The computer technologies available at present cover a wide range of needs whether the centre being developed is large or small. A variety of software and hardware are available with capacities and costs that range from those appropriate to the home computer market to those that meet the needs of satellite data managers and facilities involved in modelling of complex phenomena.

Another consideration for a data centre in selecting its computer technologies is that data will be received from a variety of computer types and operating systems. For example, the centre may have to deal with data from PCs and Unix workstations. It is important to have the ability to read and write data for the types of systems used by the centre's clients.

The following sections discuss some of the computer-based functions that will be carried out in a modern NODC and the technologies associated with them. To describe these technologies in any detail is beyond the scope of this manual and would soon be out of date. The information provided here is of a general nature and is intended to provide a list of the capabilities that a modern NODC should consider.

In what follows it is assumed that in addition to a need to deal with digital data, there is a need to deal with documents and information, including data products and publications, as well as meta data. There will also be a need for digital communications in the form of local and wide area networks. In the case of a small centre that is supporting a largely local operation there will still be a need for capabilities in document and information handling, and in networking. The difference between centres will be in the capacity of the systems that are implemented.

(i) Data Storage

One of the main functions of an NODC is to store data in digital form. There is an important choice to be made as to whether and how much data needs to be stored on-line as opposed to off-line. Data can be stored on-line on magnetic disk or optical media. Magnetic disk is by far the most common because the speed of access is greater and the initial cost is lower. Optical storage will likely become more competitive in the medium term. Data are more accessible on-line and may in fact have to be stored on-line if used in many of the commercial database management systems. On-line storage is most suitable for datasets that change frequently or from which filtered subsets need to be extracted on demand without operator intervention.

Storing data off-line on CD-ROMs is a very practical option for larger volume datasets that are not changing frequently and for which immediate extractions are not necessary. This is particularly true for datasets that are self contained such as a group of wave time-series data at a number of locations in an area. If a time-series is complete and processing is finished, the data can be stored on a CD-ROM. For particular requests, a data centre staff member can put the CD-ROM into a PC and extract data of interest for clients or simply supply a copy of the CD-ROM. This approach is less practical for a dataset that is continually being updated. Creating the new CD-ROMs will be a significant task and creating them in volume will be expensive.

The same considerations apply to high capacity cartridge type magnetic tapes as operator intervention is required to mount the tape and the read times can be long. Cartridge type magnetic tapes include "Exabyte" and "DAT" tapes. Although the transfer rates are large they must be read serially. If the data file is a long way down the tape it takes significant time to get to it. CD-ROM on the other hand is a random access device and a particular file can be found quickly.

Most large and small centres use magnetic disk for on-line storage, and a mix of CD-ROMs Exabyte tapes, or DAT tapes for off-line storage.

(ii) Data Processing

In most oceanographic data centres the data processing will be carried out on workstations. These workstations will range from inexpensive PCs to high end Unix and graphic workstations.

Some of the processing is suitable for interactive processing and some is appropriate to operating a batch queue. A few centres still maintain a mainframe for this batch processing. Batch processing is indicated for processing large numbers of records when a person is not required to make decisions if certain conditions occur. Most data centre processing systems of today will have interactive and batch processing components.

An issue that is important to the staffing and expertise that must be maintained in a data centre is the question of the use of commercially available software versus specific application software written in-house or under contract. When there is a choice use of well established commercial software is often the best choice as it will be maintained and improved by the company. However, it will usually be necessary that a significant amount of special purpose software will have to be written or otherwise developed for handling oceanographic data.

In terms of computer data processing capacity it should be noted that much of the data managed is four dimensional having both space and time attributes. It is important that computer systems have the ability to handle this complexity to ensure operations occur at an adequate speed.

The following are examples of types of special purpose software that will usually have to be written for processing oceanographic data:

- decoding and reformatting of data stored in complex coded or bit packed formats;
- algorithms to effect the more complex types of quality checking of the data;
- special analyses not provided in standard database and spread sheet applications;
- graphical outputs for complex data products;
- coding of spread sheet formulae or database summary fields.

(iii) Local Area Networking

A data centre computer environment will usually include several workstations. These workstations may be a mix of IBM PCs or clones, Unix or graphic workstations, and others. The workstations serve functions such as copying data from diskettes to on-line disk, data editing, spread sheet applications, database management, document scanning and optical character recognition, and desktop publishing. The centre would also require laser printers, and possibly a device for writing CD-ROMs.

To move the digital data from one workstation to another the centre would be well advised to implement a local area network (LAN). Ethernet LANs are almost standard equipment on most systems and the additional costs of implementing the LAN are small. A LAN is the most inexpensive and convenient way of moving data between dissimilar systems.

(iv) Wide Area Networking - The Internet and the GTS

There are two types of wide area networks (WAN) of interest to oceanographic data centres. The first type is a request-reply type of network. This type of network generally involves two or more clients exchanging data or information by specific request. An example is a user executing a file transfer protocol (ftp) to copy a file from a remote node to a local workstation. The other type of network is a broadcast network. The Global Telecommunications System (GTS) is an example of a broadcast network. Reports of meteorological and oceanographic observations from ships at sea are formed into bulletins and are "broadcast" to the world on the GTS. Centres that wish to receive them simply watch for the bulletins with headers indicating meteorological or oceanographic data and copy them off the network.

An oceanographic data centre may be interested in capturing oceanographic or surface meteorological data from the GTS. It is not common for NODCs to do this, but some do.

On the other hand, an NODC should if possible implement access to the Internet to connect to national and international centres that deal with oceanographic data and information. There is apparently an initiative to start an "Internet2" because the heavy load on the existing Internet is degrading response. However, at present the best way to reach the audience of interest to oceanographic data centres is to get on the existing Internet.

A connection to the Internet is important to an NODC because it is becoming the *defacto* standard for the distribution of oceanographic data from research programmes and some monitoring programmes. More and more oceanographic data centres are also providing data and information only via FTP over the Internet or from their World Wide Web sites. This is happening because it is an inexpensive, convenient and practical means of distribution. A dataset can be posted on an FTP or WWW site and made available to all users once. It removes the requirement to deal with many individual requests for the same data and solves the problems of incompatible digital media and formats.

In fact, the international scientific community is adopting the World Wide Web as the means of providing up to date programme and organizational information to users.

