Vol. 14, No. 1 ● Fall 2003

EARTH SYSTEM MONITOR

NOAA support for global sea level data rescue

International sea level data rescue

Patrick Caldwell
National Coastal Data Development Center
NOAA/NESDIS/NODC

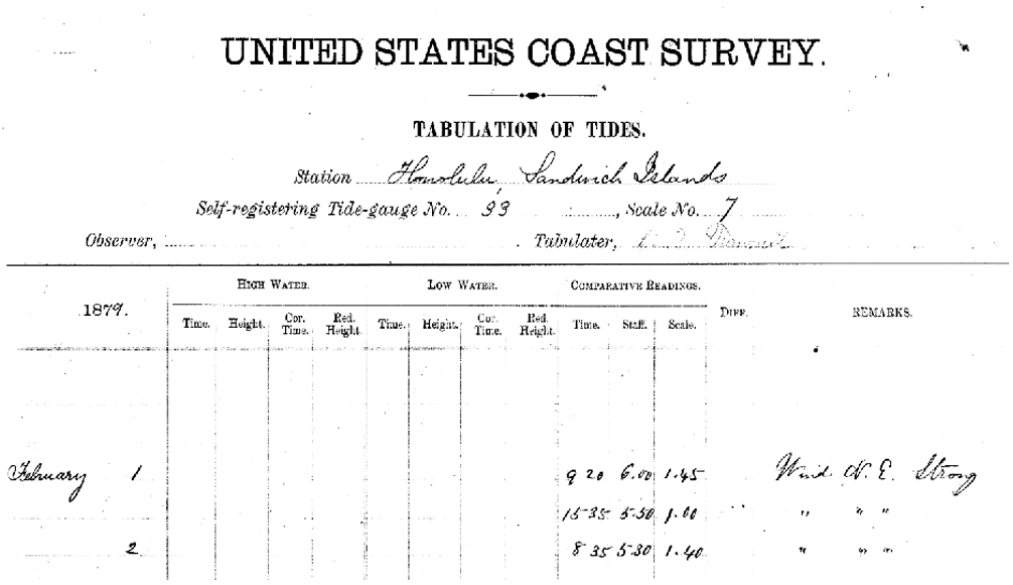
A guide to NOAA's data and information services

INSIDE

Data rescue has gained greater attention over the last decade as scientists became more interested in understanding low frequency environmental variations such as climate change. Historic data residing on dilapidating media hold valuable clues to past environmental conditions yet are at risk of loss and/or neglect. The international Global Oceanographic Data Archaeology and Rescue (GODAR) Project, initiated by the National Oceanographic Data Center (NODC) and World Data Center-Silver Spring for Oceanography (WDC) and subsequently endorsed by the Intergovernmental Oceanographic Commission (IOC), has resulted in an increase of over three million historical ocean temperature profiles, 140,000 chlorophyll profiles, 1,400,000 plankton observations, as well as many other data (IOC, 2003). As a result of recommendations made at the International GODAR Review Conference (IOC, 2003), sea level data have also been targeted for rescue.

The number of tide gauges in operation at worldwide ports increased steadily from the mid 1800s to mid 1900s in support of hydrographic objectives such as tide tables and navigational charts. The instrumentation and data collection methodology were simple and some locations still utilize a similar station configuration. The historic, standard tide gauge consists of an analog recording mechanism linked to a gear on which a cable is tied to a float that moves freely up an down with the changing water level in a securely-mounted stilling well. An essential element of each station is a tide staff, which is firmly mounted nearby the gauge and is linked geodetically to fixed benchmarks in the surrounding region. Quasi-daily observations of the tide staff are recorded along with the corresponding tide gauge measurement. The comparative readings (Figure 1) form the basis for calibration of the gauge values to a fixed datum. The analog charts were historically digitized at hourly intervals into tables stored on paper. The paper collections of the analog charts, hourly tabulations, tide staff comparative readings, geodetic survey reductions, and miscellaneous technical information regarding the station installation and maintenance, form the complete tide station data and metadata set. A significant amount of such data remains in paper form. It is desirable to make these records computer-ready to allow a more

— *continued on page 2*



UNITED STATES COAST SURVEY.

