Executive Summary

Observations are intrinsic to the mission of the National Oceanic and Atmospheric Administration (NOAA) and NOAA has a vision to see a truly integrated global observing system that will bring together all aspects of environmental monitoring on common platforms, to ensure data quality, to manage those data efficiently for the long-term, and to make these data easily and readily accessible as needed. This report summarizes plans and accomplishments in integrating data types, across disciplines, across line offices and across observational platforms, to provide capability to serve growing user demand.

NOAA has a core mission to understand and predict changes in the Earth's environment, and conserve and manage coastal and marine resources to meet the Nation's economic, social, and environmental needs. To accomplish this mission, NOAA must first monitor and observe the land, sea, atmosphere, and space through a data collection network. The data and information from this network must then be organized, permanently preserved, then provided to scientists and researchers for study and tracking of the Earth's constantly changing environment.

NOAA Environmental Data: Foundation for Earth Observations and Data Management System describes the status at the end of 2002, and future challenges for the management of an ever-expanding environmental database pouring into NOAA from a world-wide data collection network and the progress towards bringing it all together to address the complex issues of today's world. This biennial Report to Congress assesses the adequacy of the environmental data and information systems of NOAA. It updates *The Nation's Environmental Data: Treasures at Risk* provided to Congress, September 2001.[1]

NOAA oversees a spectrum of environmental data management services for diverse purposes: Ecosystem Management; Commerce and Transportation; Hazards; Climate Monitoring; and Ocean Data management. A section of this report is devoted to data management for each of these purposes. There is also a section, called cross-cutting themes, that covers all NOAA-managed environmental data: User Needs; Technology Changes; Increasing Complexities and Volumes of Data; Seamless Uses of Geospatial Data; Scientific Stewardship of Data; and Homeland Security**.** NOAA is stepping up to the challenges of providing data management and service today. The increasing demands were included in a study by the National Academy of Sciences.[2]

In ecosystem management, data from multiple and diverse research programs impose tremendous data management and access challenges to NOAA in terms of data rescue, formatting, cataloging, geographic information system mapping, and standardized online remote access. However, NOAA stands on the verge of a virtual revolution in its ecosystem data management and access capabilities.

NOAA supports commerce and transportation infrastructure by providing critical information, products, and services essential to the safe and efficient transport of goods and people at sea, in the air, and on land. This infrastructure is key to the well being of the U.S. economy. Real-time, 24x7 quality-controlled information is required, especially high resolution data, e.g., Next Generation Weather Radar (NEXRAD) (which will increase data volume by a factor of almost 30 per year in the next eight years) and side scan sonar and bottom backscatter data in support of nautical charting efforts.

Monitoring and forecasting hazards are key functions of NOAA which also depend on high-resolution, high-volume data management. NOAA is the sole United States official voice for issuing warnings during life-threatening weather situations and also provides watches and warnings for non-weather, natural and man-made hazards including severe weather, hurricanes, space weather, seismological events, wildland fires, floods, toxin release, and environmental contamination. Creative uses of new technology, combined with NOAA satellite data, on-site observations, and historical records, will continue to lead to breakthroughs in hazards detection and mitigation.

Climate monitoring, evaluation, and forecasting are critical to economic sustainability and environmental stewardship, as well as planning and responding to the quality of life changes that society will encounter in the 21st century and beyond. The challenge facing the NOAA National Data Centers is not only ingesting and processing the data, but also the convenient and timely access to the data and information. The Climate Database Modernization Program addresses access and utilization issues with the goal to make paper/film non-digital historical climate data digitally accessible online via the Internet. *World Ocean Database 2001* is a global, comprehensive, integrated, scientifically quality-controlled database, completely documented, available online and on CD-ROM, which includes profiles for over 7 million hydrographic stations.

User requirements are critical to assure that NOAA's environmental data and information can contribute significantly to the Nation's economic well being. Examples are given. NOAA has made a number of its most requested data sets available to users through the Internet to improve customer service and to reduce the costs for servicing user requests.

NOAA has been an early adopter of emerging advanced computing and communications technologies to improve data management and data availability on the Internet and next generation Internet/Internet2. A plan for improving management of NOAA's information technology resources was developed in October 2001. Over the next 15 years, current and planned remote sensing observing systems will produce volumes of environmental data on a scale not seen before. By 2017, plans for NEXRAD, Geostationary Operational Environmental Satellites, Polar-orbiting Operational Environmental Satellites, National Polar-orbiting Operational Environmental Satellite System (NPOESS), NPOESS Preparatory Project and numerous *in-situ* observation programs will increase the total data volume to more than 140 petabytes (140,000,000,000,000,000 bytes). NOAA's Comprehensive Large Array-data Stewardship System (CLASS) Project is designed to enhance NOAA's capability to predict and assess decadal to centennial climatological changes.

Geospatial data covers information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the Earth. An enterprise-wide geographic information system, the comprehensive NOAA Observing System Architecture system, was developed to enable NOAA to document its multiple observing systems and identify ways in which to evolve them in an integrated management approach. NOAA is a full federal partner in the development of framework data standards associated with the GeoSpatial One-Stop effort as well as the inventorying, documentation, data collection, data discovery, and access to these framework layers.

NOAA recognizes the need for scientific stewardship of data. Effective utilization of climate data and information and long-term stewardship requires the ability to scientifically manage the Nation's climate records and provide relevant utility to a wide range of customers. Scientific stewardship is characterized as maintaining scientific integrity and long-term utility of climate records through monitoring, improving quality, and the extraction of select key parameters from new observations and the historical records.

NOAA faces the challenge of doing business in a world of heightened security concerns and terrorist threats following the horrific events of September 11, 2001. Many of NOAA's data sets are used to support forecasting of and response to natural hazards, as well as their use in climate studies, marine applications, and a variety of industries (e.g., construction), and are essential in support of homeland security activities that include emergency response, monitoring, predicting, and modeling in times of national emergencies. Data once openly available may now be considered sensitive; posing challenges for how NOAA will maintain its "full and open" policy for data and information in a new era of heightened sensitivities. The problems facing NOAA to ensure its data are available when needed are numerous. Certain data streams must be delivered in real-time. Systems need to incorporate knowledge-based expert software. Continued enhancements will be required as new sensor data are added, and back-up delivery and archive sites must be maintained for as needed delivery of critical non-real-time data.

NOAA Administrator Vice Admiral Conrad C. Lautenbacher has a broad vision for global observations and has indicated support for an integrated global observing system, recognizing that NOAA cannot do it alone. Toward this end, NOAA is taking a prominent role, partnering with the National Aeronautics and Space Administration (NASA) and other U.S. agencies to bring this global perspective into achievable reality. The initial effort began with an Earth Observations Summit on July 31, 2003. The need for this was widely recognized and provides the benefit of a sound plan for end-to-end stewardship of environmental data. NOAA is on target to step up to this opportunity.

Synopsis: Vision for NOAA

Nobel laureate, Richard Feynman, stated "Observation is the judge."