Thus, while it is realized that not all Member States have Internet connections available, those that do should make every effort to obtain such a connection. The World Data Centres (WDCs), Responsible National Oceanographic Data Centres (RNODCs), and other data centres must continue to publish as much information on CD-ROM to meet the needs of Member States who cannot obtain an Internet connection.

(vi) Publishing CD-ROMs

The other very useful method of distributing data and information from an NODC is to publish the material on CD-ROMs. The CD-ROM is of particular interest because the format was quietly standardized while the medium was being developed for commercial sale. Thus, CD-ROMs can be treated as platform independent for most purposes. Present CD-ROMs hold more than 600 megabytes of data. Most oceanographic data can be compressed by a factor or 4 or 5 at least by most of the standard data compression applications so that in reality most CD-ROMs can hold of the order of 2 gigabytes of data.

Producing a CD-ROM is also an inexpensive operation. A blank CD-ROM for a one off copy costs of the order of $10-$15 US. Duplicate copies in volume cost of the order of $1-$2 US. A machine to write CD-ROMs connected to a PC computer will cost $3,000 to $4,000 US, including the PC software. Thus, it is not expensive for a data centre to obtain the equipment and software to establish a CD-ROM publishing activity.

CD-ROMs are an excellent medium for publishing project data sets. A project dataset is all the data collected by a single project. Thus, the user can obtain a complete set of data in a very compact and convenient form. In fact, any dataset that is complete and not being updated is a good candidate for publication on a CD-ROM. Another example is time-series data such as the data collected by weather ships over the last several decades.

Many oceanographic data centres now use the CD-ROM as the medium for *ad hoc* data requests. Most users can not deal with 9 track magnetic tapes or exabyte or dat tapes. However, most will have a CD-ROM reader.

(vii) Graphics

Graphical outputs are required for useful data presentations including geographic mapping of variables, statistical presentations, line plots, photographs, and raster images of satellite data. These presentations are used in providing data products to clients using a variety of media including hard copy, CD-ROM or electronic versions of reports, and via the World Wide Web.

Graphical presentations are also used in the data centre to view data for purposes of quality control.

There are numerous methods of producing graphical outputs. Many commercially available software packages such as spread sheets and database applications have graphical outputs in various formats. There are a variety of individual products that have been designed for simple to complex graphical development from Computer Assisted Design (Auto CAD) to complex image processing applications for satellite data. Software products such as Photoshop are also useful for manipulation of raster images.

The NODC will need colour printers to produce hard copy of graphical presentations. Other graphical presentations may be distributed through the centres World Wide Web page.

Geographic Information Systems (GIS) are useful technologies for an NODC for mapping both data distributions from inventories and the distributions of ocean variables. GIS is not an easy technology to implement and centres would be well advised to have a well developed training plan for development and maintenance of their GIS expertise.

(viii) OceanPC

OceanPC is a software package for IBM PCs and compatibles. It has been developed by the Committee on International Oceanographic Data and Information Exchange and is freely available. OceanPC is a tool for processing and displaying physical and chemical oceanographic data. It includes tools for geographic mapping of ocean variables, line plots, and other useful displays. It also contains software for maintaining inventories of cruise summaries and for entering information for the IODE Cruise Summary Report.

New data centres may wish to obtain a copy of OceanPC from the IOC Secretariat to implement some of its capabilities into their data centre operations.

6. DATA MANAGEMENT PLANNING

A national oceanographic data centre is usually organized to have a central role in the national management of oceanographic data. As such it should be prepared to take a lead role in developing national plans for ocean data management. The goal of such planning is to ensure that the data are properly processed and documented, that the data are entered into secure archives so that they are not lost to future needs, that they are distributed nationally as required to interested users, and distributed internationally according to commitments made at the intergovernmental level.

There are secondary goals in good data management planning. These include elimination of duplication of effort in processing and quality control of the data, establishing agreed standards for quality control and meta data, and establishing responsibilities for sharing the data management workload for a project. Whether formalized or not there will be at least two organizations involved in the life cycle of a type of data such as the data from a group of drifting buoys. The data collector will launch the buoys and probably process and quality control the data for his own purposes. The data will then be forwarded to the data centre for its purposes. There may well be a number of other participants involved in the data flow including international centres. It is important that the responsibilities of each organization be clearly documented and understood in order for the data flow is to proceed as planned.

It is also good data management practice to establish milestones and time-frames for critical points in the data flow so that no one will be inconvenienced by not receiving data it was depending on.

One of the most basic elements of good national data management is detailed knowledge of the data that are being collected and a positive working relationship between the NODC and the national organizations that collect the data.

Figure 2 is a graphic representation of the "data management plan" for the Global Temperature-Salinity Profile Project. This is a complex data flow. It is included as an example here because it illustrates several of the elements that should be covered in a data management plan. It also illustrates the co-operation between nine organizations in 5 countries that actively contribute to the GTSPP data flow on a daily, monthly, and annual basis. Boxes 3, 4, 6 and 7 comprise the Data Assembly Centre for the Surface Velocity Programme of the World Ocean Circulation Experiment. Thus, GTSPP serves a dual purpose.

7. PRODUCTS AND SERVICES

If all the clients for an NODC were research scientists working in specific fields, there might not be a need for a data centre to produce data products. Scientists often wish to produce their own. However, as the rather diverse client community described in Section 4 above has developed and as research programmes have become increasingly multi-disciplinary in nature, the need for data products has been recognized..

