INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION
(of UNESCO)

Eighteenth Session of the IOC Committee on International Oceanographic Data and Information
Exchange (IODE-XVIII)
Oostende, Belgium, 26-30 April 2005

IODE NATIONAL REPORT ON OCEANOGRAPHIC
DATA MANAGEMENT AND EXCHANGE
FOR THE UNITED STATES
IODE NATIONAL REPORT ON OCEANOGRAPHIC DATA MANAGEMENT AND EXCHANGE FOR THE UNITED STATES

1. **Name of Data Centre:** U.S. National Oceanographic Data Center

2. **National IODE Coordinator:**
   - Name: Kurt Schnebele
   - Address: National Oceanic & Atmospheric Administration
     National Oceanographic Data Center
     1315 East-West Highway, OC1
     Silver Spring, MD 20910
   - Tel: 1-301-713-3290 x194
   - Fax: 1-301-713-3303
   - E-mail: Kurt.J.Schnebele@noaa.gov

3. **Data Center Address:** Same as above

4. **Data Center URL:** [http://www.nodc.noaa.gov/](http://www.nodc.noaa.gov/)

5. **IODE Data Center Designation Date:** 1961

6. **Description of national data flow:**

   How does data flow operate in your country (if possible illustrate by means of one or more diagrams)? This should cover:

   1. **Metadata management:**
      - At the discovery level (e.g. do you contribute to IOC/IODE MEDI, GCMD, EDMED, another system, none?)

      The U.S. NODC currently has 188 products and datasets identified in the Global Change Master Directory. In addition, the US NODC Ocean Access System (OAS) provides online search and retrieval capabilities for more than 20,000 ocean data sets/data collections archived at the US NODC. Data accessed using the OAS is typically in the “originator’s format” or as received by the NODC. The NODC maintains minimal discovery metadata for each new data collection in the Accession Tracking Data Base (ATDB), which is searched using the OAS interface. Additional descriptive metadata for each archived data collection are maintained in the NODC Metadata Manager Repository database. Future planning is to create public-search access capability.
At the Cruise level (e.g. do you contribute to IOC/IODE Cruise Summary Reports (ROSCOPs), other in-house system, none)

The U.S. NODC itself does not create or manage ROSCOPs but receives completed forms from various U.S. academic institutions via the World Data Center for Oceanography, Silver Spring. These forms are forwarded to the International Council for the Exploration of the Sea (ICES), which are included in the ROSCOP Database.

For monitoring/operational systems (e.g. EDIOS, regional GOOS systems, etc)

The U.S. NODC provides metadata management support to a number of operational programs in addition to archiving historical data from their observing systems. Projects supported by the U.S. NODC include: Acoustic Doppler Current Profiler (ADCP) program, GTSPP, Argo, the Marine Environmental Buoy program, the Joint Archive for Sea Level, Coastal Ocean Time Series, Coral Reef Information Service (CoRIS), Pathfinder SST reprocessing, shipboard NOAA underway data, Gulf of Maine Ocean Observing System (GoMOOS), and the Central and Northern California Ocean Observing System (CeNCOOS).

2. Data tracking:

What systems are in place to track data from collecting organizations through data dissemination?

The NODC Archive Management System (AMS) provides a mechanism for tracking and maintaining copies of original data and metadata acquired by the US NODC. Digital and analog oceanographic observation data and metadata are sent to the U.S. NODC in a variety of data formats using FTP, CD-ROM, DVD, and other transfer media. Data obtained in manuscript form via the GODAR and World Ocean Database projects are digitized on site. All data acquired by NODC and products created by NODC are accessioned into the U.S. NODC archives.

The diagram below illustrates the data flow within the NODC AMS. A dataset is referred to as an Accession or Submission Information Package. A description of each dataset is created within the NODC Accession Tracking Database (ATDB), where it is assigned an accession number and where basic metadata and internal data management information are captured. For each new accession, a generic file directory structure is automatically created in a temporary storage area. Original data and metadata files are placed in the “data/0-data” subdirectory, while NODC-created information are placed in the “about/” subdirectory. When the ATDB record is complete and approved, each file in the storage area is automatically checked for viruses, data-integrity checks are calculated, and the accession is transferred to the permanent archive. If it is necessary to update an accession, the entire accessioned dataset (all files), may be checked out of the storage area.
archive, updated, and then re-archived as a new version with the same Accession Number.