The National Oceanic and Atmospheric Administration (NOAA) has a vision to see a truly integrated global observing system in place that attempts to integrate all aspects of environmental monitoring on common platforms, to ensure data quality, to manage those data efficiently for the long-term, and to make those data easily and readily accessible as needed. Vice Admiral Conrad C. Lautenbacher, Jr., U.S. Navy (Ret.), Undersecretary of Commerce for Oceans and Atmosphere and NOAA Administrator, indicated his belief that NOAA is the right agency to take a leadership role within the United States. This document reports on the status of all efforts across NOAA. It also reports on a significant accomplishment in integrating data types, across disciplines, across line offices, and across observational platforms. The integration provides identification of data for user-specified geographic locations. This capability is described in the section on "Seamless Uses of Geospatial Data" and gives the status on NOAA's progress in building a comprehensive observing architecture.

NOAA Environmental Data: Foundation for Earth Observations and Data Management System describes the current state of, and future challenges for, the management of an ever-expanding environmental database pouring into NOAA from a world-wide data collection network and the progress towards bringing it all together to address the complex issues of today's world.

This biennial Report required by Congress assesses the adequacy of the environmental data and information systems of NOAA. The Report serves as an update to *The Nation's Environmental Data: Treasures at Risk*[1] provided to Congress in September 2001.

The strength of NOAA in leading the effort to enhance the utilization and usefulness of data is its unique position among government agencies. NOAA has a core mission to understand and predict changes in the Earth's environment, and conserve and manage coastal and marine resources to meet the Nation's economic, social, and environmental needs. To accomplish this mission, NOAA must first monitor and observe the land, sea, atmosphere, and space through a data collection network. The data and information from this network must then be organized, permanently preserved, and then provided to scientists and researchers for study and tracking of the Earth's constantly changing environment. A study by the National Academy of Sciences also highlights the demands.[2]

In this first decade of the 21st century, NOAA oversees a spectrum of environmental data management services that includes:

- Ecosystem Management
- Commerce and Transportation
- Hazards
- Climate Monitoring
- Ocean Data Management

In plans for the second decade of the century, NOAA foresees cross-cutting challenges that apply to all NOAA-managed environmental data. These include:

- User Needs
- Technology Changes
- Increasing Complexities and Volumes of Data
- Seamless Uses of Geospatial Data
- Scientific Stewardship of Data
- Homeland Security

This Report addresses data management services and these challenges throughout NOAA. Success at meeting the ever-increasing data management requirements will determine NOAA's success in meeting its primary mission goals. In addition to monitoring and observing the environment, these services and challenges include:

- Understanding and describing how natural systems work together through investigation and interpretation;
- Assessing and predicting the changes of natural systems, and providing information about the future;
- Engaging, advising, and informing individuals, partners, communities, and industries as required, and facilitating in the exploitation and application of appropriate data;
- Managing coastal and ocean resources to optimize benefits to the environment, the economy, and public safety.

Ecosystem Management

Since its creation in the early 1970s, NOAA has played a leading role in coastal and ocean ecosystem management, primarily through its implementation of the Coastal Zone Management Act (CZMA), the Endangered Species Act, and the Marine Mammal Protection Act. More recently, through revision of the Magnuson-Stevens Fishery Conservation and Management Act (MSFCMA), NOAA ecosystem management has extended explicitly to marine fishery management. Under a variety of NOAA and NOAA-sponsored programs, including federal and state partnerships for coastal zone management, National Estuarine Research Reserves, National Marine Sanctuaries, Marine Protected Areas, marine fishery management, endangered species protection, ocean observing and research, and remote sensing, NOAA collects and maintains vast quantities of marine and coastal ecosystem data. These data and relevant data sets compiled by both NOAA-funded and non-NOAA-funded research programs impose tremendous data management and access challenges to NOAA in terms of data rescue, formatting, cataloging, geographic information system (GIS) mapping, and standardized online remote access. NOAA is addressing these challenges through innovative and cross-cutting initiatives among its National Marine Fisheries Service (NMFS), National Ocean Service (NOS), NOAA National Data Centers (NNDC) [consisting of the National Climatic Data Center (NCDC), National Geophysical Data Center (NGDC), and National Oceanographic Data Center (NODC)], National Coastal Data Development Center (NCDDC), and among numerous coastal states. Through these efforts and tremendous advances in computer technology, NOAA stands on the verge of a virtual revolution in its ecosystem data management and access capabilities.

Commerce and Transportation

NOAA provides critical information, products, and services essential to the safe and efficient transport of goods and people at sea, in the air, and on land. To supply these products and services, NOAA acquires a


Water Level Station at Gray Gables, Massachusetts. The station, part of NOS's National Water Level Observation program, includes a data collection platform, acoustic sensor with a protective tube, satellite antenna, various meteorological sensors and solar panel to keep the battery charged.

wide array of data, ranging from periodic to continuous, from the ocean to the land and the atmosphere, and using both fixed and mobile platforms. NOAA provides real-time observations and analyses of temperature, wind, pressure, clouds, turbulence, and icing to support the Nation's maritime community and general aviation, package, and commercial carriers. Accurate and timely navigational data and information are acquired by NOAA and distributed to mariners navigating our waters and to key players active in port and harbor development, from the ice-covered waters in Alaska to the small estuaries along the southeast coast. NOAA's atmospheric and radar observing systems form the backbone of an expanding public/government partnership for observing the Nation's airports and air corridors that make up the National Aerospace System.

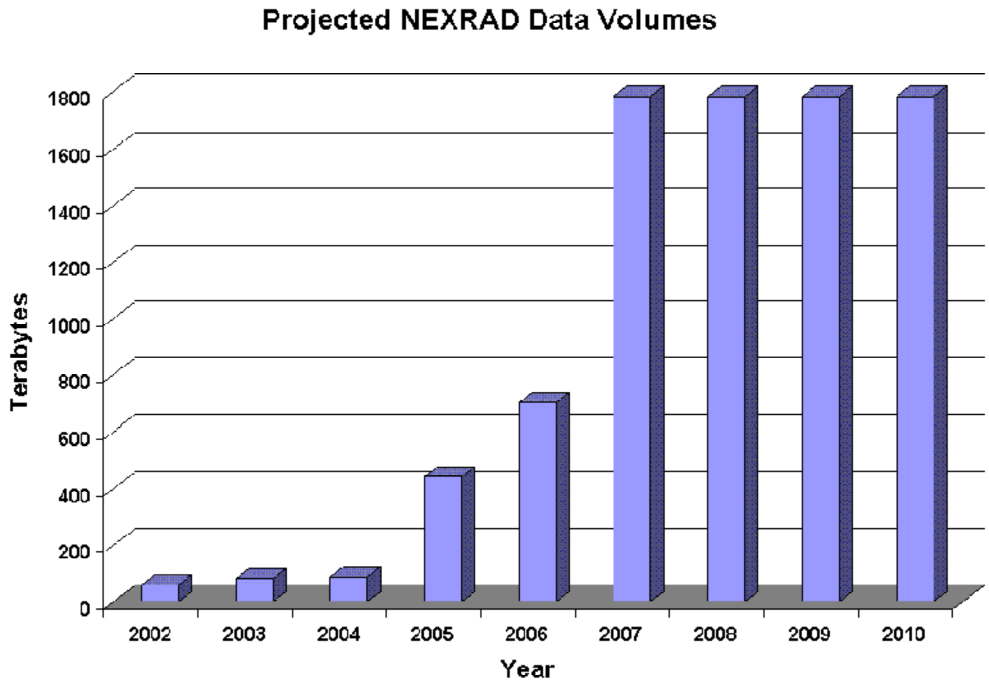
Today, two-thirds of everything purchased by American consumers arrives by ship, and this trade is conservatively expected to double by 2020. In addition, the U.S. coastal recreation and tourism industry, with over 17 million recreational boats, has an annual economic value of about $24 billion to the country. On land, the infrastructure for transportation contains 3.9 million miles of roads, 120,000 miles of railroad tracks, 200,000 miles of gas and oil pipelines, 5,400 airports, and 6,600 miles of urban transit rail.[3,4,5] This infrastructure is key to the well-being of the U.S. economy.