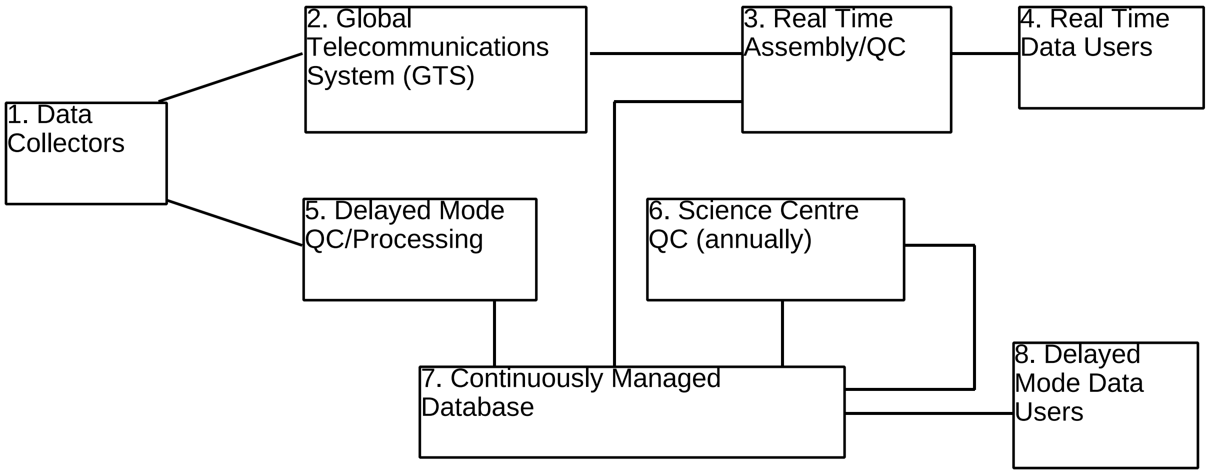
There will always be difficulties in defining useful data products for clients in oceanography. The data will often not be there to prepare the presentation that will neatly and completely answer the question at hand. The user may not be able to specify exactly the requirement because of a lack of understanding of what can be produced. A data centre must work with the clients to develop appropriate data products. The process will frequently be iterative in that initial attempts may need refinement before a satisfactory result is achieved. It is important that the data centre work closely with the client to understand the real need before developing the product.

Some examples of data products include atlases with maps of distributions of ocean variables, exceedance diagrams of significant wave height in an area, tidal predictions or tables of harmonic constituents for tidal predictions, or datasets of ocean observations filtered by area, time-frame and variables observed. Each of these products is designed for a specific use and has been developed for specific classes of users. A data centre should look at its client community, consult with them to determine their most important needs, and then develop the appropriate products.

Once again it could be helpful for the new NODC to review the World Wide Web sites of existing NODCs for ideas and examples of data and information products that have been found useful to clients in other countries.

The same considerations apply to the development of data centre services. It is not possible to write down a standard list of services all of which should be provided to clients by all NODCs in the world. The services provided by an NODC will depend on the clientele (national and international) that the centre is supporting.

The importance of maintaining close working contacts with the user community and working together to develop data centre products and services can not be over-emphasized.



In Box 1 data are collected from ships at sea or moored or drifting buoys. Low resolution data are encoded in the IGOSS BATHY or TESAC formats and entered onto the GTS. The full resolution data are forwarded to the shore based organizations operating the collection programs (Box 5).

In Box 2 the data are circulated on the GTS network of the World Weather Watch of the World Meteorological Organization.

In Box 3 the data are gathered from 3 points on the GTS by a centre in Canada that assembles the data, removes duplicates and applies the agreed GTSPP quality control. The data are collected from 3 points on the GTS (Canada, Germany, U.S.A.) to permit monitoring of the performance of the GTS in routing data around the world. The data are passed from this centre to real time data users around the world (Box 4) and to the Continuously Managed Database in the U.S.A (Box 7).

In Box 5 the full resolution data are processed by the labratory that collected it. The data are calibrated, passed though complex quality assessment programmes and are forwarded to the Continuously Managed Database.

In Box 7 the data are entered into a database according to a priority scheme that specifies that certain versions of the data replace other versions. The latest and best observations replace earlier and cruder versions. For example the full resolution calibrated versions of the data in the delayed mode stream will replace the low resolution uncalibrated BATHY and TESAC versions.

In Box 6 data for a year are received by three oceanographic laboratories in the U.S.A. and Australia. These centres apply sophisticated quality control procedures that check for consistency with other data from the geographical area. The data are then returned to the CMD and replace the versions of the data that have not had scientific quality control. The centre that operates the CMD also provides data and data products on request to users (Box 8)that did not require the data in real time and were prepared to wait for the more complete and better quality data a year or two later.

Figure 2. The GTSPP Data Management Plan

While it is not appropriate to specify in full the services which all data centres should provide, it is possible to write a list of some services and that will be common to most centres. These services would include:

- developing and protecting national archives of oceanographic data.

- providing regular directory and inventory information on its data and information holdings. The IOC Marine Environmental Data Inventory (MEDI) is an example of this type of product.

- providing subsets of its data and data products to clients in a clearly documented format and with available supporting information and meta data attached.

- performing a level of quality control on its data holdings that will identify and flag easily detected errors and inconsistencies.

- acting as a referral centre for other national and international oceanographic data and information.

- receiving data for specific processing followed by delivery of the processed data or products on to someone else as part of a scheduled data management activity with a group of partners.

8. QUALITY CONTROL

One of the most important activities that contribute to the performance and reputation of a data centre is the quality control applied to its holdings. There are several classes of errors that can occur in data. Some can be detected easily, others can not. Obvious errors such as an impossible date, time, latitude or longitude are easily detected. Some of them are easy to fix. More subtle errors such as an error in latitude or longitude for an observation can be detected because a ship did not have time to get to the new position from the old one in the elapsed time between stations. Checks on the values of variables against climatology may also be indicators that the position or date of the observation may be incorrect.

These sorts of errors are the types of errors that a data centre should look for in incoming data. The tests can be automated and applied by computer. Software can be developed that allow a person to interact with the data to correct some of the errors. Other errors will just be flagged as incorrect or suspicious when a correction can not be made with a high level of confidence.

The most subtle errors that might occur in data cannot usually be detected by a data centre. An instrument may be slightly off calibration but the values it produces are within the variability of a climatology for the area. This type of error can only be detected if the scientist or technician re-calibrates the instrument and discovers the problem.

Clients will rapidly lose confidence in a data centre that distributes data with a significant number of obvious errors in it. Therefore it is important that an NODC take the matter of quality control seriously and apply at least a level of automated tests to detect impossible values and the more obvious inconsistencies such as significant deviations from climatology.

Some data centres use graphic workstations to view profiles of ocean variables, ship tracks, and time-series of variables in order to detect errors or flag possible inconsistencies. This activity is relatively labour intensive and not all centres will have the resources to carry out a visual inspection of their data. However, it must be realized that at this time such an inspection provides the best and most reliable quality control.

The other problem is that there is no adequate commercially available software for quality control of ocean data. The software must be developed by data centres and oceanographic laboratories. Unfortunately, it has been the experience with IOC that most of this software is computer system dependent and cannot be easily shared. However, the algorithms can be shared. There are several publications that detail how various QC systems work. These publications are listed in Annex A.