An automated process exports ATDB metadata elements to an XML file that conforms to the FGDC Content Standard for Digital Geospatial Metadata (CSDGM), where additional descriptive metadata may be added. All archived data may be searched by the general public online at http://www.nodc.noaa.gov/Archive/Search/. Searches can be performed using nearly two dozen parameters, including data submitting and collecting institutions. Data that are identified as relevant to the information needs of the search can be downloaded as a whole or as individual files. In many cases, an NODC Standard Product that is processed to a common data format is more suitable than multiple originator data sets in a variety of formats. The NODC User Services Team can assist customers with making the best choice of data for a particular need.

7. What is the structure of marine data management in your country:

1. How many organizations are involved?

Thus far, over 300 U.S. federal, state, and local organizations, universities, and commercial research organizations have provided data or other products to the U.S. NODC.
2. Who does what?

The U.S. NODC consists of five (5) divisions: the Coastal Ocean Laboratory (COL), the National Coastal Data Development Center (NCDDC), the Library and Information Services Division (LISD), the Ocean Climate Laboratory (OCL), and the Information Systems and Management Division (ISMD).

The COL participates in the acquisition, processing and ingesting of global oceanographic data into the NODC ocean data archives and metadata databases. It develops and maintains online product databases and other derived products (e.g., GTSSP, Argo, Shipboard Sensor Database, NOAA Marine Environmental Buoy Data) for use by diverse stakeholders. The COL performs research and develops new methods and techniques for using coastal physical, biological, and satellite-sensed oceanographic data to determine the role of various processes within the world coastal ocean system. The COL is active in efforts to prepare for ocean data from the next generation of NOAA operational satellites and in international efforts to generate new merged satellite products. The COL defines and develops new NODC coastal ocean data and information products, provides customer services for all NODC products and coordinates most NODC outreach activities.

The NCDDC mission is to support ecosystem stewardship by providing access to the nation's coastal data resources. NCDDC achieves this capability through the integration of diverse coastal data distributed in multiple repositories and provides these data to users via the Internet using established and emerging technologies. We accomplish this by maintaining a searchable metadata catalog of coastal data, developing gateways to data repositories and using middleware technology that provides data in user specified formats.

The LISD provides library services for all NOAA components and special programs, such as Ocean Exploration, Marine Protected Areas, Coral Reefs, etc. The LISD acquires and catalogues information resources, provides reference and circulation services, and manages diverse print, photographic, audio-visual, and electronic resources. The LISD coordinates the activities of 2 regional and 28 field offices. The Central Library in Silver Spring, Maryland, maintains an extensive collection of historical Coast and Geodetic Survey materials (from 1807), US Weather Bureau and foreign meteorological data materials (from the 1830's), and a collection of maps and rare books (from 1485).

The objectives of the OCL are to improve the quality of the U.S. NODC's oceanographic data archives by using the data to perform scientific analyses; develop improved ocean climatologies for annual, seasonal, and monthly compositing periods; investigate interannual-to-decadal ocean climate variability using historical oceanographic data; build scientifically, quality-controlled global oceanographic databases; and facilitate the international exchange of oceanographic data through the auspices of the World Data Center for Oceanography, Silver Spring. OCL also directs the international Global Oceanographic Data Archaeology and Rescue
(GODAR) project and the World Ocean Database project for the Intergovernmental Oceanographic Commission.

The ISMD provides the IT infrastructure backbone to support acquisition, ingest processing, quality control and long-term preservation of oceanographic data at the U.S. NODC. The ISMD develops processes to assure that data are scanned for viruses; cryptographic checksums are generated and stored with the original data files so data integrity can be monitored and verified over extended time periods and across generations of storage technologies; and copies of accessioned data are written to near-line mass storage and removable media for off-site storage.

3. What data goes where?

The U.S. NODC manages the world's largest collection of publicly available oceanographic data, which includes in situ and remotely sensed physical, chemical, and biological oceanographic data from coastal and deep ocean areas.

Original copies of new data acquired by the NODC are archived in the centrally-maintained Archive Management System (AMS, described above). Data and metadata about original data sets and NODC products are added to the AMS in a distributed fashion by personnel from the Coastal Ocean Laboratory, Ocean Climate Laboratory, and the Information Systems Management Division.

