Section 3.

Prospects for Success: Barriers and Challenges

For NOAA to successfully respond to the changes discussed in the previous section, several significant barriers and challenges must be overcome:

- Data management resources have been declining while the volume of data has been growing.
- Current data ingest and archiving methodologies and processes cannot handle the massive amounts of new data coming into NOAA for archive.
- Valuable data on aging media are essentially unusable and are at risk of being lost forever. Capital investments are needed to implement new storage technologies.
- Old systems are failing to meet the demands of industry for timely data. Observing systems need upgrades to insure the continuation of environmental monitoring for critical data sets.
- Most existing data sets do not conform to National Spatial Data Infrastructure (NSDI) specifications and do not have associated Federal Geographic Data Committee (FGDC)-compliant metadata. Limited resources have inhibited the adoption of the NSDI policies and it will take considerable resources to become FGDC-compliant.

The Nation is losing the data resources in which it has already invested. Academia, especially global-change researchers (and in turn industry, policy makers, and the public), are losing access to data that the Federal government has spent billions of dollars to collect.

Supporting New Environmental Observing Programs

NOAA plays an integral role in meeting the data management needs of programs that are attempting to address the status and trends of Earth's environment, both nationally and internationally. New large-scale, long-term environmental observing programs are in various stages of planning. Once in place, these new programs will put great demands on NOAA data management activities, requiring higher processing and storage capacities in addition to improved access capabilities.

By working closely with national and international organizations, NOAA is leveraging resources to produce and analyze data to achieve NOAA's goals, and to further the interests of the U.S.

However, NOAA's resource limitations are threatening its ability to be a successful partner in these cases, and many other opportunities to leverage additional cooperation are at risk.

For example, the Climate Variability (CLIVAR) Program will address the oceanic and atmospheric phenomena that result in climate change. CLIVAR has its roots mainly in the World Ocean Circulation Experiment (WOCE) and will build upon what was learned during the WOCE. Activities in CLIVAR data management—coupled with the ongoing WOCE data management activities—will overlap. The Global Ocean Observing System (GOOS), a worldwide network of *in situ* measuring networks coupled with satellite remote sensing, and the U.S. Coastal GOOS, will have a combined effect of overwhelming the present capacity of the NOAA National Data Centers (NNDC).

Acceleration of Technology Change

Advances in information technologies are occurring at ever-increasing rates. Moore's Law—first described by Intel, Inc., co-founder Gordon Moore—states that computer chip performance doubles every 18 months. Storage technology companies have both promised and delivered a doubling of media storage capacity every 2 years. Networking companies are doubling transmission rates annually.

Technological advances have enabled the collection of increasing amounts of environmental data. However, the immense volume of data coming into NOAA is outstripping the ability to process these data sets for archive, and subsequently make them accessible. NOAA has not had the resources to respond to the advances and implement the needed ongoing technology refreshment plans. NOAA is being buried under the very technology that is enabling the improvements in environmental data collection.

It was once acceptable to migrate archived data to new technologies approximately every 10 years. Now migration must begin every 2 to 3 years, or data will be lost within the old technology. As data volumes increase, the time required to accomplish data migration increases. Undertaking a policy of ongoing mass-storage technology refreshment is needed to eliminate very costly, large-scale total system replacements.

Technology changes also affect NOAA's ability to deliver data to users. Some user groups find it difficult to keep up with technology due to their own funding issues. In these cases, NOAA must be able to provide data via a wide variety of alternative delivery media.

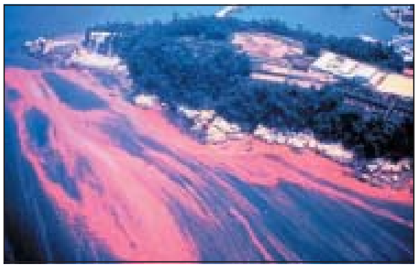


Photo credit: Sulsan Aviation Company.

Harmful Algal Blooms Cost the Fishing Industry Millions

Each year harmful blooms of algae in the oceans and estuaries of the U.S. cause massive fish kills, which can cost the fishing industry millions of dollars a year and raise the price of fish in the supermarket. These algal blooms can also be harmful to human health.

NNDC databases help researchers to understand the conditions in which the blooms occur, but more work is needed. With the appropriate resources, the NNDC could serve as a central repository for data and information from historical blooms that can be used by researchers and planners to predict, or reduce the severity of, these algal blooms. Savings of millions of dollars are possible for American industry from this type of research.

Paleoclimate records—derived from tree rings, ice cores, corals, ocean and lake sediments, and other natural recorders—can provide the long-term perspective on the physics of Earth's climate that is required to improve our understanding of climate and climate change.