PART II. INTERNATIONAL SERVICES

Section 1 gives a short overview of the international organizations that manage aspects of the exchange of international oceanographic data and information.

Section 2 discusses the Intergovernmental Oceanographic Commission which is a lead organization for international exchange of ocean data and information.

Section 3 lists and briefly describes the subsidiary bodies of IOC that are most relevant to the work of a national oceanographic data centre.

Section 4 describes the system of oceanographic data centres. Some of these data centres are operated on behalf of the international community by NODCs.

Section 5 describes how an NODC can participate in international exchange activities and contribute to the global science programmes.

Section 6 is a brief summary to put Part II in context.

1. OVERVIEW OF INTERNATIONAL DATA MANAGEMENT ORGANIZATIONS

There are 3 organizations that have major roles in the management, international exchange, and archiving of oceanographic data and information.

(i) The Intergovernmental Oceanographic Commission (IOC)

The IOC is a part of UNESCO, the United Nations Educational, Scientific, and Cultural Organization. As part of its mandate IOC is responsible for promoting international co-operation in carrying out scientific studies of the oceans and for co-ordination of international oceanographic data and information exchange. IOC also supports data management for many of the global scientific experiments through its subsidiary bodies including IGOSS, IODE, GLOSS, etc.

(ii) The International Council of Scientific Unions (ICSU)

The World Data Centre System was established by the International Council of Scientific Unions (ICSU) in the early 1960s to ensure that data collected during the International Geophysical Year would be available to scientists in all nations. There are several WDCs that deal with oceanographic data. There are WDCs - Oceanography A, B1, and D in the USA, Russia, and China respectively. The USA and Russia also operate WDCs for marine geophysical and geological data.

(iii) The World Meteorological Organization (WMO)

The WMO is primarily concerned with meteorological data collection and exchange. Its programmes cover both land and sea. The marine programmes are managed by the WMO Commission on Marine Meteorology. WMO is primarily interested in surface oceanographic data for use in weather and climate modelling. WMO co-sponsors with IOC several programmes concerned with ocean data collection and management.

The programmes of the IOC will be discussed in detail below. Programmes that are joint with WMO or ICSU will be identified.

2. THE INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION

The oceans cover about 70% of the earth’s surface. They influence in a very significant way both global environmental change and sustainable development. It is therefore essential that we understand and are able to predict global and regional ocean conditions and interactions with the atmosphere, biosphere and land. The Intergovernmental Oceanographic Commission of UNESCO was founded in 1960 because of this need. The goal of IOC is to promote enhanced study of the oceans through cooperative programmes. At the present time the IOC has 125 Member States.

The IOC has two types of subsidiary bodies. They are scientific/technical bodies, and regional bodies. Regional bodies include regional committees and sub-commissions. Scientific and technical bodies have been created at several levels to develop and manage the programme of the IOC. It has already been noted that several of these subsidiary scientific and technical bodies are co-sponsored or co-operative programmes with WMO and ICSU.

For purposes of this manual only the subsidiary bodies that have global responsibilities for international data and information management will be discussed. For more information on the other subsidiary bodies and on other matters concerning IOC visit the IOC World Wide Web Homepage, see the publications listed in Annex A, or contact the IOC Secretariat as described in Annex B.

3. SUBSIDIARY BODIES OF IOC WITH GLOBAL DATA MANAGEMENT RESPONSIBILITIES

The following subsidiary bodies of IOC have responsibilities for developing and co-ordinating programmes for the management and exchange of oceanographic data and information. The groups meet on a regular basis to review requirements and plan the development of the programmes in their area of responsibility. The frequency of these meetings vary. The proceedings of the meetings are recorded in the form of summary reports. Copies of the summary reports are widely circulated among Member States of IOC or are available from the Secretariat in Paris. Some of them can be found on the IOC WWW site.

The last 3 subsidiary bodies, GEBCO, GIPME, and GLOSS, are data producers, data users and managers of global datasets.

(i) IOC-WMO Joint Committee for the Integrated Global Ocean Services System (IGOSS)

The Joint Committee for IGOSS sponsored by IOC and WMO is responsible for procedures, formats, and standards for several types of ocean data collected and circulated by telecommunications means in operational time-frames. Operational time-frames are generally considered to be from hours to of the order of one month. The data are collected by moored and drifting buoys or ships, encoded into agreed code forms and circulated on the Global Telecommunications System. Variables that are used to forecast weather and severe conditions that threaten human and environmental safety are of particular interest to IGOSS. IGOSS also provides support to the global ocean experiments by helping to circulate the data quickly so that it can be used in managing the experiments. Arrangements have been made for the archival of IGOSS data in the IODE system so that it will not be lost.

The Joint Committee for IGOSS has the following subsidiary bodies.

- IGOSS Group of Experts on Operations and Technical Applications.
- IGOSS Group of Experts on Scientific Matters.

(ii) International Co-ordination Group for the Tsunami Warning System in the Pacific (ITSU)

This Co-ordination Group manages and further develops as necessary a network of water level gauges in the Pacific Ocean and a supporting processing and warning system. The system detects and provides early warnings of tsunamis caused by earthquakes or other catastrophic events. This system has been in operation for almost 3 decades. It is of primary interest to Pacific nations that are vulnerable to tsunamis.

(iii) Data Buoy Co-operation Panel (DBCP)

The Data Buoy Co-operation Panel was jointly established in 1985 by WMO and IOC, then under the name of *Drifting* Buoy Co-operation Panel. The principal objectives of the panel are: (i) to achieve the optimum use of any data buoy deployments being undertaken worldwide and an increase in the amount and quality of buoy data available to meet the objectives of major IOC and WMO programmes; and (ii) to encourage and support the establishment of “action groups”, in particular programmes or regional applications to effect the desired co-operation in data buoy activties.

The DBCP holds annual meetings in conjunction with the Argos Joint Tariff Agreement meetings where the costs of platform location and data collection for governmental users are negotiated with Collecte-Localisation-Satellites/Service Argos, the provider of the necessary staellite services.

(iv) Global Ocean Observing System (GOOS)

GOOS is an internationally co-ordinated system for systematic operational data collection (measurements), data analysis, exchange of data, generation of data products, technology development and technology transfer. It will use a globally co-ordinated, scientifically based strategy to monitor and forecast envirnomental and climate change globally, regionally and nationally. Data will be generated by remote-sensing and by repeat sampling using sea-surface and subsurface instrumentation in the open sea and coastal regions worldwide. Major physical, chemical and biological variables have been or are being identified that can be used to provide an integrated assessment of the current health of the oceans and their living resources, and early warning of deteriorations. GOOS will provide a mechanism and an infrastructure for data and information to be made available on various time-scales to participating nations, thereby strengthening individual national observing capabilities.