TABULATION OF TIDES.

Station Honolulu, Sandwich Islands

Self-registering Tide-gauge No. 93 , Scale No. 1

Observer, . Tabulater,

1879.	High Water. Time.	Height.	Cor. Time.	Red. Height.	Low Water. Time.	Height.	Cor. Time.	Red. Height.	Comparative Readings. Time.	Staff.	Scale.	Diff.	Remarks.
February 1									9 20	6.00	1.45		Wind N.E. Strong
									15 35	5 50	1.60		" " "
2									8 35	5 30	1.40		" " "

▲ **Figure 1.** An example of a comparative reading sheet obtained from the NOS archives.

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration

Sea level data rescue, *from page 1*
modern means of quality control and analysis, provide an additional backup for posterity, and facilitate data exchange.

The NODC and the National Coastal Data Development Center (NCDDC) under the auspices of the WDC collaborate with various international agencies and programs to increase the quality and quantity and facilitate the exchange of science-ready sea level data. This effort is lead by the Joint Archive for Sea Level (JASL), a partnership between the NODC, NCDDC, and the University of Hawaii (UH). The JASL was formed in 1987 to bridge the WDC to the Tropical Ocean Global Atmosphere (TOGA) Sea Level Center (SLC), which was directed by Dr. Klaus Wyrtki. The TOGA SLC grew into the UHSLC in the 1990s and the focus expanded from the tropics (Caldwell, 1992) toward the poles in support of various international science programs such as the joint US/France TOPEX/Poseidon sea surface topography experiment and the World Ocean Circulation Experiment. Contributions of hourly sea level data to the JASL have been provided by over 60 agencies representing over 70 countries. A vast majority of the data are received electronically although some have arrived in paper form. The JASL research quality data set is made readily available via the Internet (*http://uhslc.soest.hawaii.edu*) and is annually contributed to key international sea level centers (WDC, the Permanent Service for Mean Sea Level (PSMSL), and the Climate Variability and Predictability Project (CLIVAR) Delayed-Mode Sea Level Data Center, co-located at the British Oceanographic Data Center (BODC)). As of July 2003, the JASL holds 9,213 years of quality-assured, documented data from 412 locations worldwide.

Within the international community, the Global Sea Level Observing System (GLOSS) is a focal point of collaboration among data collection agencies, hydrographic offices, data centers, and scientific research institutes. GLOSS is supported by the Joint Technical Commission for Oceanography and Marine Meteorology of the World Meteorological Organisation and the IOC. At the GLOSS Group of Experts meeting in May 1999 (GE-6), support was dedicated to the growing GODAR effort, and subsequently, a GLOSS Archeology Project was initiated. A letter seeking information on the availability of historic data sets was prepared by Dr. Lesley Rickards of the BODC and sent on behalf of GLOSS to sea level data collection agencies and data centers. The JASL is noted in the letter as the primary center for digitizing hourly tabulations. This effort has led to the release of several data sets. In support of GODAR under funding from the NOAA Atlantic CLIVAR program, Dr. Gary Mitchum, an active GLOSS participant and professor at the University of South Florida, produced a list in 2002 on worldwide stations known to have had or still maintain tide gauges. This unpublished summary was compared to lists of existing holdings of the international data centers to highlight potentially available data sets. The JASL has utilized this report to augment data discovery.

The JASL has been one of the more active international sea level data centers that have volunteered for the arduous task of digitizing the paper records. Data rescue at the JASL has been ongoing for over a decade. The JASL is co-located with the UHSLC; thus, a pool of UH students are accessible for key entry. Upon digitization, the JASL performs quality control, assessment, and preparation of the data for the posterity and online archives. The amount of hourly data rescued to date totals 372 years from 34 locations (Table, *on page 4*). That represents over three million hourly values.