Appropriate technology is needed to maximize both the quality and the efficiency with which NOAA manages these data to ensure they are accurate, reliable, secure, and readily accessible to meet the stringent demands of the transportation sector. Real-time, 24 hours per day, 7 days per week (24x7) quality-controlled information dissemination is rapidly becoming the norm.

With the many recent advances in technology, significantly larger data sets are being acquired using remote sensing techniques. The additional information resulting from the upcoming improved capabilities of Next Generation Weather Radar (NEXRAD) will increase the amount of (uncompressed) data by a factor of almost 30, from 60 terabytes (TB) per year in 2002 to almost 1,800 TB [or 1.8 petabytes (PB)] per year by 2010. Side scan sonar and bottom backscatter data sets acquired in support of nautical charting efforts can result in up to 600 gigabytes (GB) of data per day per vessel. These increases will significantly impact NOAA's data storage, data handling, and communication infrastructure.

Hazards

Each day in 2003, NOAA ingests approximately 254 GB of new satellite data, compared with 65 GB in 1999. New NOAA and non-NOAA satellite programs will increase data volumes at a staggering rate. By 2017, it is expected that this volume will increase to over 30 TB per day.



Most of these data are used for hazard, weather, and climate monitoring and forecasting applications. Data are also collected from individual locations on the ground, in the ocean, or in the air. On-site observations contribute to the accuracy of satellite data, and add critical information to existing data compilations. NOAA's historical data collections can be applied to both real-time and long-term aspects of hazards monitoring, such as climate change.

Accurate climate monitoring and forecasting depend on high-resolution, high-volume data management. The exponentially increasing volume of satellite data is the largest challenge associated with NOAA's hazards mandate.

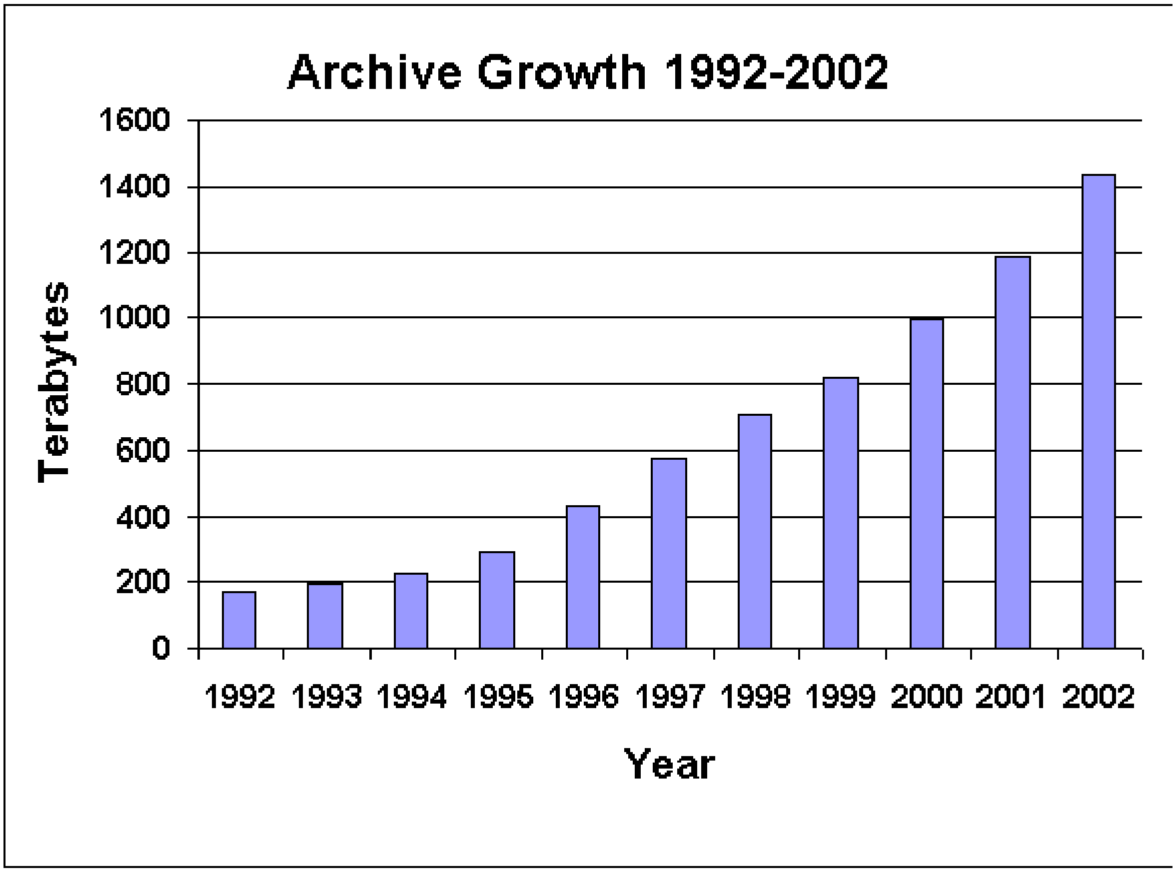
NOAA is the sole United States official "voice" for issuing warnings during life-threatening weather situations. NOAA also provides watches and warnings for non-weather, natural, and man-made hazards. Hazards may include severe weather, hurricanes, space weather, seismological events, wildland fires, floods, toxin release, and environmental contamination.

During the occurrence of an extreme hazards event, NOAA data and information assist first responders in emergency management. In the recovery period after an event, NOAA researchers study the scientific processes that occurred, leading to improved forecasting.

Creative uses of new technology, combined with NOAA satellite data, on-site observations, and historical records, will continue to lead to breakthroughs in hazards detection and mitigation. These scientific advances will save lives, property, and money.

Climate Monitoring

Climate monitoring, evaluation, and forecasting are critical to economic sustainability and environmental stewardship, as well as planning and responding to the quality-of-life changes that society will encounter in the 21st century and beyond.



At the heart of today's modern climate monitoring is the need to ingest, process, access, archive, and analyze enormous numbers of observations. In 1996, 400 TB made up the total digital archive volume at NOAA's NCDC. By 1999, volume had doubled. Currently, the total digital archive of the three NNDCs is about 1.4 PB. Over the next ten years, new satellites and improved weather radars will dramatically increase by several orders of magnitude the volume of data and information. Projections through 2010 indicate an increase by a factor of 25, an estimated volume of over 43 PB per year.