GOOS will be established by Member States and implemented largely through nationally or regionally owned and operated facilities and services. It is intended that the data collection and management aspects of GOOS be built upon the existing systems of IGOSS, IODE, GLOSS, DBCP and others. Co-ordination will be provided by the IOC in co-operation with the other sponsors, WMO, UNEP and ICSU. GOOS is based on the principle that all countries should participate and that particiapnts should make certain commitments according to their capabilities, so that all countries can both contribute and benefit.

There are two high level committees for the development of GOOS. the Intergovernmental Committee for the Global Ocean Observing System (I-GOOS) and, from the end of 1997, the GOOS Steering Committee (GSC). I-GOOS is the body that co-ordinates Member States efforts for establishing GOOS, and the GSC is the body that provides the scientific and technical design for GOOS and develops the plans for implementing it.

GOOS, IGOSS and IODE representatives are meeting under the umbrella of the Joint GCOS/GOOS/GTOS Data and Information Management Panel (J-DIMP) to devise and implement a GOOS data and information management plan. As GOOS defines its data collection requirements and the processing and distribution systems begin to take form, NODCs will have the opportunity to contribute to the operations of GOOS.

(v) Committee on International Oceanographic Data and Information Exchange (IODE)

The Committee on IODE is the IOC subsidiary body that will be of primary interest to NODCs in Member States. In fact, most of the participants at the IODE Committee meetings are directors or staff of NODCs in the Member States of IOC.

IODE committee meetings are held about every 3 to 4 years. The meetings develop the agreements, procedures and means of sharing and exchanging ocean data and information among the Member States of IOC. The meetings also work to set standards for data and meta data so that future users will be able to clearly understand the circumstances of the data collection and will be able to use the data properly.

The Committee on IODE is a scientific/technical committee. Many of the agenda items dealt with at the meetings are very technical. To insure that the technical experts to discuss these matters are available at the meetings, Member States are strongly encouraged to include directors and staff of their NODCs on their delegations.

When a new NODC is planned in a Member State of IOC, that NODC should expect significant involvement in IOC programmes and in IODE meetings. Many of the NODCs in operation at the present time have received as much as 50% of the data included in their database from international exchange through the work of the Committee on IODE.

IODE also provides data management support to the global science experiments including components of the World Climate Research Programme. IODE is identified as one of the existing systems that will provide support to the Global Ocean Observing System.

During the last 10 years the Committee on IODE had the following subsidiary bodies that assisted in the development and management of the IODE programme:

- Task Team on Remotely Sensed Oceanographic Data;
- Task Team on Development of IODE Data Centre Services;
- Task Team on Exchange of Marine Geological and Geophysical Data;
- Task Team on Marine Biological Data;
- Task Team on Oceanographic Data Quality Control;
- Group of Experts on Responsible National Oceanographic Data Centres (RNODCs) and Climate Data Services;
- Group of Experts on Technical Aspects of Data Exchange;
- Group of Experts on Marine Information Management (MIM).

These subsidiary bodies are being changed from time to time by the decisions of the Committee on IODE. For the latest available information on the composition of the Committee visit the IOC World Wide Web site or see the current publication of the IODE Handbook Which is updated and reprinted after sessions of the IODE Committee.

(vi) General Bathymetric Chart of the Oceans (GEBCO)

Major achievements of GEBCO over the past years have been the digitization of the contours of the 5th Edition and the preparation of the GEBCO Digital Atlas. Now the "GEBCO Digital Atlas (GDA)" is available on Compact Disc (CD-ROM). This provides the results of the interpretative contouring in a computer compatible form so that they can be widely used and flexibly manipulated.

Data on ocean bathymetry are very useful to NODCs for placing depth contours on maps of ocean variables. The data can also be useful for scientists developing numerical models of ocean processes.

(vii) Committee on the Global Investigation of Pollution in the Marine Environment (GIPME)

GIPME investigations focus primarily on the coastal zone and shelf seas but also deal, where appropriate, with the open ocean e.g., the open ocean baseline study. The Programme assesses the presence of contaminants and their effects on human health, marine ecosystems, and marine resources and amenities, both living and non-living.

The GIPME programme includes data acquisition on the levels and fluxes of contaminants. These data may be of interest to a new NODC.

(viii) Global Sea-Level Observing System (GLOSS)

GLOSS is an international system initiated in 1985 and co-ordinated by the IOC. GLOSS provides high-quality standardized sea-level data from a global network of sea-level stations. The measuring system has become known as GLOSS because it provides data for deriving the Global Level of the Sea Surface, a smooth level after averaging out waves, tides and short-period meteorological events.

GLOSS collects an important dataset on sea- level. Data from GLOSS stations are processed and archived at the Permanent Service for Mean Sea-Level (PSMSL) in the UK. These data are available to NODCs in a number of forms.

4. THE SYSTEM OF OCEANOGRAPHIC DATA CENTRES

(i) World Data Centres

As previously discussed, the ICSU system of World Data Centres has 3 centres devoted to Oceanography and 2 devoted to Marine Geology and Geophysics (MGG). The programmes of the WDCs are managed by the ICSU Panel on World Data Centres. Annex A provides a reference to the ICSU Guide to the WDC System. Annex B contains additional information on contacting the WDCs for Oceanography.

The Directors of the WDCs for Oceanography work closely with the IODE Committee. In fact the Directors are Officers of IODE. The IODE Officers meet intersessionally to monitor development of the programmes and provide guidance in meeting the goals set for an intersessional period. The Chairman of IODE periodically attends meetings of the ICSU Panel on WDCs to provide further co-ordination of the programmes of the two organizations.

NODCs and RNODCs are expected to transfer copies of data that meets the requirements for international exchange to the WDCs on a regular basis.

(ii) Responsible National Oceanographic Data Centres (RNODC)

The concept of the Responsible National Oceanographic Data Centre was developed to provide a means to supply assistance to the World Data Centre System. With the increasing complexity, variety and volume of oceanographic data it was recognized that the acquisition, processing, and archiving requirements were beyond the capacity of the existing WDCs. The IODE Committee responded with the creation of a number of RNODCs. An RNODC may take on a responsibility for a particular data type. An example of this is the RNODC for Drifting Buoy Data. Alternatively an RNODC can take on the responsibility for creating a multi-disciplinary archive of data for a given geographical area. An example of this type of centre is the RNODC for the Southern Oceans.