The paper records have been received from a variety of sources as shown on the Table. During a technical visit to South and Central America in 1991, the author negotiated the release of historic data from three locations in Colombia and three in Panama. The

— continued on page 5

Patrick Caldwell
Hawaii Liaison
NOAA/NESDIS/NODC/NCDDC
Department of Oceanography
University of Hawaii at Manoa
Honolulu, Hawaii 96822
E-mail: Patrick.caldwell@noaa.gov

EARTH SYSTEM MONITOR

The *Earth System Monitor* (ISSN 1068-2678) is published quarterly by the NOAA National Oceanographic Data Center. Past issues are available online at *http://www.nodc.noaa.gov/General/NODCPubs/*

Questions, comments, or suggestions for articles, as well as requests for subscriptions and changes of address, should be directed to the editor.

The mailing address for the *Earth System Monitor* is:

National Oceanographic Data Center
NOAA/NESDIS E/OC1
SSMC3, 4th Floor
1315 East-West Highway
Silver Spring, MD 20910-3282

EDITOR
R. Torstenson
Telephone: 301-713-3281 ext.107
Fax: 301-713-3302
E-mail: Roger.Torstenson@noaa.gov

U.S. DEPARTMENT OF COMMERCE
Donald Evans, Secretary

National Oceanic and Atmospheric Administration
Conrad C. Lautenbacher, Jr.,
Under Secretary and Administrator

NOAA/NESDIS SARSAT rescues

In the space of just four days in late September, the Search and Rescue Satellite-Aided Tracking (SARSAT) system led to the rescue of nine individuals in three separate incidents.

In the first incident, on September 23, five rescues were recorded when SARSAT detected a 406 MHz signal 25 nautical miles (NM) north of the island of Saipan. The crew of the fishing boat *Santa Remedios* activated their emergency beacon when their boat began taking on water and sank. The five people onboard abandoned ship into a life raft. The U.S. Navy launched a helicopter from Guam to the scene, which located the life raft and transported the crew safely to the island of Saipan.

In the second incident, on September 26, one rescue was recorded when SARSAT detected a 121.5/243 MHz signal, 115 nautical miles northwest of Anchorage, Alaska. An aircraft had crashed and burned, but the pilot escaped without injury. The U.S. Coast Guard notified a local Alaska Civil Air Patrol (CAP) squadron which launched a plane to the scene. CAP located the crash site and arranged for the pilot to be picked up.

In the third incident, on September 26, three rescues were recorded when the U.S. Coast Guard received a Mayday signal from the sailing boat *Endless Journey* when they began taking on water. The crew was asked to activate their 406 MHz beacon for location purposes. The U.S. Coast Guard launched a helicopter to the scene, which located three crew members on board a life raft 60 nautical miles northeast of Nicaragua. The crew members abandoned their vessel after their pumps could not keep up with the flooding, and the vessel sank. The crew members were safely transported into port.

In November, 20 rescues were recorded in one week in five separate incidents. Three rescues were recorded just east of Fort Bennet, Louisiana. Five rescues took place in one incident in the Bahamas off Freeport. There were seven rescues 65 nautical miles southeast of New York. Three rescues 10 nautical miles north of Port Mansfield, Texas, and two rescues were recorded 211 nautical miles southeast of Nantucket, Massachusetts.

News briefs

NOAA's Coral Reef Information System (CoRIS) site recognized

NOAA's CoRIS website was selected as an exemplary site for educators, by the Eisenhower National Clearinghouse (ENC). The list is published monthly at ENC Online at: *http://enc.org/weblinks/dd/*. ENC is funded by the United States Department of Education and administered by The Ohio State University. ENC collects both physical and virtual resources useful to math and science educators. As part of its mission to serve educators across the nation, each month a dozen sites are chosen to be highlighted. These sites must have current and accurate math and/or science content. They must support school improvement efforts and have useful multimedia features or helpful navigation.