In situ profile data, particularly temperature, salinity, and nutrients profile data may be processed into one or more of the product databases, i.e., World Ocean Database, GTSP, and Argo and into analysis products such as the World Ocean Atlas. Satellite observation data may be processed into one or more products, e.g., 4 km AVHRR Pathfinder Version 5.0 SST product and the GODAE High-Resolution SST Pilot Project’s Long-Term Stewardship and Reanalysis Facility. Some non-profile in situ data types, e.g., current meter observations, oceanographic/meteorological buoy observations, are processed into the NODC Time Series Database or Marine Environmental Buoy database. Coral reef data, which may include quantitative and descriptive biology of coral reef taxa and communities (measurements of occurrences, life stages, pathology, productivity) and non-biological measurements (sea water chemistry, turbidity, temperature, salinity, currents, bio-optical data, and substratum features) are archived and included in the Coral Reef Information System (CoRIS), a NOAA project to provide a centralized clearinghouse for coral data and information resources within NOAA.

Data that do not fit into one or more of the many products produced by the NODC are archived and maintained ‘as is’ in the AMS file management subsystem. These data, while not ‘processed’ by NODC in the traditional sense of value-added quality assurance and/or quality control analyses, are described, maintained and made available to the public though the Ocean Archive System.

The LISD maintains a research collection of more than 1.5 million documents and Internet links to scientific, technical, and legislative information with a
comprehensive coverage of hydrographic surveying (from 1820); oceanography, meteorology, and hydrology (from 1870); living marine resources (from 1970 with selected coverage from 1870); and meteorological satellite applications (from 1960). Historical documents on the history of sciences date from 1485.

The NCDDC provides coastal communities with development services for coastal data as the Center implements the mission of coastal data stewardship. An important feature of the NCDDC services is the development of delivery tools merged into the data streams. Examples are: (1) the Harmful Algal Blooms Observing System (HABSOS); (2) the Coastal Risk Atlas (CRA); and (3) the Coastal Ocean Observing System (COOS). These examples demonstrate the unique integration of data access and data service. Data partnerships have been developed with customers and suppliers by maintaining these effective data services as a developmental process. The Coastal Ecosystem Program has emerged as a major theme from an initial project in support of regional Gulf of Mexico biological researchers. Elements of Coastal Ecosystem services include data and mapping services for hypoxia data, an online data mapping service and gateway access to significant ecosystem data resources. Electronic collaborations provided a tool for implementing a comprehensive data support service the NOAA Ocean Exploration Program. A Digital Atlas was the first assignment by the Ocean Exploration office as a developmental action item that has been adopted as the standard for Ocean Exploration data services.

The OCL manages the following projects and databases: (1) the World Ocean Database (WOD) series continues to be updated with historical and modern oceanographic data. The WOD01, released in March 2002, includes new variables and data from additional instruments, such as autonomous pinniped bathythermograph (instrumented marine mammals), drifting buoys, profiling floats, and undulating oceanographic recorders; (2) the World Ocean Atlas (WOA) series also continues to be updated. The WOA01 contains ASCII data of statistics and objectively analyzed fields for one-degree and five-degree squares generated from the WOD01 observed and standard level flagged data. The ocean variables included in the atlas are: in-situ temperature, salinity, dissolved oxygen, apparent oxygen utilization, percent oxygen saturation, dissolved inorganic nutrients (phosphate, nitrate, and silicate), chlorophyll at standard depth levels, and plankton biomass sampled from 0 - 200 meters; (3) the Global Ocean Data Archaeology and Rescue (GODAR) and World Ocean Database (WOD) projects, which continue to increase the volume of historical and modern oceanographic data that can be made available to researchers; (4) the OCL also supports World Data Center for Oceanography, Silver Spring (WDC), which is operated under the auspices of the U.S. Academy of Sciences, and is one of three WDCs for Oceanography.

4. Are there data for which there is no home?

NODC has the capability to archive a variety of data but are not being well handled, i.e., underwater video, biogeographical species level data.
5. What gets passed on to other organizations?

Oceanographic data archived by the U.S. National Oceanographic Data Center and the World Data Center for Oceanography, Silver Spring are, in nearly all cases, available to a diverse clientele of users in any country without restriction or limitation. In rare instances, data may only be obtained from the WDC at the request of the IOC or ICSU.

As participants in the World Data Center system, data contributed to the NODC/WDC are assumed to have no restrictions or limitations and will be widely available for redistribution. The NODC/WDC may refuse to archive data contributions that contain copyrighted materials, proprietary software, or that contain other distribution constraints or limitations.