The challenge facing the NNDCs is not only the ingest and processing of the data, but also providing convenient and timely access to the data and information. Millions of paper pages and thousands of feet of microfilm/microfiche of recorded instrument measurements and other information dating back hundreds of years currently are under the stewardship of NOAA's NCDC. Much of this historical information has been or is being converted to digital form.

Digital databases, such as wind speed and direction, precipitation, temperature, and pressure are far more useful than paper and microfilm records. These databases support many disciplines, including economic research, engineering, risk management, and passive (i.e., solar, wind) energy enterprises. The Climate Database Modernization Program (CDMP) addresses access and utilization issues. The Program's goal is to make historical paper and film non-digital climate data digitally accessible online via the Internet. Access to historical climate information will require a combination of optically scanned digital images of records and digitized (manually keyed) observations. Thirty million documents have been imaged and many thousands of observations manually keyed or digitized from shipping records, America's military forts, major U.S. cities, lighthouses, weather ships, and other sources.

The deployment of a new generation of satellites over the next ten years, including: the National Aeronautics and Space Administration's (NASA's) Earth Observing System (EOS); Next Generation Geostationary Operational Environmental Satellite (GOES); joint Department of Defense, NASA, and NOAA National Polar-orbiting Operational Environmental Satellite System (NPOESS); and the dipolar and phased array enhancements to the operational NEXRAD present major data management challenges (i.e., stewardship and customer access) for the NNDCs.

Climate Data Modernization Program (CDMP) goes International.
The Climate Data Modernization Program (CDMP) hosted a workshop, *Digitizing of Analog Charts*, in Asheville, North Carolina July 9-10, 2002. The goal of the workshop was to identify viable low-cost methods of digitizing the large numbers of hourly precipitation and other analog data under the stewardship meteorological communities in many countries. During the months of May and December 2002, NOAA staff visited Vietnam and countries in Africa (Kenya, Malawi, Mozambique, Niger, Senegal, and Zambia) to teach local meteorological services how to image historical weather records using digital cameras.

Today, data and information are not easily accessible or easily integrated to provide a comprehensive and multi-disciplined picture of climate trends and change. In the future, scientists and decision makers will need to rapidly access and utilize climate data and records dating back decades and centuries in order to create a total, inclusive portrait of the Earth's dynamic climate system.

Ocean Data Management

The oceans are the great regulator of climate on the decadal to centennial time scales. Ocean observations are not as extensive as land observations, nor as well structured and organized as weather observing systems. Therefore, they are more difficult to collect and archive. Ocean data, in many formats, are collected during numerous, usually unrelated, expeditions conducted by many countries and different oceanographic institutions. NOAA's NODC leads several ocean database-building activities. The result of this work has been a substantial increase in the amount of historical oceanographic data available to the ocean and climate research communities. Two projects have added more than three million digitized ocean temperature profiles to the databases, which are being accessed by the international research community studying the role of the ocean in the Earth's climate system.

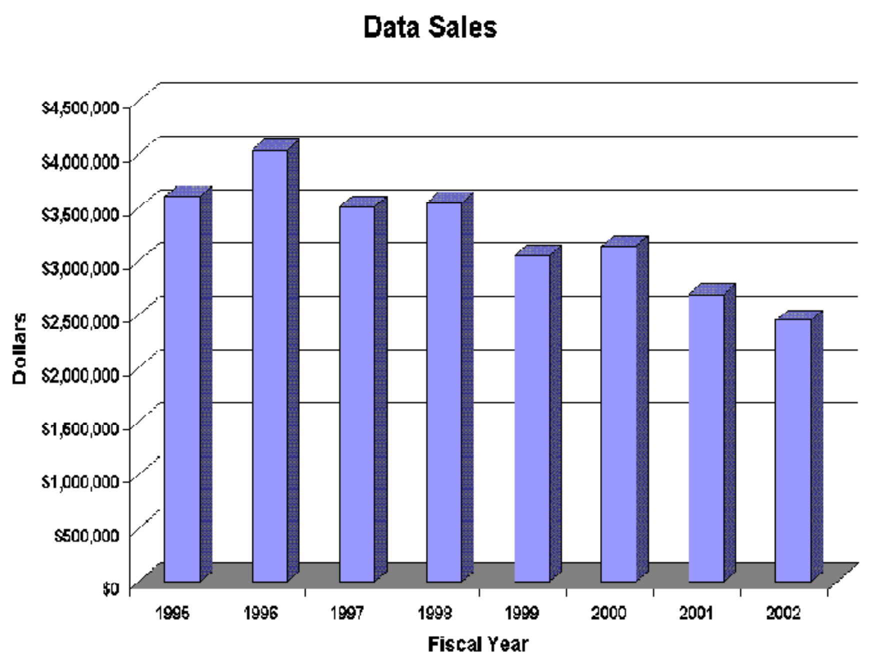
The Global Oceanographic Data Archaeology and Rescue Project (GODAR), sponsored by the Intergovernmental Oceanographic Commission (IOC), locates and digitizes historical oceanographic data, both coastal and open ocean. Much data exist in manuscript form and/or data in electronic form that are not part of an internationally available database. Under the project, these data are integrated into digitally accessible world ocean databases. NOAA's NODC leads the *World Ocean Database Project*. *World Ocean Database 2001* (WODO1) is a global, comprehensive, integrated, scientifically quality-controlled database that is completely documented and available online and on computer disc - read only memory (CD-ROM). The WOD01 includes profiles for over 7 million hydrographic stations that feature observations such as temperature, salinity, inorganic nutrients (nitrate, nitrite, phosphate, silicate), oxygen, carbon alkalinity, dissolved inorganic carbon, pH, chlorophyll, and biology (zooplanknton and phytoplankton).

User Needs

Environmental scientists and advisors have a critical need for a long environmental record of historical and recent global data. These data are used to assess long-term trends, evaluate current status, and predict future conditions and events. This information is necessary to help answer questions from the public, industry, and policy makers such as the following:

- Is global climate change really occurring?
- Are the oceans getting warmer?
- What is the status of the Nation’s ecosystems?
- How do El Niño and La Niña affect the weather?
- How long will a drought last?
- How severe will winter be this year?