First rescue in contiguous U.S. using Personal Locator Beacon

The first-ever rescue of a person using a Personal Locator Beacon (PLB) in the contiguous United States was recorded on November 14, when SARSAT detected a 406 MHz PLB distress signal in Watertown, New York. Mr. Carl Skalak manually activated his PLB after he became isolated and low on food while camping in extreme weather conditions. The U.S. Air Force notified the New York State Warning Point and the Herkimer County Sheriff. The sheriff's office passed all information to the Watson East Triangle Forest Service. The forest service and Fort Drum coordinated to launch a helicopter and medical team to assist in the search. The helicopter arrived on the scene and received a flashlight signal from Mr. Skalak who was then rescued. This rescue was the first using PLB technology in the contiguous U.S. since they became available for nationwide use on July 1, 2003. NOAA's SARSAT Program was instrumental in getting PLBs available to the public and has worked closely with the AFRCC in the implementation of this important life-saving technology.

Unprecedented solar activity impacts satellite operations

On November 4, the GOES Space Environment Monitor recorded an X28 flare. However, since it occurred on the limb of the sun, most of the high-energy particles were aimed at right angles to the Earth, so the Earth was dealt only a glancing blow. It was the grand finale in a series of flares from two extremely active regions. In just four weeks these two regions produced some 12 X-class flares and 43 M-class flares. This solar activity produced a long period of activity in the near-Earth space environment. NGDC maintains the GOES SEM and SXI archives and a satellite anomaly database for retrospective studies of these events.

Anomalies in satellite and instrument operations during these four weeks continue to be reported. An interesting consequence of the solar flares was they produced so many high-energy particles that these particles contaminated the instrumentation on NASA's Advanced Composition Explorer (ACE) satellite that was designed to measure them. On October 24, the Japanese lost contact with the ADEOS II satellite at 80 km altitude. On October 28, the Japanese data relay satellite, Kodama, went into safe mode. While in transfer orbit, the Small Missions for Advanced Research in Technology satellite of the European Space (ESA) reported various problems in transfer orbit. The solar arrays on commercial geostationary TV satellites showed a sharp degradation in power output from the solar panels to the onboard systems since October 20. Several CCD-based instruments on NASA and NOAA satellites showed "blinding snow" caused by energetic particle contamination in the imagery.

NSIDC glacier analysis featured in *Science News*

An article titled "On Thinning Ice," by S. Perkins, a writer for *Science News*, reports on retreating and thinning glaciers and their impact. The article cites work by NSIDC Director Roger Barry, Dr. Tatiana Khromova (Institute of Geography, Moscow), and Dr. Mark Dyurgerov (Institute of Arctic and Alpine Research, University of Colorado), on glacial shrinkage since 1977 in Central Asia. The work was published August 15 in Geophysical Research Letters.