6. What regional links and data centers are there?

Within the U.S. NODC organization, we operate the following regional sites: an Eastern Liaison in Charleston, South Carolina, which is responsible for the entire East Coast of the U.S.; a Western Liaison in Seattle, Washington, which is responsible for the entire West Coast of the U.S. as well as Alaska; a Gulf of Mexico Liaison in Stennis Space Center, Mississippi, which works closely with organizations in Texas, Louisiana, Mississippi, and Alabama; and a Pacific Islands Liaison in Honolulu, Hawaii, which is responsible for Hawaii and the U.S.-affiliated islands in the Pacific.

The NOAA Central Library is networked to 29 regional and field libraries and has access to their specialized collections.

8. What are the strengths and problems of the present arrangements nationally, regionally and internationally?

Strengths:
(1) Mechanisms do exist to plan and implement increasing international data interoperability by leveraging IODE, GOOS, GEOSS, GODAE, etc.

Problems:
(1) Loosely federated organizations – No coherent data management and communications strategy for effectively integrating the wide variety of complex marine environmental measurements and observations across disciplines, institutions, and temporal and spatial scales.
(2) Data are heterogeneous – Classes of data range from huge satellite track records to multi-dimensional model outputs to Lagrangian drifters to polygonal geographic regions and point measurements. Variables are from diverse disciplines and are unevenly distributed in time and space.
(3) There are large distributed holdings – Data cannot be centrally managed at a single location; data must reside and be managed at many distinct locations (some of which contain vast volumes of data).
9. What improvements could be made nationally, regionally and internationally?

**Background:** A Data Management and Communications (DMAC) Steering Committee was established in the spring of 2002 by the Ocean.US Office to develop a detailed, phased implementation plan that will lead to an effective data management and communications component of the Integrated Ocean Observing System (IOOS), which is the U.S. contribution to the international Global Ocean Observing System (GOOS) and to the Global Earth Observing System of Systems (GEOSS). Ocean.US is the national office for planning and coordinating the development of IOOS. This plan, reviewed at state, national, and international levels, by private industry and academia, by technical and scientific experts as well as by the broader marine environmental data supplier and user communities, was published in May 2004 (http://dmac.ocean.us/dacsc/imp_plan.jsp). This plan focuses on enhancing the interoperability of existing IOOS components though development of a common Data Communications Infrastructure. The infrastructure will consist of standards and protocols for metadata, data discovery, transport, on-line browse, and long-term archive.

The recommendations in the DMAC Plan are being implemented at national, regional, and international levels to address the problems identified in Item 8 above. Recommendations include: participation, at all levels, in a community-based development of interoperable data and metadata standards and adoption of these standards to support data discovery by users, data delivery to users, and ultimately the ingestion of new data from data suppliers (users).

10. What future national activities are planned?

Implementation of the IOOS/DMAC Plan is underway to build an initial operating capability. Major activities leading up to this capability will include: the establishment of accepted community-based data and metadata standards, best-management practices, and appropriate software tools; enhancement of hardware and software for networking, data assembly and archival; and pilot projects to demonstrate capabilities, test new technologies, and establish interoperability among existing components of IOOS.

The immediate priorities for DMAC are:

1. Establishing a DMAC Steering Team to help guide the development and adoption of data and metadata management standards and best practices.
2. Initiating technical working groups, under the Steering Team, to address metadata/data discovery; data transport; data archive; quality assurance/quality control; real-time data, data assembly, systems engineering, and other areas.
3. Engaging system engineering services to initiate development of well-organized documentation, centralized coordination of assistance to IOOS data suppliers and product generators, and software life-cycle planning for critical DMAC components.
4. Data suppliers, who wish to begin participation in IOOS immediately, should adopt the initial DMAC Plan recommendations: a) create FGDC (ISO) compliant metadata; b)
enable data discovery by sharing metadata with designated IOOS metadata facilities; c) make data available through the OPeNDAP data access protocol and/or participate in enterprise GIS; d) make data browseable on-line through the Live Access Server; and e) ensure that designated IOOS archive centers have plans in place for long-term archiving of the contributed data.

11. What national, regional or international projects is your NODC involved in (both IODE and non-IODE). Examples: Argo, GTSP, EDMED, EDIOS, Sea-Search, GODAR, etc.)

The U.S. NODC is involved in a variety of national, regional, and international projects including Acoustic Doppler Current Profiler (ADCP) program, GTSP, Argo, the Marine Environmental Buoy program, the Joint Archive for Sea Level, Coastal Ocean Time Series, Pathfinder, shipboard underway data, Gulf of Maine Ocean Observing System (GoMOOS), Central and Northern California Ocean Observing System (CeNCOOS), the Global Oceanographic Data Archaeology and Rescue (GODAR) project, and the World Ocean Database Project.