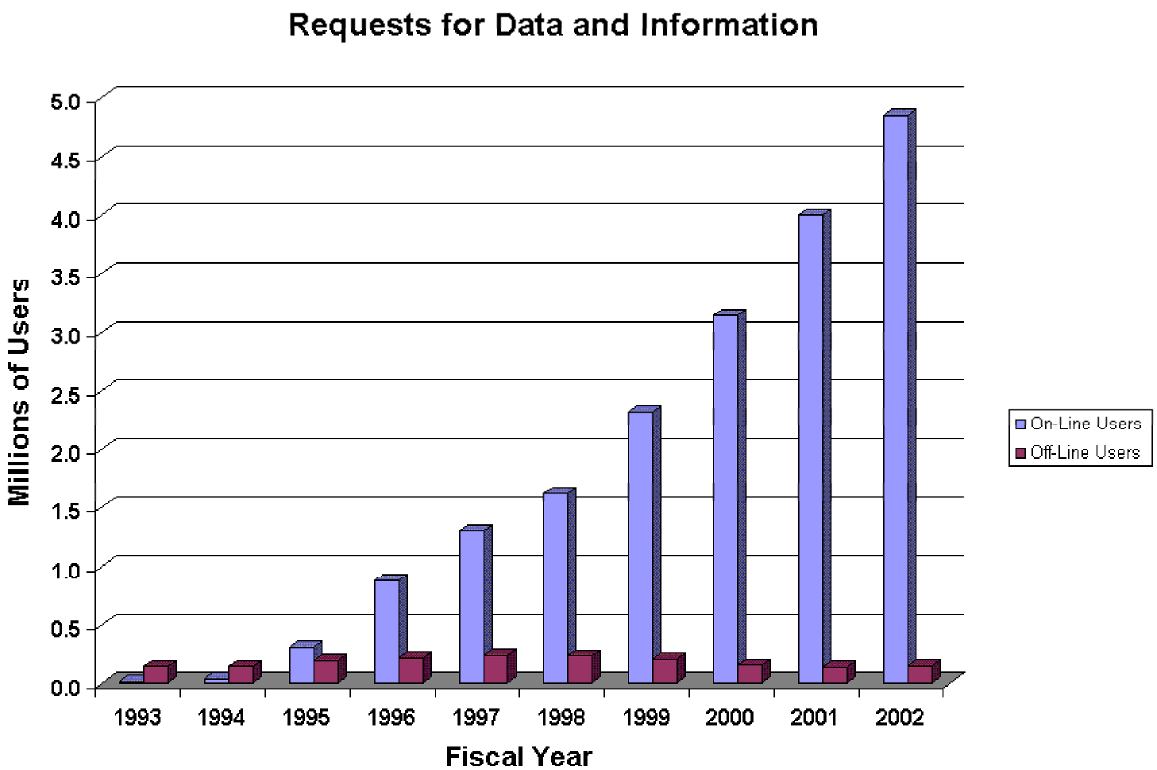
To fully support its users, NOAA must listen to their needs, partner with the broader research and operational communities, and capitalize on new science and technology that accelerate service improvements customers expect. NOAA must do so in a responsive and cost-effective manner.



The graph depicts total data sales revenues at the three NOAA data centers (NCDC, NGDC, and NODC). Revenues (cost recovery) have exhibited a gradual decline since the mid 1990s. This decline results from a NOAA commitment to e-government business practices. Data access costs via the web are either much lower or free to constituents. This has resulted in lower data sales revenues even as the customer’s demand for environmental data has increased.

It is clear that use of NOAA’s environmental data and information can contribute significantly to the Nation’s economic well being. Examples of areas where use of these data may contribute include disaster preparedness, agriculture, and construction. It is estimated that better preparation, response, and mitigation could reduce the average cost of storm-related disasters by 10 percent, or $700 million savings per year.[6] Economic benefits to U.S. agriculture from improved El Niño forecasts by altering planting decisions could be $300 million annually.[7] In construction, Air-Freezing Index research has allowed the construction industry to build shallower frost-protected foundations that save $330 million annually.[8]

Over the past 20 years, there has been a tremendous increase in the number of business users who request data and information from NOAA. The initiation of an online store service for ordering data products and the acceptance of credit cards as a payment option have contributed to the changing user profile by



simplifying the process of obtaining data and information from NOAA. A small sample of the business uses of NOAA environmental data include attorneys and consulting meteorologists, the insurance industry, utilities, agribusiness, marine transportation, the fishing industry, and financial service institutions.

With the growth of the Internet, users have to come to expect online ordering and online search and browse capabilities, with electronic file transfers for data delivery. Users are no longer content to wait days or weeks for their data or information requests to arrive at NOAA, be subsequently processed, and mailed back to the user.

In response to user demand, NOAA has made a number of its most requested data sets available to users through the Internet to improve customer service and to reduce the costs for servicing user requests. By the end of 2001, many customers were both using new NOAA online ordering systems and receiving environmental data from online NOAA sources.

As online access to NOAA's data expands, the user's average level of technical sophistication and scientific expertise is changing. Online users are searching for information and answers to specific questions rather than access to data. The needs of the business community and industry are becoming more complex, seeking interrelated data and supporting information and documentation, rather than just seeking one particular type of data.

Technology Changes

NOAA has been an early adopter of emerging advanced computing and communications technologies to improve data management and data availability on the Internet and next generation Internet/Internet2.

In the mid 1990s, it became apparent that the lack of adequate Internet connectivity would be a major limiting factor in NOAA's ability to continue its delivery of data and information products to the public. Stretched network resources would dictate which new technologies would be incorporated into NOAA's operational systems and which would not.

In order that NOAA could continue to meet the expectations of the ever-growing user community, and to support its in-house network application requirements, a plan for improving management of NOAA's information technology resources was developed in October 2001, and specific objectives were identified.[9] These included:

- Developing a robust, common network environment with adequate bandwidth;
- Adopting technologies to improve customer service;
- Providing improved information technology security;
- Improving management of information technology resources, including maintaining a baseline of NOAA web servers; and
- Using high performance computing resources to run high-end, environmental system modeling applications with the rapidly increasing amounts of newly available data.

Operating at speeds about 1,000 times faster than the typical dial-up connection, or 100 times faster than the typical home digital subscriber line, NOAA's second generation network connections support the development of advanced applications that cannot be implemented using the current Internet. NOAA must address the challenge of transferring massive amounts of data in real-time.

One example of a NOAA Internet2 application currently is under development by NOAA's National Severe Storms Laboratory (NSSL). In collaboration with other NOAA, university, and private industry

