International Workshop Looks at Ocean Data Management Issues

James Churigin
Chairman, Ocean Climate Data Workshop Organizing Committee

With support from the Earth System Data and Information Management Program Office, NOAA's National Oceanographic Data Center (NODC) organized the Ocean Climate Data Workshop held to promote dialogue between ocean scientists and data managers on improved data delivery systems needed for climate change research.

The primary workshop sponsor was the Intergovernmental Oceanographic Commission (IOC). Additional support came from the Commission of European Communities (CEC), the International Council for the Exploration of the Sea (ICES), the International Council of Scientific Unions (ICSU), the Scientific Committee on Oceanic Research (SCOR), and the World Meteorological Organization (WMO).

Hosted jointly by NOAA and NASA, the workshop was held February 18-21 at NASA's Goddard Space Flight Center, Greenbelt, Md. Over 100 registered participants from 18 countries attended the full 3-1/2 day workshop, while additional observers and part-time attendees brought total attendance to about 125.

Dr. John Knauss, Under Secretary for Oceans and Atmosphere, introduced the workshop with a talk on "The Constancy of the Ocean" in which he described how, in his lifetime, the old scientific paradigm that the deep ocean was unchanging has been overthrown by new evidence.

New Director Takes Helm at NODC

Bruce C. Douglas has been appointed Director of the National Oceanographic Data Center (NODC). He succeeds Gregory W. Withee, who served as NODC Director from 1985 until 1991. Withee is now Deputy Assistant Administrator for Environmental Information Services, NESDIS, and heads the NOAA-wide Earth System Data and Information Management Program.

Before coming to NODC, Douglas served as Chief of the Geosciences Laboratory within NOAA's National Ocean Service. A graduate of the University of California at Los Angeles (B.A., mathematics; M.A., astronomy), he is the author of numerous scientific papers in the fields of satellite geodesy and oceanography. He also developed NOAA's program to process and analyze altimeter data from the U.S. Navy Geodesic Satellite (Geosat). Douglas and his wife Virginia live in Bethesda, Maryland.
Ocean Data Workshop, from page 1

knowledge of ocean variability. The workshop was dedicated to the memory of the late Professor Henry Stommel.

Objectives

The workshop was organized to:
- Identify opportunities for improving data management for ocean climate research;
- Find ways to improve access to marine data;
- Outline the characteristics of data management systems needed to support ocean monitoring and prediction; and
- Provide guidelines for improved data services.

Issues and Recommendations

At the concluding session, a summary of the 10 major issues identified by the workshop was prepared for submission to the IOC Executive Committee for review and development of recommendations by the IOC and the other sponsoring organizations. Following is a synopsis of these issues and findings.

Continuing Liaison Between Data Managers and Scientists

The "case studies" presented at the workshop show that the data management systems that were working the best have a common trait: They were developed by teams of data managers and research scientists assembled in the early stages of project planning. Examples include WOCE Data Assembly Centres, the Joint Analysis Centers operating within the Global Temperature-Salinity Pilot Project (GTSP), and the Topical Centers of the Joint Global Ocean Flux Study (JGOFS). Conversely, projects that considered data management as a totally separate activity with lower priority often failed to provide the service required to meet scientific objectives. Therefore, groups within the IOC and other international organizations should be encouraged to:
- Publicize successful data/scientist collaborations to serve as models in planning for the future.
- Reduce adversarial situations where data managers and scientists appear to be in competition.
- Foster collaboration for high data quality and more timely data submission.

Importance of Historical Data

Progress in ocean climate studies will depend on filling spatial and temporal gaps in the data record. Some historical data are in danger of being lost because of deterioration in their present state, however, and an immediate rescue effort is needed. An ongoing multilateral effort known as "Data Archaeology" is working to locate and preserve historical data, but the workshop recommended that current ad hoc efforts be expanded to an international data rescue and recovery project with full support of IOC member states. It was noted that data rescue operations must recover not only the numerical values, but the data documentation (metadata) as well.

Role and Importance of World Data Centers (WDCs)

The consensus of workshop participants was that the World Data Centers play an important international role in the sharing of scientific data and information and that this role would increase in importance as global change problems grow in number and complexity. The workshop recommended that:
- The role and responsibilities of the World Data Center System be re-examined in light of present plans for climate and global change experiments.
- The WDCs for Oceanography (in the United States, Russia, and China) receive encouragement and support for their project to harmonize data holdings and to produce a unified semi-annual catalogue (possibly to be made available on an electronic bulletin board).
- The WDCs continue to promote free access to data and continue their policy of freely exchanging data. The sharing of -continued on page 3
data is of growing importance to ocean climate programs.
- The WDCs address the issue of access to satellite derived data and data products, as well as multi-disciplinary data sets.