Summary of hourly sea level records that have been digitized by the JASL

Station	Country	Ocean	Provider	Years	Number of Years
Cartagena	Colombia	Atlantic	IDEAM	1951-1994	44
Tumaco	Colombia	Pacific	IDEAM	1951-1979	21
Buenaventura	Colombia	Pacific	IDEAM	1951-1979	21
Acajutla	El Salvador	Pacific	IGNES	1963-1971, 1986-1991	15
Puerto Armuelles	Panama	Pacific	IGNP	1983-1998	16
Palau	Belau	Pacific	JODC	1929-1936	7
Lungsurannaga	Borneo	Pacific	JODC	1942-1943	2
Balikpapan	Borneo	Pacific	JODC	1942-1943	2
Samarinda	Borneo	Pacific	JODC	1943-1944	2
Pantuan	Borneo	Pacific	JODC	1943-1944	2
Bayor	Borneo	Pacific	JODC	1943-1944	2
Pointe-a-Pitre,FA	France	Atlantic	METEOF	1993-1998	6
Belem	Brazil	Atlantic	NOS	1955-1968	14
Fortaleza	Brazil	Atlantic	NOS	1955-1968	14
Recife	Brazil	Atlantic	NOS	1955-1967	13
Salvador	Brazil	Atlantic	NOS	1955-1964	10
Canavierias	Brazil	Atlantic	NOS	1956-1961	6
Rio de Janeiro	Brazil	Atlantic	NOS	1955-1968	14
Puerto Cortes	Honduras	Atlantic	NOS	1948-1968	21
Puerto Castilla	Honduras	Atlantic	NOS	1955-1968	24
Belize	British Honduras	Atlantic	NOS	1964-1967	4
Port Royal	Jamaica	Atlantic	NOS	1965-1968	4
Guantanamo	Cuba	Atlantic	NOS	1937-1948	12
Fajado, PR	USA	Atlantic	NOS	1921-1923	3
Bermuda	UK	Atlantic	NOS	1933-1949	17
Bridgetown	Barbados	Atlantic	NOS	1968-1970	3
Honolulu, HI	USA	Pacific	NOS	1877-1882, 1899-1901	8
Palmyra	USA	Pacific	NOS	1947-1950	4
La Union	El Salvador	Pacific	NOS	1954-1966	13
Corinto	Nicaragua	Pacific	NOS	1967	2
Cristobal	Panama	Atlantic	PCC	1907-1974	68
Balboa	Panama	Pacific	PCC	1907-1974	68
Point Fortin	Trinidad and Tobago	Atlantic	TTHU	1987-1996	10
Port-O-Spain	Trinidad and Tobago	Atlantic	TTHU	1984-1992	9

Acronyms for providers are as follows:
IGNES Instituto Geografico Nacional de El Salvador
IGNP Instituto Geografico Nacional de Panama
IDEAM Instituto de Hidrologia, Meteorologia y Estudios Ambientales de Colombia
JODC Japan Oceanographic Data Center
METEOF Direction de la Meteorologie, Meteo-France, Antilles-Guyana
NOS NOAA National Ocean Service
PCC Panama Canal Commission
TTIMA Trinidad and Tobago Institute of Marine Affairs

▲ **Figure 2.** The IAGS site in Salvador, Brazil.

Sea level data rescue, *from page 2*

sites on either side of the Panama Canal started recording sea level at the beginning of the 1900s. Prior to the release of the paper data, hourly data were only available at the canal sites from the mid 1970s onward. Thus, these series were extended backwards by over 50 years and now are some of the longest records held at the JASL. While monthly data had previously been available through the PSMSL, a more thorough quality control was made possible by acquiring the hourly data. In addition, the availability of daily data augments studies of high frequency signals such as the passage of Kelvin waves associated with the El Nino-Southern Oscillation phenomena.

The Japan Oceanographic Data Center (JODC) has contributed several sets of hourly tabulations to the JASL in support of the GLOSS Archeology

Project. Data for Palau, Belau in the western tropical Pacific from 1929-1937 were discovered in a library at the Kyoto University. The data are of high quality with few gaps. JODC also made available historic hourly tabulations from five locations in Borneo from 1942-44.

One of the largest sources for historic tabulations in paper or microfiche form is the NOAA National Ocean Service (NOS). During the mid 1970s, a large batch of over 1000 years of hourly water level heights for 55 sites were digitized under the guidance of Dr. Doug Luther, who was in graduate school at the Massachusetts Institute for Technology and used the data for various studies (Luther, 1980; Luther, 1982; Miller et al., 1993). These data were ingested into the JASL. NOS became active in data rescue during the 1990s and digitized over 500 years of data for over 80 locations, mostly in the continental USA. For sites that overlapped with the Luther set, the JASL was updated with finalized NOS data, which have a higher degree of calibration and quality control.

The NOS historic holdings extend beyond the USA. During the mid 1940s through early 1970s, the US Coast and Geodetic Survey (USCGS), which was absorbed into NOS during the 1970s, maintained a network of tide gauges throughout the tropical Pacific, the Caribbean, and South and Central America. Sites from the latter region were referred to as the Inter American Geodetic Survey (IAGS). The stations were equipped with standard tide gauges using analog or punched paper tape recording mechanisms and were well maintained (Figure 2). Significant amounts of IAGS hourly tabulations remained in paper form when the author visited NOS in 1997 to gather an inventory for data recovery.