Status of NOAA Environmental Data Management

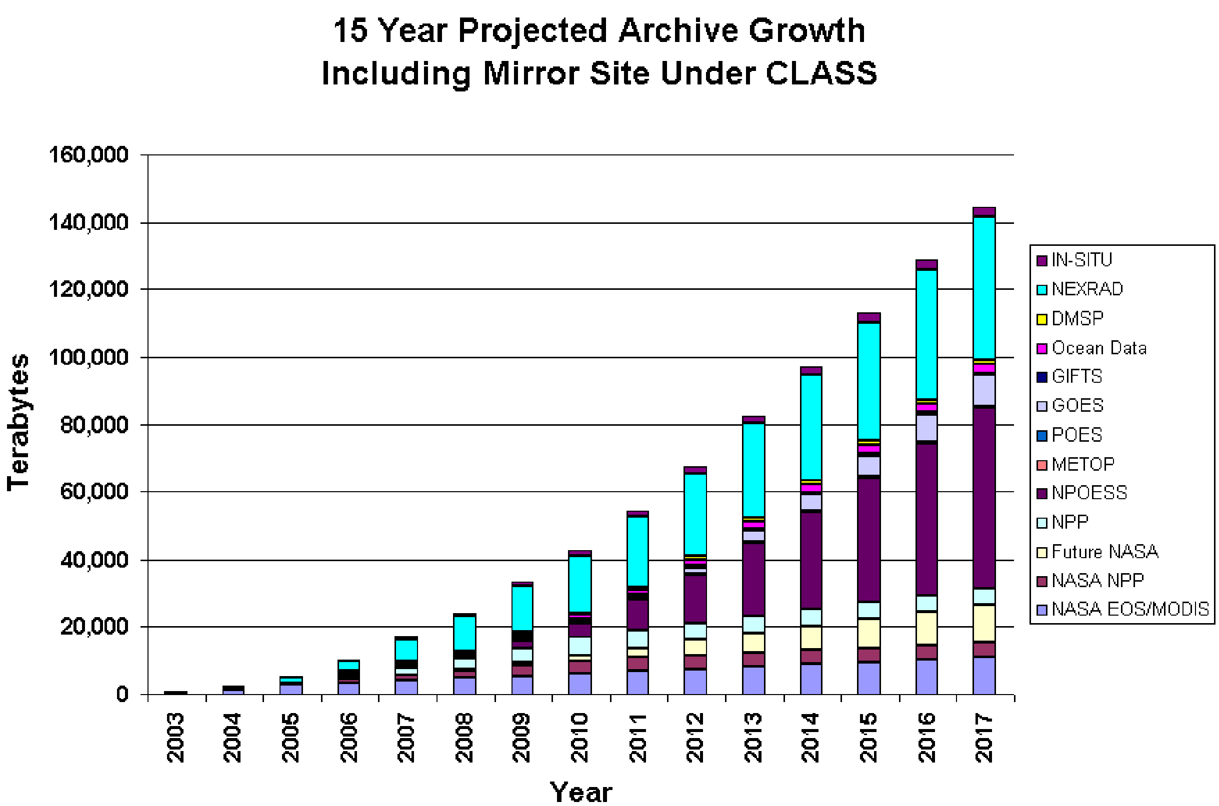
	Data Sets and Observations	End-to-End Environmental Data Management Functions							
		Planning	Collect or Rescue	Ingest	Metadata & Cataloging	Calibrate & Validate	Store	Access	Migrate
On-going	In Situ - Centers of Data	✓	✗	✓	✗	✗	☢	☢	☢
On-going	In Situ - NOAA National Data Centers	✓	✓	✓	✓	✓	☢	☢	☢
On-going	COOP/USHCN	✓	✓	✓	✓	✓	☢	☢	☢
On-going	GHCN	✓	✓	✓	✓	☢	☢	☢	☢
On-going	CARDS/COADS	✓	✓	✓	✓	☢	☢	☢	☢
On-going	DMSP	✓	✓	✓	✓	☢	☢	☢	✗
On-going	POES	✓	✓	✓	☢	☢	☢	✗	✗
On-going	ASOS	✓	☢	☢	☢	☢	☢	✗	✗
On-going	NEXRAD	✓	☢	☢	☢	✗	✗	✗	✗
On-going	GOES	✓	☢	☢	☢	✗	✗	✗	✗
Future	New In Situ Land & Ocean Observing Systems	✓	✗	✗	✗	✗	✗	✗	✗
Future	NPP	☢	✗	✗	✗	✗	✗	✗	✗
Future	NPOESS	☢	✗	✗	✗	✗	✗	✗	✗
Future	EOS	☢	✗	✗	✗	✗	✗	✗	✗

✓ = Can Do With Current Resources
☢ = Need Incremental Resources
✗ = Requires Substantial Additional Resources

Valid as of December 31, 2002

partners, NSSL's experimental Collaborative Radar Acquisition Field Test (CRAFT) project provides real-time compression and real-time delivery of NEXRAD data from multiple radars to users including NOAA's NCDC. This automation has dramatically increased the ability for NOAA to successfully archive as much as 30 percent more NEXRAD data than was previously possible (65 percent to 95 percent). NOAA's National Weather Service (NWS) has decided to adopt a CRAFT-like approach for all the radars in the continental United States. NSSL is helping to transfer this technology to the NWS.

Another prototyping effort to use large-scale networking within NOAA has been undertaken by NOAA's NGDC. In partnership with NASA, the NGDC has been testing and prototyping the operations involved in transferring and ingesting hundreds of gigabytes of satellite data per day.



Increasing Complexities and Volumes of Data

NOAA must ensure and maintain a healthy infrastructure that is capable of the ingest, archive, curation, access, and dissemination of its data sets. Under NOAA's archive and access architecture, its data management strategy is to accomplish two goals: (1) replace the assortment of aging and inefficient systems it currently operates; and (2) incorporate all of the data management functions shown in the accompanying "Status of NOAA Environmental Data Management" table. The table contents are based upon the status as of December 31, 2002. (See Appendix C for definitions of data sets.)

Over the next 15 years, current and planned remote sensing observing systems will produce volumes of environmental data on a scale that has not been seen before. Data from these systems will be preserved and made available by NOAA to support a myriad of users. By the year 2017, plans for the current NEXRAD, GOES, Polar-orbiting Operational Environmental Satellite (POES) for NOAA, including the Defense Meteorological Satellite Program (DMSP), and EOS/Moderate Resolution Imaging Spectrometer (MODIS) campaigns, future European Meteorological Operational Satellite (METOP), NPOESS, and NPOESS Preparatory Program (NPP) campaigns and numerous *in-situ* observation programs will increase the total data volume (primary and backup copies) to more than 140 PB.

NOAA's Comprehensive Large Array-data Stewardship System (CLASS) Project is designed to enhance NOAA's capability to predict and assess decadal to centennial climatological changes. It will provide various environmental data and information archive and access services to the Nation through the effective application of modern, proven techniques and technology.

A large portion of the Nation's current archive of environmental data is stored and maintained by NOAA's NNDCs and Satellite Active Archive (SAA). To prepare for the large increases in data volumes over the next 15 years, NOAA must increase the data-handling capacity and capabilities of its Data Centers. The CLASS project will afford efficient management of high volumes (petabytes) of data critical to the Climate Change Science Program, United States Global Change Research Program, and scientific community. It includes the development and implementation of standardized archive methods, which will be integrated with a robust, large-volume, rapid-access storage and retrieval system that is capable of receiving a user's online data request, automatically processing that request, and providing the requested data via the most appropriate medium. This system will provide standardization in security, media, interfaces, timeliness, formats, and processes for the very large data sets produced by satellites and radars.

Effective systems must be in place to support scientific research in the government, commercial, educational, and private sector communities, as well as address data management issues that are associated with massive volumes of data. NOAA will implement an architecture for an integrated, national environmental data access and archive system to support the ingest, archive, access, and distribution of its environmental data and information. The size, number, and frequency of data sets to be stored and distributed will require significant expansion of capacity for moving, storing, processing, and distributing the data.

Another area of major emphasis is the Nation's marine sanctuaries. Both the volume and the complexity of the data being gathered from these areas will increase, and management of these new data sets will require additional investments that are presently unidentified. In addition, NOAA is involved in characterizing and mapping freshwater, estuarine, and marine species communities and their habitats on local, regional, and national scales to meet mandated responsibilities for Essential Fish Habitat regulations. This effort is growing rapidly and requires significant collection and management of spatial data in GIS-compatible formats.