Evolution of Data

Workshop discussions covered a full range of technical matters associated with the collection and dissemination of data and meta-data. It was recognized that many of these items will require assembling relatively small groups of experts to make specific recommendations aimed at solving particular problems. Issues include:
- Problems associated with the increasing size of data sets.
- Increasing complexity of data.
- Need for correlation of data sets across disciplinary lines.
- Technical problems associated with the storage and retrieval of satellite derived observations.
- Need to formulate an IOC strategy for development of data systems required to support an operational ocean observing system.

Participation of Developing Countries in Ocean Climate Programs

Participation by developing countries in ocean climate projects is often hampered by their inability to get access to data and data products. Computer tools demonstrated at the workshop, however, show that low cost solutions are available today. The real problem is getting hardware and software to the right places and providing adequate training to the users. Technological developments such as CD-ROM, user-friendly software, and low cost computers, along with the cooperation of IOC member states, should ease the present situation considerably.

Developing countries and regions should devise their own data management strategies to maximize technology and data access. Developing countries should also be asked to play a role in data rescue efforts.

Model Data

The workshop recognized that air-sea interaction models and forecasting models are of growing importance to ocean climate projects. These models are used by modelers and generators of data (or pseudo-data). This subject was discussed in enough detail that it is an excellent candidate for future follow-up discussions. Issues that were discussed include:
- Modelers need data input and generate data output. Output is now considered a research product but may be needed by others. Should it be archived? For how long? Where?
- Problems associated with storing model output need further discussion. Do you archive all model output? Or just selected products?
- Should model output be considered as part of a data set or complementary to it?
- How important are the meta-data that accompany model output and what should they contain?

Data Quality

The importance of data quality was a key theme in workshop talks and discussions. Many of the ongoing climate related projects (e.g. WOCE) have very high quality standards. While some modelers may have ways of filtering data of lesser quality, others require data that has been fully processed and quality assured. Among the issues discussed:
- To ensure that the best quality data are obtained without duplication of effort, careful coordination is required throughout the data acquisition process until they are made available for general dissemination.
- The Global Temperature-Salinity Pilot Project (see the article on page 4 of this issue) was cited as an excellent example of collaboration among data centers and researchers to produce a high quality data set available for the international community.
- While there is much to learn from the meteorological example, participants felt that oceanography does not have the "forecasting" base used by that community and must develop its own strategies for building data sets needed by climate change research projects.

Funding

The discussions centered on how necessary tasks would get done, and how funding could be obtained to perform them. Traditionally, oceanography has been supported by research funds. As we move toward an operational ocean observing system, how do countries receive funding for this system, while still maintaining the strong research base that will be required? Workshop participants could not answer this question, but were quite aware that it is critical to the future development of an ocean observing system.

The Global Ocean Observing System (GOOS)

While many workshop discussions and recommendations involved elements related to GOOS, several points specifically concerned that program. Speakers involved in the development of GOOS stated that a strong, effective data management program is at the heart of a successful ocean observing system. It was also pointed out that GOOS requires coordination among a number of IOC groups and between IOC and a number of other international -continued on page 8
International cooperation improves the flow of ocean data

**GTSPP: The Global Temperature-Salinity Pilot Project**

*Douglas Hamilton*
Chief, Data Processing Branch
National Oceanographic Data Center
NOAA/NESDIS

For nearly one hundred years scientists around the world have been routinely measuring profiles of ocean temperature and salinity versus depth. From temperature and salinity (T-S) conditions and their variations in time and space, researchers have been able to learn much about the ocean's circulation, chemistry, and biology. In addition, it is now clear that this one-hundred year historical record of T-S measurements is extremely valuable for studying and understanding changes in the earth's climate.

In the late 1950s and 1960s it became apparent that oceanographers of all countries would benefit from sharing available ocean temperature and salinity data. That sharing among scientists now occurs primarily through two channels.

First, real-time T-S data are transmitted over the Global Telecommunications System (GTS), which was originally established to disseminate weather data internationally. Second, the same observations are also recorded on computer tapes or diskettes on ship and later sent as "delayed mode" data to national data centers in many countries of the world. The delayed mode data are usually higher in quality and resolution than the same observations which were sent previously in real time.

Nations manage the sharing of real-time data through the Integrated Global Ocean Services System (IGOSS), and of delayed mode data through the Committee on International Oceanographic Data Exchange (IODE). Through these groups agreements are made and standards for exchange are established. Overall, the management of these two data streams has worked well. Until recently, however, there was little coordination of the two data streams, even though they originate from the same source.