An RNODC is established by the IODE Officers on behalf of the Committee and terms of reference are agreed. The terms of reference clearly define the responsibilities for data and services. NODCs will be expected to provide copies of relevant data to the RNODCs. In return, NODCs can request data and data products in the area of responsibility of the RNODC and receive them promptly and at no more than a nominal charge for reproduction.

(iii) National Oceanographic Data Centres

National Oceanographic Data Centres are the engine of the international data exchange system. NODCs are depended upon to process and make available the data that are archived internationally in the ICSU-IOC system. It is almost always NODCs that take on the responsibility to operate an RNODC in support of IOC and the ICSU WDCs.

The various activities of NODCs in international data and information exchange are discussed in more detail in Section 5 below.

(iv) ICES Regional Data Centre

The data centre of the International Council for the Exploration of the Seas has for many years provided data centre services to several Member States of ICES. ICES co-ordinates its data management programmes with IOC and ICSU by participating in meetings of the Committee on IODE and contributes to many IODE programmes. Several NODCs of IODE in North Atlantic region also provide members to the ICES Data Management Committee.

(v) Permanent Service for Mean Sea-Level

The Permanent Service for Mean Sea-Level (PSMSL) is a focal point for global mean sea level data. Member States of IOC and others supply mean sea-level data to the PSMSL on a regular basis. The data is available to users interested in sea-level change and related climate problems.

5. INTERNATIONAL RESPONSIBILITIES OF NODCs

As was stated earlier the first responsibility of a national oceanographic data centre must be to provide a good national service to its clients. Once the national services and functions have been specified, the centre should look at what its international role will be. The centre should not spend more of its resources on international activities than on national activities. Up to 15% to 20% might be a reasonable limit on the resources to be devoted to international programmes. It is worth stressing once again that a centre must operate a strong national programme before it can expect to operate a good international programme.

The following paragraphs describe typical international activities that can be undertaken by a national oceanographic data centre. A centre might not take part in all such activities. For example, a new centre or a small centre would probably not be in a position to take on the responsibilities of an RNODC.

(i) Participate in the International Exchange of Oceanographic Data and Information

The IOC/ICSU Manual on International Oceanographic Data Exchange is the best source of the guidelines and procedures for international exchange. (See Annex A for references to this and other relevant publications)

Participating in the international exchange of oceanographic data according to the policies of the Committee on IODE is the most important international activity for an NODC.

This activity is the one that produces the largest benefit for the national centre. The benefit of course is data and information relevant to the centres area of interest; home waters. It is very cost efficient to receive a station of oceanographic data by exchange rather than having to collect it. It probably cost US$1 to receive a station of data by exchange as opposed to US$3,000 to $5,000 or more to collect it from a ship. Also, as already mentioned, some Member States of IOC have doubled the amount of data they hold for their area of interest through international exchange.

Participating in international exchange means supplying national data to the World Data Centre System and to the RNODCs that support the WDCs. Centres should take full advantage of the

WDCs and the RNODCs to acquire data collected by others in their area of interest. This may take the form of requests for specific data identified from catalogues, an annual request for any new data for the area of interest of the centre, or in some cases a negotiated agreement for regular updates. The onus is on the national centre to keep informed of what is available from the international system, and to make the approach to the appropriate international centres.

There are two routes that can be followed in international exchange. In addition to exchanging data through the IOC/ICSU data centre systems, a nation can negotiate bilateral exchange agreements with other countries. These bilateral agreements are not part of the international arrangements of IOC. They concern only the countries involved. Normally, an NODC would have some bilateral agreements as well as participating in the usual international exchanges. Bilateral agreements might for example, be arranged when neighbouring countries have a sea area of common interest.

(ii) Participate in the Meetings of the Committee on International Oceanographic Data and Information Exchange

It is important for a centre to participate in the meeting of the Committee on IODE and its subsidiary bodies. The discussions in plenary, in the working groups and drafting groups that are established during the sessions, and the informal discussions in the hallways cover far more than can be captured in the summary reports of the meetings.

The meetings also present a very good opportunity to discuss common problems with colleagues who may have already solved them. There is an opportunity to make contacts for bilateral arrangements for sharing of software and techniques, or for sharing the work and costs of a development project.

When choosing the delegation for a meeting of the Committee on IODE, Member States should include technical experts on the agenda items of the meeting that are of interest to them. It is also highly desirable to have an expert on marine information management in the delegation. If the data centre does not have such an expert, one can often be found in a marine science library.

The benefits of attending the meeting will significantly outweigh the costs of travel and subsistence.

(iii) Participate in the Development of the International Inventories of IODE

There are 3 international inventories that are operated by IODE. The Marine Environmental Data Inventory (MEDI) lists the holders and characteristics of environmental databases in Member States of IOC. The NODC in the Member State can assist MEDI by collecting information on oceanographic databases held in its country, as well as its own databases and supply the information to the IOC for incorporation in MEDI.

The Cruise Summary Report (CSR) is a form used to record information about data collected from research vessels. The World Data Centres and the ICES Data Centre maintain computerized inventories of CSRs. Member States can obtain information on what data have been collected in their area of interest or indeed anywhere in the world. Arrangements can then be made on a bilateral basis or through IODE channels to receive copies of the data. NODCs should promote the use of the CSRs by the data collecting organizations in their country and then compile them and provide copies to the WDC with which they normally deal. It should be noted that the WDCs exchange data and information on a regular basis to keep their holdings alike. An NODC only needs to submit its data and information to one WDC or to an RNODC that will supply the data to the WDCs.

Software is available from ICES to assist Member States in completing and managing CSRs.

The third inventory of interest is operated by the IOC Secretariat. A national oceanographic programme (NOP) is a programme that will collect oceanographic data that a Member State intends to submit for international exchange. The NOP announcement is a description of a mission to collect certain data in a certain ocean area. The announcements are forwarded to the IOC Secretariat and are circulated by mail and posted on the World Wide Web. This serves two purposes. First of all it alerts the international exchange system about data which will be coming. Secondly, it provides opportunities for interested parties to be aware of upcoming data collection activities and to negotiate participation if the activity is of mutual interest. An NODC should promote the preparation of NOP announcements for national activities and compile and submit them to the IOC Secretariat on a regular basis.