Data rescue has been motivated by several regional projects such as the Caribbean Planning for Adaption to Climate Change (CPACC) beginning in 1997 and the Water Level Observation Network for Latin America (RONMAC in Spanish) starting in 1999. The NOS and JASL formed a partnership to bring the historic USCGS data for these regions into digital form. Mr. Doug Martin of NOS arranged for the photocopying of the original hourly tabulations and comparative readings of nine sites totaling 69 years, which were subsequently digitized by the JASL. Additional data for these projects were received from agencies overseas. Mrs. Shelley-Ann Jules-Moore, data manager for CPACC based at the of the Institute of Marine Affairs, Trinadad and Tobago, provided 20 years of data from two sites on Trinidad. In support of RONMAC, Mr. Luis Hernandez of the National Geographical Institute of El Salvador contributed hourly tabulations for Acajutla, a GLOSS site. This allowed the time series to be extended backwards by a decade as well as replacing a four-year gap in the late 1980s.

At the GLOSS GE-7 meeting in Honolulu in April 2001, it was noted that Brazil was a region with low data availability. Since the NOS inventory included Brazilian sites, a request was sent to Mr. Leonard Hickman, who made arrangements for the transfer of this historic set to the JASL for digitization. The set included six stations within the IAGS network for the years of roughly 1955-1968. These data are of good quality and have significantly increased the holdings of science-ready hourly data for Brazil. Additional IAGS Brazilian data held at the US National Archives are expected to be released. Technical metadata for the IAGS sites in Brazil are held by the Brazilian Insti-

— continued on page 6

Sea level data rescue, *from page 5*

tute of Geography and Statistics (IBGE in Portuguese). Dr. Roberto Teixeira Luz of the IBGE provided a complete set for Salvador, including the photograph in Figure 2, and will release copies for the other stations.

The NOS is presently involved in data rescue as part of the Climate Database Modernization Program (CDMP), sponsored by the National Climatic Data Center, to digitize microfiche records, which include the station maintenance reports, leveling, comparative readings, summaries, and original hourly height tabulation. In the first year of this project about 253,000 images were digitized and indexed. This year NOS expects to complete an additional 370,000 images. The images are managed by the Web Store Search Retrieve Display (WSSRD), which will be made available to the public upon project completion. NOS also aims to image all of the station packages that contain the history of maintenance and leveling at about 6,000 locations. These images will also be indexed for access through WSSRD.

NOS is also rescuing ancillary data from tide stations and surrounding waters. Daily measurements and monthly means and extremes of water density and air and water temperature, totaling 4,400 station years, are targeted for imaging and indexing. Moreover, approximately 158,000 images of tidal current data and analysis from NOS historic circulation surveys dating from the 1840s to the 1980s have been digitized under the CDMP.

Data rescue expands the temporal and spatial coverage of national and international data center holdings, which can be accessed for a variety of studies. Sea level data are particularly important for understanding decadal variations and climate change, yet must be available from numerous worldwide locations with long series in order to compute sound statistics. NOAA supports the recovery of global historic sea level data through the efforts of the JASL and NOS, which in turn contribute to international scientific activities such as CLIVAR, GLOSS, and GODAR.

References

Caldwell, P., 1992. Building an archive of tropical sea level data. Earth System Monitor, December 1992. pp 2-4.

IOC, 2003: Proceedings of the "International Global Oceanographic Data Archeology and Rescue (GODAR) Review Conference." IOC Workshop Report XXX Silver Spring, MD, USA, July 1999.