The Global Temperature-Salinity Pilot Project was organized in the late 1980s by IGOSS and IODE as a combined trial project to improve the timeliness, quantity, and quality of ocean temperature and salinity data. This is being accomplished by:

- improving the flow of real-time data from GTS,
- developing and implementing standard procedures for data quality control,
- building and maintaining a "continuously managed database" (CMD) of global ocean temperature-salinity data, and
- including in the CMD data quality flags and changes resulting from analysis of the data by scientists in research centers around the world.

GTSPP operations were initiated in January 1991 when real-time ocean temperature and salinity data began flowing to the U.S. National Oceanographic Data Center (NODC) from Canada's Marine Environmental Data Service (MEDS). MEDS provides initial quality control for the real-time data. NODC provides quality control for the delayed mode data and merges the incoming data into the database (see figure 1).

NODC divides the global data stream into three files, one for each of the major ocean basins: Atlantic, Pacific, and Indian. Each month, data for each of the ocean basins is transmitted via computer network to one of three

---

**Figure 1.** Real-time data transmitted over the Global Telecommunications System (GTS) are assembled and quality controlled by the Marine Environmental Data Service, Canada, which relays them to the U.S. NODC over a computer network. The NODC processes and quality controls delayed mode data and maintains the GTSPP database. [Note: NODCs = National Oceanographic Data Centers, RNODCs = Responsible National Oceanographic Data Centers, WDCs = World Data Centers]
science centers responsible for scientific analysis and quality control of the data. These centers are:
- **Atlantic Ocean**: NOAA Atlantic Oceanographic and Meteorological Laboratory, Miami Florida,
- **Pacific Ocean**: Scripps Institution of Oceanography, La Jolla, California,
- **Indian Ocean**: Commonwealth Scientific and Industrial Research Organization Division of Oceanography, Hobart, Tasmania, Australia.

In addition, the data are transmitted to the TOGA Subsurface Thermal Data Center, Brest France, which also maintains a global TS data file. The real-time data received each day are also immediately forwarded to the NOAA's Climate Analysis Center where they are used in an upper ocean temperature model.

As of May 1992, the GTSP data base held 59,727 real-time profiles. The NODC fulfills requests for these data and recently provided selected GTSPP data to researchers at the University of Maryland and the University of Colorado. In the future, the delayed-mode data submitted to NODC will be merged into the database.

GTSPP is a data management project in which development and operations are ongoing at the same time. Even with operations underway, we have continued to develop and enhance systems to accept and check the data, create a database, set up data communications, and develop working procedures with MEDS and the three science centers.

We have learned some lessons in the process of implementing GTSPP at the NODC. In distributed data management projects such as this, it is essential to:
- clearly define the contents, structures, and relationships of the data to be managed;
- discuss as many details as possible to ensure that all participants understand their responsibilities; and
- discuss and agree on those data fields that uniquely define one record and establish rules for placing or modifying data in those fields. We learned recently, for example, that what we thought was a “unique” station number occasionally is not.

Through GTSPP we hope to improve existing data management systems at NODC. We will soon transfer the GTSP database to the database machine system acquired for project Poseidon, NODC's effort to develop an integrated management/access system for its data archives. Over the next year we also plan to move much of our processing from mainframe to workstation computers. ■
NOAA/EPA Library Catalog Available Online

The holdings of NOAA and Environmental Protection Agency (EPA) libraries can now be searched via personal computer access to NOAALINC, the NOAA Library and Information Network Catalog. NOAALINC is a CD-ROM database of bibliographic records for more than one million books, technical reports, serials, and microforms.

The database covers most of the holdings of NOAA and EPA libraries throughout the United States. These collections are strong in aquatic science, marine ecology, microbiology, oceanography, meteorology, climatology, chemistry, geology, remote sensing, hazardous wastes, pesticides, toxicology, pollution, and law.

Just set your terminal emulation to ANSI. Set the baud rate as high as it will go up to 9600. Your communications protocol is 8 bits, 1 stop bit, no parity. It is best to use a communications software program that does not remap your keyboard.

To access NOAALINC, dial 301-770-7375 if you are in the Washington, D.C. area or 800-352-7281 for the rest of the country. Press <Enter> at least twice to establish a "handshake". This gives NOAALINC an opportunity to recognize your call and determine the speed at which the communication is taking place. Type: ANSI as the User ID. To search NOAALINC, simply follow the instructions on the screen.