As a service to the international community, the Data Information Unit (DIU) of the University of Delaware in the USA provides a searchable on-line database on planned research cruises. This database includes information on the cruises described on the IOC CSRs and other sources. The system may be consulted on the World Wide Web at the address given in Annex B.

(iv) Operate a Responsible National Oceanographic Data Centre

Taking on the responsibility of operating an RNODC can be a major task. However, it is essential that Member States of IOC who can find the resources make a contribution to international data management and exchange by doing so. Operating an RNODC will usually result in a centre having to acquire, process and archive data that is outside it ocean area of interest. However, there are benefits. The centre may be more successful in acquiring data for its own area of interest more quickly than in would otherwise.

The other benefit is that other Member States and their NODCs will operate other RNODCs which will collect and process other types of data. Any centre can then acquire these other types of data for its area of interest. It is basically a method of sharing the workload. Overall efficiency is increased because not every centre has to process and manage all the various data collected in its area of interest.

When looking at the possibility of operating an RNODC a centre should review its expertise and abilities. If the centre already has a strong capability in managing a particular type of data and has in place the software systems to process, quality control, and retrieve the data, then taking on the RNODC responsibility will not require a large software development and maintenance effort. The centre merely needs to cope with a larger volume of data than provided through national sources. This will reduce significantly the burden of operating the centre.

In any case a centre should consider carefully whether to take on management responsibility for a data type outside its expertise or for which it has no experience or existing capability.

(v) Assist with Data Management for an International Science Programme

Large programmes such as the World Climate Research Programme and its component experiments present another opportunity for an NODC to provide an international service. An example of this is the World Ocean Circulation Experiment (WOCE). WOCE established a data management programme based on the concept of Data Assembly Centres (DACs). A DAC is a two component entity. A data centre in co-operation with an oceanographic laboratory will form a DAC. The oceanographic laboratory provides the scientific expertise with the data and often provides analysis and quality control of the data. The data centre provides the data management expertise, archival and dissemination of the data, or more processing and QC depending on the arrangements with the scientific centre. Some complex DACs include more than one data centre and several scientific organizations. The GTSPP data flow in Figure 2 is an example of a complex DAC arrangement.

Several IODE NODCs have taken on responsibilities as part of a WOCE DAC. WOCE is only one example of NODCs supporting international experiments. JGOFS is another one but operates differently.

Support to data management for global science experiments is another function that centres should consider when deciding how best to support international science. Of course, any such support by a centre is considered as part of that countries contribution to the international experiment.

6. IN SUMMARY

Part II of this manual has provided an overview of the organization of international oceanographic data management. It has also presented ways in which NODCs can participate and contribute to this activity. It is important that a new centre participate at some level in international exchange. Otherwise it will lose opportunities for access to data that would be important to it national services and to the expertise of colleagues around the world who can contribute information, advice, and technology to the development of the centre.

ANNEX A

Publications of Interest to a National Oceanographic Data Centre

I. IOC Manuals and Guides Series

Manuals and Guides N$^{o.}$ 1
Guide to IGOSS Data Archives and Exchange (BATHY and TESAC).
Source: MAN.-GUIDES-IOC. 1993. N$^{o.}$ 1 2nd rev. ed., 26 pp
Publication year 1993
---Manuals and Guides N$^{o.}$ 3
Guide to Operational Procedures for the Collection and Exchange of IGOSS Data.
Source: MAN.-GUIDES-IOC. 1988. N$^{o.}$ 3, 72 pp
Publication year 1988
---Manuals and Guides N$^{o.}$ 5
Guide to Establishing a National Oceanographic Data Centre
Publication year 1997
--Manual and Guides N$^{o.}$ 9
Manual on International Oceanographic Data Exchange (Fifth Edition)
Publication Date 1991.
--Manuals and Guides N$^{o.}$ 14
Manual on Sea-level Measurement and Interpretation. Volume 2. Emerging Technologies
Source: MAN.-GUIDES-IOC 1994 N$^{o.}$ 14, 73 pp
Publication year 1994
--Manuals and Guides N$^{o.}$ 16
Marine Environmental Data Information Referral Catalogue.
Source: MAN.-GUIDES-IOC. 1993. N$^{o.}$ 16, 157 pp
Publication year 1993
--Manuals and Guides N$^{o.}$ 17
GF3: A General Formatting System for Geo-referenced Data, Volume 1, Introductory Guide to the GF3 Formatting System
Source: MAN.-GUIDES-IOC 1993 Vol. 1, N$^{o.}$ 17, 36 pp
Publication year 1993
--Manuals and Guides N$^{o.}$ 17
GF3: A General Formatting System for Geo-referenced Data, Volume 2, Technical Description of the GF3 Format and Code Tables
111 pp
Publication year 1987
--Manuals and Guides No. 17
GF3: A General Formatting System for Geo-referenced Data, Volume 3, The GF3 Code Tables
Publication year 1987
--Manuals and Guides No. 17
A General Formatting System for Geo-referenced Data, User Guide to the GF3-PROC Software: Volume 4.
Source: MAN.-GUIDES-IOC. 1989. N$^{o.}$ 17, 27 pp
Publication year 1989
--Manuals and Guides N$^{o.}$ 17
GF3. A General Formatting System for Geo-referenced Data. Volume 5., Reference Manual for the GF3-proc software.
Source MAN.-GUIDES-IOC 1992 N$^{o.}$ 17/5, 74 pp Report number IOC SC-92-WS-16 (SC92WS16)
Publication year 1992
--Manuals and Guides N$^{o.}$ 17