Less than 50 percent of the hydrographic surveys are available in digital format for use in GIS applications. Paper chart and source data archives need to be scanned, geographically rectified, cataloged, supplemented by metadata, and made accessible to users to facilitate charting efforts, environmental research, and environmental resource management. Although a great deal of historical tide, water-level gauge height, and water current data have been the subject of previous data rescue efforts, many more data remain in need of rescue.

Rescue efforts are underway to convert more than 10,000 historical shoreline maps into digital format. However, the rescue of 600,000 aerial photographic images of coastal areas, airports, and some sites of natural disasters has not yet been addressed. Much historical data of importance to understanding fisheries stocks (e.g., commercial vessel logbooks and old data sheets from fisheries research vessels) are in danger of being lost due to deteriorating physical media.

The increasing volumes of incoming data are outgrowing NOAA's storage capabilities and technologies, and they are affecting NOAA's data-rescue capabilities as well. With new media and mass storage technologies being introduced at an ever-accelerating pace, NOAA's ingest and storage capabilities – as well as its media migration capabilities – are falling farther and farther behind modern technology. There is an urgent need to migrate much of our archived data to new media before these data are unrecoverable.

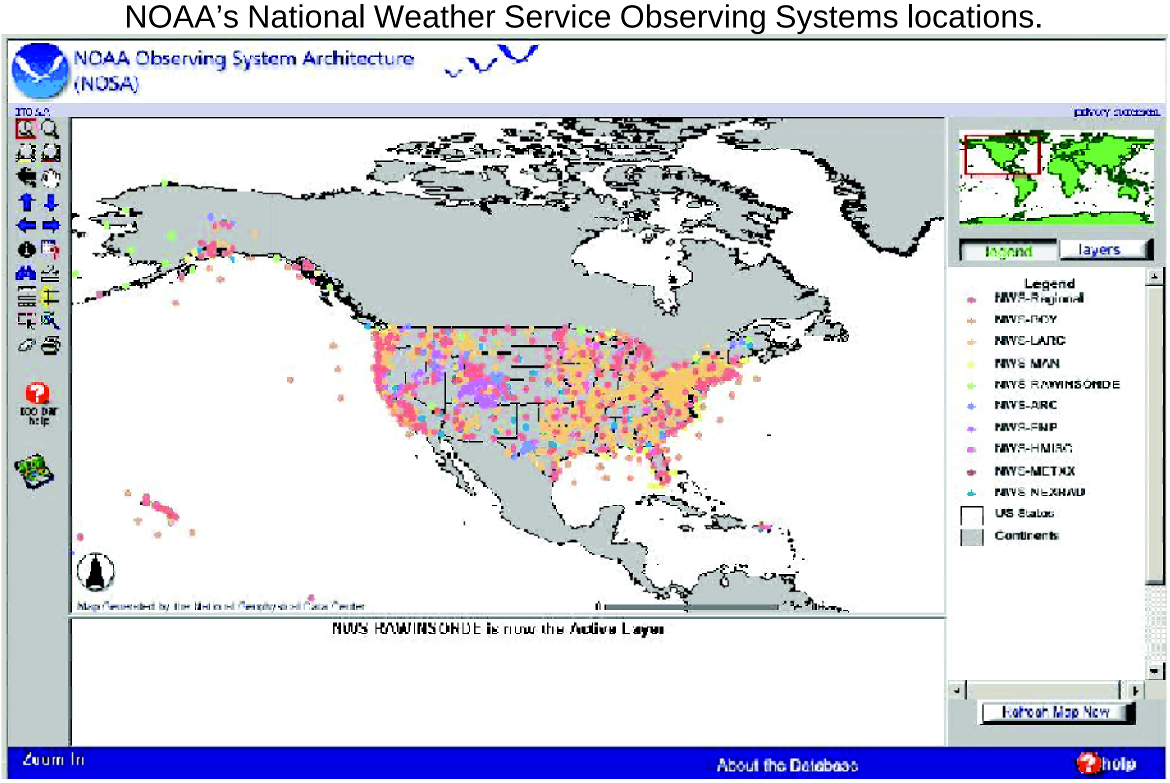
Seamless Uses of Geospatial Data

Geospatial data covers information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the Earth. This information may be derived from, among other things, remote sensing, mapping, and surveying technologies.

To group spatially-enabled data from a variety of environmental observations into a seamless database requires common horizontal and vertical geographic references. Unfortunately, the variety of geodetic and tidal vertical datums in use around the Nation has severely limited the usefulness of bathymetric, topographic, and shoreline data.

Such data can only be integrated if they are referenced to the same geodetic datum. Thus, for NOAA to integrate bathymetric and topographic data sets with those data sets from other mapping agencies, the data must be easily and accurately transformed to a common vertical datum. Important applications, such

NOAA's National Weather Service Observing Systems locations.



as storm surge modeling, hurricane evacuation planning, and the permitting activities of coastal resource managers depend upon integration in a standardized GIS. This capability does not presently exist in NOAA, and must be developed.

One impetus for the need for an enterprise-wide GIS in NOAA has been the development of the comprehensive NOAA Observing System Architecture (NOSA). NOSA was developed to enable NOAA to document its multiple observing systems and identify ways in which to evolve them in an integrated management approach. A key component of NOSA is a GIS capability that allows users to pose complex queries of a geospatially-enabled database of observing systems, resulting in easily interpretable maps of information.

The core principles for a NOAA Enterprise GIS will build upon existing standards and efforts, most importantly the Office of Management and Budget's (OMB's) E-Government Geospatial One-Stop portal, and the interoperability standards and protocols developed in consultation with the OpenGIS Consortium. Geospatial One-Stop is a part of the new OMB E-Government initiative to improve effectiveness, efficiency, and customer service throughout the Federal Government. The strategy, adopted by the President's Management Council in October 2001 implements the "Expanding Electronic Government" reform outlined in the President's Management Agenda.

NOAA is a full federal partner in the development of framework data standards associated with the GeoSpatial One-Stop effort as well as the inventorying, documentation, data collection, data discovery, and access to these framework layers. The Geospatial Information One-Stop Project will enable citizens and government to go to a single location to access Federal Government and other geospatial data assets. There are a number of challenges relating to Geospatial One-Stop that include:

- Better involvement of state and local governments and the private sector in an effective Geospatial One-Stop data standards development, data collaboration, and portal design process while maintaining traceability to business requirements;
- Better facilitation of improved geospatial data access and collaboration via the Geospatial One-Stop portal and other mechanisms;
- The development of policies regarding appropriate private sector use of the Geospatial One-Stop portal, for example, those dealing with the licensing of data to vendors;
- The development of interoperable web GIS interfaces and services, such as mapping and analysis, for the portal; and
- Anticipation of user demands for geospatial data access through the portal.

The success of Geospatial One-Stop will be dependent upon cooperation and collaboration between different agencies of the Federal Government; and between the federal, state, and local governments, and the private sector (e.g., utilities).