Luther, D.S., 1980: Observations of long period waves in the tropical oceans and atmosphere (Ph.D. Thesis). Woods Hole Oceanographic Institution Technical Report #80-17, 210 pp.Luther, D.S., 1982: Evidence of a 4-6 day barotropic, planetary oscillation of the Pacific Ocean. J. Phys. Oceanogr., 12, 644-657.Miller, A.J., D.S. Luther and M.C. Hendershott, 1993: The fortnightly and monthly tides: Resonant Rossby waves or nearly equilibrium gravity waves? J. Phys. Oceanogr., 23, 879-897. ■

Call for Papers

Third GOES-R Users' Conference

Define the GOES-R Future Today!

Boulder, Colorado, May 10-13, 2004

The third generation of the National Oceanic and Atmospheric Administration (NOAA) Geostationary Operational Environmental Satellites (GOES) will begin with GOES-R (in 2012). An advanced imager, hyperspectral sounder, coastal water sensor, lightning mapper, solar imager and space environment monitor will provide critical atmospheric, oceanic, climatic, solar and space data.

The focus of the Third GOES-R Users' Conference will be to help users prepare to acquire, digest and use GOES-R data and to assist with the advanced planning of requirements, products, communications, and distribution of data. The conference will consist of speaker presentations, poster sessions, and facilitated breakout sessions.

Abstracts are invited for poster papers focusing on GOES-R potential applications, GOES-R as part of the Global Observing System and the smooth transition to GOES-R.

The deadline for abstracts is January 28, 2004.

Please see the web site for details: *http://www.osd.noaa.gov/announcement/index.htm.*

Pat Viets
Short & Associates, Inc.

Introducing the U.S. NODC Archive Management System

Stewardship of the nation's oceanographic data archive

Donald W. Collins, Steven B. Rutz, H. Lee Dantzler, Eric J. Ogata, Francis J. Mitchell, Joseph Shirley, and Thaila Thailambal
National Oceanographic Data Center
NOAA/NESDIS

The stewardship of the nation's oceanographic data archive is an essential responsibility of the U.S. National Oceanographic Data Center (NODC). At issue are the long-term preservation, integrity, and accessibility of irreplaceable observational data through multiple technological and scientific generations. Recognizing these challenges, the NODC has implemented new processes to ensure that its data archive stewardship responsibilities are met, that online data discovery and retrieval services are expanded, and that adequate supporting metadata are available to guide use of the provided data. The new NODC Archive Management System (AMS) enables datasets to be accessioned, archived, and disseminated in a Web-enabled, browser-based environment (*http://www.nodc.noaa.gov/Archive/Search/*). To date, over 20,000 unique accessioned datasets, ranging from individual observations to large collections by major programs, are included in the system. This is the first time that such a large amount of originator's oceanographic data has been brought into a single data management system.

Using the NODC AMS, the ocean community will be able to search historic observation data in a new way, with links to related cruise reports and other gray literature. In addition, ocean scientists can more easily fulfill contractual obligations for submitting federally funded observation data and can now verify the existence of previous data submissions.

The NODC AMS consists of four primary components:

1) Accession Tracking Database (ATDB) - generates a unique accession (i.e., a tracking) number for each data submission, captures basic metadata and internal data management information, allows metadata to be exported into XML files that follow the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM) uses a controlled vocabulary for dataset descriptions, and provides a mechanism for NODC data managers to oversee the management of each dataset;

2) Archive File Management System - generates a uniform directory tree structure for each unique, original dataset submission; creates message digests for file validation; performs virus checks; implements dataset versioning; and provides for automated backups;

3) NOAA Metadata Repository (NMR) - stores and manages FGDC CSDGM-compliant metadata for each accessioned dataset, and allows the import and export of metadata using XML; and

4) NODC Ocean Archive System - allows for the search and dissemination of datasets archived at NODC via a web browser.

One of the initial considerations of the AMS development effort was to determine the viability of open source software and utilities to develop a robust system for managing a mission-critical information system. To test this capability, the NODC decided to use PostgreSQL database and Linux servers for the ATDB, File Management System, and Ocean Archive System components. User interfaces, file management tasks, and database search and retrieval tasks for these components were developed using Perl CGI scripts and standard SQL commands. Interactive user interfaces are all browser-based and function in established browsers (e.g., Mozilla, Netscape, and Internet Explorer). Internal users and the database administrator may use SQL commands to perform advanced searches, to create management reports on the information in the ATDB, and to perform necessary database maintenance tasks.