GF3. A General Formatting System for Geo-referenced Data. Volume 6., Quick Reference Sheets for GF3 and GF3-Proc.
22 pp
Publication year 1989
--Manuals and Guides N^o.^ 18
User Guide for the Exchange of Measured Wave Data.
81 pp
Publication year 1987
--Manuals and Guides N^o.^ 19
Guide to IGOSS Specialized Oceanographic Centres (SOCs)
17 pp
Publication year 1988
--Manuals and Guides N^o.^ 20
Guide to Drifting Data Buoys.
Source: MAN.-GUIDES-IOC. 1988. N^o.^ 20, 69 pp
Publication year 1988
--Manuals and Guides N^o.^ 22
GTSPP Real-time Quality Control Manual.
Source: MAN.-GUIDES-IOC. 1990. N^o.^ 22, 121 pp
Publication year 1990
--Manuals and Guides N^o.^ 24
Guide to Satellite Remote Sensing of the Marine Environment.
Source: MAN.-GUIDES-IOC. 1992. N^o.^ 24, 184 pp
Publication year 1992
--Manuals and Guides N^o.^ 25
Standard and Reference Materials for Marine Science
Source: MAN.-GUIDES-IOC 1993 N^o.^ 25 rev. ed., 579 pp
Publication year 1993
--Manuals and Guides N^o.^ 26
Manual of Quality Control Procedures for Validation of Oceanographic Data.
Source: MAN.-GUIDES-IOC. 1993. N^o.^ 26, 436 pp
Publication year 1993
--Manuals and Guides N^o.^ 29
Protocols for the Joint Global Ocean Flux Study (JGOFS) Core Measurements
Source: MAN.-GUIDES-IOC 1994 N^o.^ 29, 126 pp
Publication year 1994

II. IOC Technical Report Series

IOC Technical Series N^o.^ 42

Calculation of New Depth Equations for Expendable Bathythermographs Using a Temperature-error-free Method (application to Sippican/TSK T-7, T-6 and T-4 XBTs)
Author(s):Hanawa,-K.; Rual,-P.; Bailey,-R.; Sy,-A.; Szabados,-M. Author Affiliation Tohoku Univ., Dep. Geophys., Sendai 980, Japan
Source: TECH.-SER.-IOC 1994 N^o.^ 40, 50 pp
Publication year 1994

IOC Technical Series N$^{o.}$ 40

Oceanic Interdecadal Climate Variability
Source: TECH.-SER.-IOC 1992 N$^{o.}$ 40, 40 pp
Publication year 1992

III. The ICSU Guide to the WDC System

The ICSU Guide to the World Data Centre System
109 pages
Publication year 1996

ANNEX B

Contacting the IOC, and other International Data Sources

I. IOC Secretariat

The IOC Secretariat is based in Paris at 1, rue Miollis in the 15th arrondissement (south-west Paris).

The postal address is:

Intergovernmental Oceanographic Commission
UNESCO
1, rue Miollis
75732 Paris Cedex 15
France

The central fax number **(33) (1) 45 68 58 12**

The address of the IOC Homepage on the World Wide Web is **http://www.unesco.org/ioc/**

II. World Data Centres

(i) World Data Centre A (Oceanography)

Director, Dr. S. Levitus
NODC/NOAA E/OC5
1315 East West Highway, Room 4362
Silver Spring , MD 20910-3282
U.S.A.
Tel: (301) 713 3294
Fax: (301) 713 3303
E-mail: slevitus@nodc.noaa.gov
Data Centre URL: http://www.nodc.noaa.gov/NODC-dataexch.html

(ii) World Data Centre-B1 (Oceanography)

Deputy-Director RNIIGMI-WDC, Dr. V. I. Smirnov,
Koroleva 6, Kaluga Region
RF Obninsk 08439
RUSSIAN FEDERATION
Tel: <7> (08439) 25925
Fax: <7> (095) 255 22 25
Tlx: 412633 INFOR SU
E-mail: wdcb@adonis.iasnet.com

(iii) World Data Centre-D (Oceanography)

Director
National Marine Data & Information Service
State Oceanic Administration (SOA)
93, Liu Wei Rd, Hedong district
Tianjin 300171
PEOPLE'S REPUBLIC OF CHINA
Tel: <86> (22) 430 52 13-401
Fax: <86> (22) 430 44 08
Tlx: 23138 NODC CN
E-mail: houwt@bepc2.ihep.ac.cn

III. The Data and Information Unit (DIU)

Dr. Ferris Webster
Ocean Information Centre
Graduate College of Marine Studies
University of Delaware
700 Pilottown Road
Lewes, DE 19958
U.S.A.
Tel: (302) 645 4278
Fax: (302) 645 4007
E-mail: oceanic@diu.cms.udel.edu
WWW Server for the DIU: www.cms.udel.edu
WWW Server for ship schedules: www.cms.udel.edu/ships

ANNEX C

LIST OF ACRONYMS

BATHY	GTS code for transmitting an ocean temperature profile
CAD	Computer Assisted Design
CD-ROM	A write once-read many times optical disk for storing data and information.
CMD	Continuously Managed Database
CMM	(WMO) Commission on Marine Meteorology
CSR	Cruise Summary Report
DAC	Data Assembly Centre (element of WOCE data management system)
DBCP	Drifting Buoy Co-operation Panel
DIU	Data Information Unit (University of Delaware, USA)
DNA	Designated National Agency (for IODE)
FTP or ftp	File Transfer Protocol
GE	Group of Experts
GEBCO	General Bathymetric Chart of the Oceans
GDA	GEBCO Digital Atlas
GIPME	Global Investigation of Pollution in the Marine Environment
GIS	Geographic Information System
GLOSS	Global Sea-Level Observing System
GOOS	Global Ocean Observing System
GTS	Global Telecommunications System
GTSPP	Global Temperature-Salinity Profile Programme
ICES	International Council for the Exploration of the Seas
ICSU	International Council of Scientific Unions
I-GOOS	IOC Committee for the Global Ocean Observing System
IGOSS	Integrated Global Ocean Services System
IOC	Intergovernmental Oceanographic Commission
IODE	International Oceanographic Data and Information Exchange
ITSU	International Co-ordination Group for the Tsunami Warning System in the Pacific
JGOFS	Joint Global Ocean Flux Study
J-GOOS	Joint Scientific Committee for the Global Ocean Observing System
LAN	Local Area Network
MEDI	Marine Environmental Data Inventory

MGG	Marine Geology and Geophysics
MIM	Marine Information Management
NODC	National Oceanographic Data Centre
NOP	National Oceanographic Programme
PC	Personal Computer (usually used to describe an IBM or IBM compatible PC)
PSMSL	Permanent Service for Mean Sea-Level
QC	Quality Control
RNODC	Responsible National Oceanographic Data Centre
TESAC	GTS code for transmitting an ocean temperature, salinity, and currents profile
WAN	Wide Area Network
WDC	World Data Centre
WMO	World Meteorological Organization
WOCE	World Ocean Circulation Experiment
WWW	World Wide Web

[end of document]