What the database contains

The NOAALINC database contains bibliographic information on books, technical reports, technical memoranda and serials. There is title, author, subject and keyword access to each record. It also indicates which NOAA or EPA library holds the item.

What the database does not contain

The database does not hold contents pages of journals, abstracts, or chapters of books. To find these, you should search the various periodical indexes located in the libraries.

Why use the database

The database is useful for comprehensive searches as well as for browsing a subject area to see what has been published. This database is also a good source for government documents that sometimes are not well covered in other databases. These include reports from various NOAA regional centers and offices that are not widely distributed. NOAA employees holding reports not already in the database can have them added if they furnish at least one copy to the NOAA Central Library.

How to search

The database is easy to use. Just follow the instructions on the screen. You may search by keyword, phrases or combined sets of words (two words combined with an ampersand, e.g., CLAMS & GEORGIA). Use your communications software function keys to either print any record or save any record to your computer. NOAALINC workstations are also available at a number of NOAA and EPA libraries and information centers. A listing of locations is available from your nearest NOAA or EPA library.

To obtain the books and technical reports listed in NOAALINC, have your librarian request them by interlibrary loan or call either the Central Library (301-443-8332 or FTS 443-8332) or one of the NOAA Regional Libraries (Seattle at 206-526-392-6241 or FTS 392-6241; Miami at 305-361-1429 or FTS 350-1429). For technical questions about using NOAALINC, please call Lynda Kuntz of the NOAA Central Library staff at 301-443-8288.

- Lynda Kuntz
NOAA Central Library
NOAA/NESDIS E/O/C4
Washington Science Center, Bldg. 4
Rockville, MD 20852

Upcoming Conference on Applications of Satellite Remote Sensing by Developing Countries


One of the major conference themes is management of the resources of the Earth and its environment. The conference will discuss the role of satellite remote sensing in resource management, environmental assessment, and global change studies. Panels of experts from the developing world will address their particular needs concerning satellite remote sensing for the management of resources and environment, as well as for global change studies. The intent is to increase awareness of opportunities for participation in space activities by developing countries.

Although the conference is open to all interested parties, priority will be given to participants from developing countries. It is of special interest to professionals from developing countries representing the following disciplines: management of natural resources; environmental monitoring, assessment, and protection; global change studies; policy making in government agencies related to environmental monitoring and remote sensing; and satellite remote sensing technology and applications.

For further information or to apply to participate, contact: Betty Howard, NOAA/NESDIS International Affairs, Federal Building 4, Room 0110, Washington, DC 20233. Telephone: 301-736-4588. FAX: 301-736-5828.
NOAA Symposium on Scientific Workstations

Putting new technology to work at the "Earth System Agency"

Robert H. Kidwell
Chief, Information Resources Management Staff
NOAA

In April the Information Resources Management Staff of NOAA's Information Systems and Finance Office sponsored a three and one-half day Symposium on Scientific Workstations. Because NOAA's scientific computing requirements will increasingly be met through use of networked workstations, this technology is expected to have a tremendous impact on NOAA's computational infrastructure.

With this in mind, the Symposium was organized to:
- provide the opportunity for NOAA scientists to learn about current technology and how it will develop in the future,
- spark the creation of people-networks of scientists who share common work interests and face common technical problems, and
- elicit feedback from the participants - through informal comments and communications, through discussion, and through the working-group process - to provide a foundation for further action on an agency-wide basis.

The Symposium opened with a half-day tutorial on scientific workstation technology to give Symposium participants the necessary background to understand the issues involved.

The following three full days were thematically organized. Day 1 was an in-depth exploration of "What's Happening in NOAA" with regard to workstation technology. Day 2 explored "What's Happening in Industry," and Day 3 was devoted to an attempt to reach a consensus on future agency action with regard to the new technology.

What's Happening in NOAA?

Day 1 was kicked off with a keynote talk by Ray Kammer, Deputy Under Secretary for Oceans and Atmosphere. Mr. Kammer functions as NOAA's chief operating officer and oversees the day-to-day operation of NOAA. His remarks focused on NOAA's role as the "Earth System Agency" which must be able to collect, organize, analyze, visualize, and disseminate large quantities of data. To do this NOAA must maximize its use of information technology, modernize its computing infrastructure, and ensure that it can inter-operate. He stressed the importance of organizing NOAA's cross-cutting programs so that NOAA can work together to meet this mission efficiently.

Mr. Kammer's keynote remarks were followed by some two dozen different presentations representing each of NOAA's line offices.

What's Happening in Industry?