File Management System functions provide a significant improvement in NODC's ability to assure that data files are migrated through time and technological changes. The computation of MD5 hashes for each file in each directory prior to and following the creation of a backup copy on archival media provides a high level of confidence that any changes to a data file are detected. NODC is also able to create and manage incremental changes to the data contents of an accessioned data set by providing a mechanism to make multiple versions of the data set. Each version is encapsulated separately from any other version of the same data set and all versions are maintained in the archives. Accessioned data sets are also backed up to a local system and a regional backup facility (less than 60 miles away). A remote backup facility (more than 60 miles away) is being negotiated.

The Ocean Archive System is the browser-based tool for users to search and discover data sets archived at NODC. Information elements from the ATDB, such as observation date range, sea area, project, and data-collecting institution are used to create a query. A

— continued on page 8

Donald W. Collins
Database Management Division
National Oceanographic Data Center
1315 East West Hwy, E/OC1
SSMC3, 4th Floor
Silver Spring, Maryland 20910
E-mail: Donald.collins@noaa.gov

***NODC archive system**, from page 7*
controlled vocabulary, accessed by drop-down lists, is used for most ATDB elements, ensuring more precise data entry and higher-precision results. Because the Ocean Archive System only searches for original data sets in the NODC archives, the NODC provides a number of other search utilities to specific products, such as the NODC World Ocean Database (via WODSelect). The current functionality of the Ocean Archive System allows a user to obtain the entire directory for an accession. When fully implemented, the Ocean Archive System will allow a user to drill down in the archive file structure to retrieve a specific file within an accession.

In contrast to the other AMS components, the NOAA Metadata Repository (NMR) component is comprised of an Oracle database developed by Blue Angel Technologies, Inc. The database model is based on the FGDC CSDGM. The browser-based NMR interface was developed by a team from NODC and the National Geophysical Data Center (NGDC) using Java Servlets. The NOAA National Data Centers (NODC, NGDC, and the National Climatic Data Center) share the capabilities and management responsibilities of this component. The NMR accepts structured descriptive metadata in XML format that is consistent with the CSDGM. Descriptive metadata content can be edited, parsed, approved and published using NMR tools. Descriptive metadata from the NMR for each accessioned data set is used to populate the NOAA Server metadata clearinghouse. The NOAA Server clearinghouse coordinates and exchanges descriptive metadata with other clearinghouses, such as the NASA Global Change Master Directory (GCMD) and the FGDC Geospatial Clearinghouse.

A significant effort was required to migrate historic accessioned data from legacy file management systems to the current AMS. It was also necessary to move metadata from the legacy database to the ATDB and NMR. Thus far, more than 16,000 NODC accessions have been moved from multiple servers and directories into the AMS File Management System. In addition, a series of programs were developed to move metadata for more than 19~~,,000~~ 600 accessions from a legacy Sybase database into the ATDB and NMR. Work continues on developing the linkage between the contents of the AMS databases.

The NODC AMS represents a significant step forward as the nation moves towards implementing an integrated ocean observing system, with the concomitant increases in observations data flow. The test of using open source software to develop a mission-critical database appears to be successful. NODC personnel have been using operational versions of the ATDB and File Management System internally for approximately one year with few problems. Future work includes improving metadata content for all datasets, allowing remote accessioning of data sets, and incorporating GCMD valid keywords into the system. The AMS enables NODC to be compliant with the data archiving and access requirements described in the U.S. Integrated Ocean Observing System (IOOS) Plan for Data Management and Communications (DMAC). ■

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
Publication Distribution Facility
1315 East-West Highway
Silver Spring, MD 20910-3282
ATTN: Earth System Monitor

Address Correction Requested
OFFICIAL BUSINESS
Penalty for Private Use $300