Scientific Stewardship of Data

Effective utilization of climate data and information and long-term stewardship requires the ability to manage the Nation's climate records scientifically and provide relevant utility to a wide range of customers. Scientific stewardship can be characterized as maintaining scientific integrity and long-term utility of climate records through monitoring, improving quality, and the extraction of select key parameters from new observations and the historical records. Scientific stewardship is principally a data management discipline designed to maximize the benefit of data beyond the initial and immediate short term uses. It is an important process encompassing the transformation of data to meaningful information, information to knowledge, and, ultimately, knowledge to understanding. Knowledge and understanding enhance the formulation of sound economic and environmental policies and decisions. The concept of scientific stewardship within NOAA means providing environmental data and information services necessary to answer global change science questions of the highest priority, both now and in the future. The practice of scientific stewardship is essential for maximum utility and long-term preservation of data records, and will lead to increased levels of confidence in climate model projections.

Scientific stewardship, with an emphasis upon satellite and radar data and information, is a new era in data management consisting of an integrated suite of functions to preserve and exploit the full scientific value of environmental data and information entrusted to NOAA. These functions provide effective observing system design, careful monitoring of observing systems performance for use in long-term applications, improved quality control, generation of authoritative long-term quality data and records, assessments on the state of the environment, and archive and access to data, metadata, products, and services. Successful implementation of scientific stewardship will ensure maximum use and public benefit of environmental data and information, now and in the future.

Maximizing the public benefit of the Nation's operational satellites, radars, and past proxy data sets requires implementing the scientific stewardship program. The involvement and consensus of the research community are essential to accomplish these objectives. Implementation of stewardship of environmental data from the NOAA observing systems will ensure the quality, usefulness, and accessibility of this national information resource treasure for current and future generations.

Homeland Security

NOAA, as with many federal, state, and local agencies, as well as private industry, faces the challenge of doing business in a world of heightened security concerns and terrorist threats following the horrific events of September 11, 2001. Clearly, NOAA found itself in a new situation regarding data and information needs, customers, and procedures. Many of NOAA's data sets are used to support forecasting of and response to natural hazards, as well as their use in climate studies, marine applications, and a variety of industries (e.g., construction), and are essential in support of homeland security activities that include emergency response, monitoring, predicting, and modeling in times of national emergencies.

For example, as gateways to our largest cities and industries, U.S. seaports are vulnerable choke points and strategic targets for attack. The U.S. economy is dependent upon the uninterrupted and efficient flow of goods and services. Commercial ports also serve as logistical centers for rapid deployment of U.S. military forces and material, and must therefore remain open and protected. NOAA can offer assistance that is of great benefit to the U.S. Coast Guard, the U.S. Navy, and port authorities through its mapping and charting products.

NOAA's models of atmospheric, oceanographic, and water quality conditions provide crucial advance data for re-routing of vessel traffic, port condition forecasts, and low visibility navigation to keep traffic moving. They also are critical to air and water dispersion efforts. As an example, NOAA works closely with first responders by providing modeling information for incidents such as oil spills, toxic atmospheric conditions, and smoke plume trajectory so that decisions related to population and environmental safety can be made. Data once openly available may now be considered sensitive; posing challenges for how

NOAA will maintain its “full and open” policy for data and information in a new era of heightened sensitivities.

NOAA is taking these initial steps:

- The agency has implemented interdiction software as a way of trying to limit selling data and information to known terrorists;
- It is reviewing its Internet sites for sensitive information (launch schedules, modeling programs, etc.) which might be used adversely. This type of information is being removed or put in protected areas and customers are being dealt with off-line;
- It is improving protections for Internet sites through information technology methods that provide mission partners with their products and services while ensuring security of the information being delivered.

The problems NOAA faces to ensure its data are available when needed are numerous. Certain data streams must be delivered in real-time and, therefore, clean and uninterrupted communications are essential. Robust and redundant lines of communication must be established in advance of any emergency and tested on a regular basis. Systems must be put in place for 24x7 quality assurance of real-time data, based upon defined parameters, with the ability to discontinue data dissemination if quality or accuracy are in question. Systems need to incorporate knowledge-based expert software. Continued enhancements will be required as new sensor data are added. Also, back-up delivery and archive sites must be maintained to deliver the critical non-real-time data in a timely manner.

Conclusion: Vision for Earth

The vision NOAA Administrator Vice Admiral Lautenbacher stated in the introduction is part of his broader vision for global observations. He indicated his support for an integrated global observing system in separate speeches delivered to the Intergovernmental Ocean Commission and the World Meteorological Organization in June 2002; a key excerpt from those speeches follows:

> **I strongly believe that NOAA is the right agency to take a leadership role within the United States, but we know full well that we cannot do this alone. The global observation effort for climate is far too enormous for one organization, or even one country, to undertake alone. We must work together. Perhaps the greatest challenge is to develop one integrated observation plan for the atmosphere, ocean, and land which everyone can support. The Global Climate Observing System and Global Ocean Observing System, working with the Integrated Global Observing Strategy Partners and others, have developed international consensus on overall needs. There is, however, much work still to be done. This challenge lies in our ability to provide one coherent plan which integrates space and *in-situ* observations across those three elements.**

Toward this end, NOAA is taking a prominent role, partnering with NASA and other U.S. agencies to bring this global perspective into achievable reality. The initial effort began with an Earth Observations Summit on July 31, 2003. The need for this was widely recognized and provides the benefit of a sound plan for end-to-end stewardship of environmental data. This is a challenge and NOAA is on target to step up to this opportunity.

[Page left intentionally blank]

Introduction

The National Oceanic and Atmospheric Administration (NOAA) has a core mission to understand and predict changes in the Earth's environment and to conserve and manage coastal and marine resources to meet the Nation's economic, social, and environmental needs. The data and information from this network must then be organized, permanently preserved, then provided to scientists and researchers for study and tracking of the Earth's constantly changing environment.

NOAA Environmental Data: Foundation for Earth Observations and Data Management System describes the current state of, and future challenges for, the management of an ever-expanding environmental database pouring into NOAA from a world-wide data collection network and the progress towards bringing it all together to address the complex issues of today's world.

As required biennially by Congress, this Report assesses the adequacy of the environmental data and information systems of NOAA. The Report serves as an update to *The Nation's Environmental Data: Treasures at Risk*[1] provided to Congress in September 2001.

In this first decade of the 21st century, NOAA oversees a spectrum of environmental data management services that includes:

- Ecosystem Management
- Commerce and Transportation
- Hazards
- Climate Monitoring
- Ocean Data Management

As we plan for the second decade of the century, NOAA foresees cross-cutting challenges that apply to all NOAA-managed environmental data. These include:

- User Needs
- Technology Changes
- Increasing Complexities and Volumes of Data
- Seamless Uses of Geospatial Data
- Scientific Stewardship of Data
- Homeland Security

This Report addresses these data management services and challenges throughout NOAA. Our success at meeting the ever-increasing data management requirements will determine NOAA's success in meeting its primary mission goals. In addition to monitoring and observing the environment through investigation and interpretation, these services and challenges include understanding and describing how natural systems work together; assessing and predicting the changes of natural systems, and providing information about the future; engaging, advising and informing individuals, partners, communities, and industries as required and facilitating in the exploitation and application of appropriate data; managing coastal and ocean resources to optimize benefits to the environment, the economy, and public safety.