Day 2 was led off by Don Brown, a well-known industry analyst. He gave a 90-minute presentation on the workstation marketplace entitled Crisis in the Mini/Micro Market. This was followed by two presentations, Open System Standards: A Perspective, and Overview of Graphics Standards, given by representatives from the National Institute of Standards and Technology. The first clarified the concept of Open System Environments and how specific FIPS (Federal Information Processing Standards) flow from such a concept. The second presentation clarified the alphabet soup of graphics standards (GKS, PHIGS, CGM, IGES, CG1). Copies of the viewgraphs from these presentations were requested by a number of Symposium participants. The viewgraphs and other handouts will be mailed to all participants.

The rest of Day 2 was devoted to vendor demonstrations and presentations. Unlike trade shows, which often feature hundreds of vendors, the Symposium featured a small number of representative vendors, each of which was invited to maintain an exhibit. In addition, each vendor made a one-hour presentation on a topic (or several) of general interest to Symposium participants.

What Should We Be Doing?

The final day of the Symposium was devoted to planning NOAA's response to the developing technology. It opened with a panel discussion on Lessons Learned: What Other Agencies Have Done - NASA, EPA, NSA, and USGS. Four distinguished speakers told of their own agencies' experiences with workstation procurements, acquisition, implementation, training, and maintenance.

After this stimulating and provocative discussion, participants turned to the focal point of the last day: the working group session. Two days earlier, in his keynote remarks, Deputy Under Secretary Ray Kammer was very supportive regarding the need for funding to accomplish modernization. But he emphasized that NOAA management requires the help of NOAA's scientists and scientific managers to formulate the necessary technical plans. He challenged the group to develop a written proposal for agency action and promised to review and respond to it personally. The Symposium was planned to culminate with a working group session that would be the genesis of such a proposal for future NOAA-wide action.

The working group process began with a "brainstorming" session. The goal was to collect as many ideas for action as possible. The ground rules were simple:

-continued on page 8
all ideas were welcome, and it was recognized that everyone was speaking for themselves and not necessarily their organization. This prompted a lively discussion that resulted in more than 30 ideas being recorded. During the lunch break these ideas were consolidated into a list of specific areas in which NOAA-wide action would be appropriate.

During the afternoon session, these areas were discussed and more clearly defined. In an attempt to determine if a consensus existed, the session ended with the Symposium participants voting for their three highest priority actions.

The results proved interesting. The need for NOAA-wide planning and management of networks is clearly a high priority. Fortunately, work is going on in this area through the NOAA Network Advisory Review Board (NARB). [See the article on the NARB in the February 1992 issue of the Earth System Monitor.]

Another priority area is improving the communication and sharing of information within NOAA. One specific suggestion was to establish a NOAA electronic bulletin board through which information on technical matters such as scientific workstations could be rapidly disseminated. The NOAA IRM Staff is looking into the feasibility of setting up a bulletin board. Their initial idea is to operate the board on the VAX 6520 located at Suitland, Maryland. They would welcome other suggestions.

Many participants indicated through their votes that a centralized source for information on procuring workstations is needed. This would include information on existing contracts that can be used, sample RFPs (Requests for Proposal), and samples of required IRM documentation. The IRM Staff plans to be a repository of this type of information. NOAA organizations planning to buy workstations can call the IRM Staff to get information and advice.

Many other areas in which action is desirable were identified. These will be included in a Symposium summary and proposed action plan, along with specific proposals, that is being prepared for the Deputy Under Secretary. Some of the potential action areas are: 1) establish block contracts for workstations, 2) prepare site licenses for workstation software, 3) set up a NOAA sponsored workstation training program, and 4) create a NOAA "SWAT" team to provide technical advice to NOAA organizations planning for and installing networks of workstations.

The Symposium accomplished its goal of being a "kick-off" event for coordinated NOAA action. For more information or if you have suggestions, call the IRM Staff at 301-443-3478.

Ocean Data Workshop, from page 3
organizations such as WMO, SCOR, ICSU, and the United Nations Environmental Program (UNEP).

Communications
Workshop participants expressed concern over the adequacy of communication networks as required by both present research programs and potential monitoring activities and suggested further study of specific technical and policy issues. They agreed that the electronic mail system used widely by the oceanographic community has been, and will continue to be, an essential part of the international communication system.

Conclusion
This workshop differed markedly from those where a data manager was invited to a science meeting, or where a scientist was invited to a data meeting. The balance between scientists and data managers at this workshop resulted in interactions that were mutually beneficial to most participants. Although it was recommended that another workshop be held in two or three years, it was felt that too many topics were covered in this first meeting and that the next one should be more narrowly focused to enable more specific recommendations to be developed.