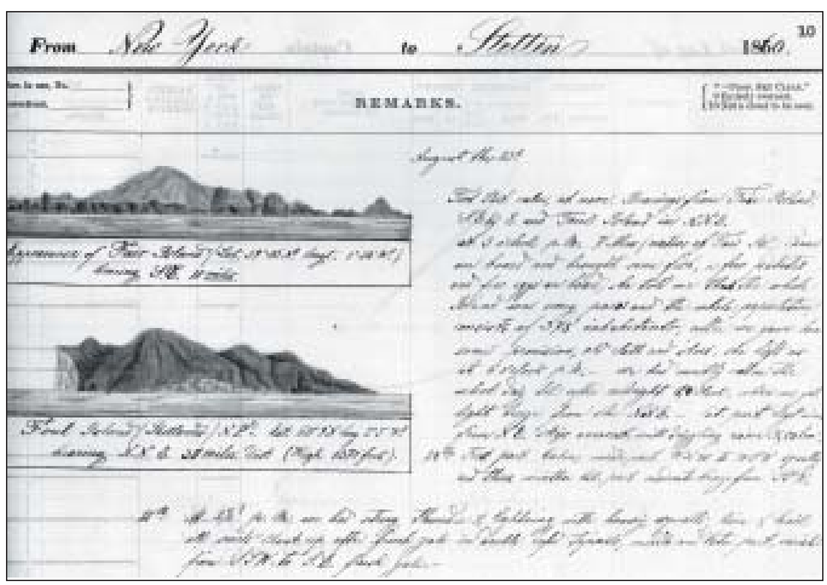
Deteriorating Historical Archives

NOAA maintains hundreds of millions of environmental records from a myriad of sources. Proxy data from sources such as ice core samples, coral skeletons, and tree-ring samples provide insight into various environmental events over hundreds and thousands of years. *In situ* weather observations that have been manually recorded at hundreds of locations over time—beginning with Thomas Jefferson in the 18th Century and including today's volunteer network—provide a 200-year record of the Nation's environment. Data collected by remote sensing systems, such as radar and satellites, have been providing additional global environmental records for the past 40 years. Hydrographic and fish surveys covering waterways provide a history from the mid-1800s.

These data often provide highly credible, century-scale, environmental data. Without preservation of the media, the data are in danger of being lost. Without digitization, the data cannot be used effectively.

Paper becomes brown and brittle and eventually crumples to dust. Microfilm and microfiche eventually go through a chemical deterioration process, making the data unreadable. Digital records contained on tape and disk media will eventually deteriorate with age, or these records will be unreadable by newer technology mass-storage subsystems.

Frequently, the preservation of heritage data suffers from a lack of support because its value is not fully appreciated. Recent experiences have shown that data originally captured for one purpose can have totally different applications years later—applications that could not



Page from the log of the Prussian brig Elise*, dated 1860. Old logbooks contain valuable and historical observation data which are being lost due to crumbling paper media.*

have been foreseen when the data were originally gathered. For example:

- The NOAA Advanced Very High Resolution Radiometer (AVHRR) was designed to produce cloud imagery and sea-surface temperature data to support the National Weather Service weather-forecasting operations. The data are now also used to derive a “greenness index” that shows the health of vegetation in the U.S. related to drought conditions, and fire susceptibility and risk for areas such as the great plains of the U.S. AVHRR data are also used for ice-mass studies to detect ice movement and the formation of large icebergs, and to measure ice-surface temperature. Over the years, the AVHRR data have been reprocessed using new algorithms, producing new data sets. One such project, jointly undertaken by NOAA and NASA, reprocessed more than 14 years of AVHRR data to produce consistent sets of oceanic, land, and atmospheric data for global change research.

- Recently, a researcher reprocessed some of the early GOES data from the late 1970s using a new algorithm. He discovered that he could extract high-resolution wind data that were not originally available from that data set. This finding showed that GOES data can provide an entire 20-year wind climatology for the U.S. However, more than 20 years of the GOES satellite archive is stored on obsolete and rapidly deteriorating tape media. The equipment required to retrieve the data on these tapes is becoming very difficult and costly to maintain.

- Oceanographic and biological data have been collected during fish and hydrographic surveys, and from light ships and tidal stations since the mid-19th Century. These data have the potential to support research in the emerging field of oceanographic oscillations—decadal shifts in basin-wide, or even global, oceanographic conditions. These shifts could have profound biological and climatic impacts, such as sharp declines in endangered Pacific salmon stocks and Steller sea lions, as well as declines in North Atlantic groundfish. However, most of the historical data needed to conduct the relevant research remains inaccessible in old logbooks that are deteriorating and are in danger of being lost.

This underscores the need to archive and preserve data. Only with such an archive is it possible to produce a consistent data set of scientifically important parameters over the longest possible period, providing the ability to reprocess the data repeatedly. Migration to new digital media is a requirement to increase data accessibility, and to enable the integration of historical data sets into the global

Scientists review seismic data to determine the best location for a sea floor core. Analysis of sediments is crucial to unlocking the mysteries of a region’s climatic history.

databases that are required for today's global change research projects.

Aging Observing Systems

Some of the Nation's most basic observing systems are aging, and are at risk of failure. Perhaps the most sought-after and valuable data that support climate-monitoring and sustainable economic activities are collected and reported by the nationwide, volunteer-based COOP Network. NOAA receives more than 100,000 data requests each year from industry, business, and government users who depend upon these data.

The COOP Network provides valuable input to daily weather forecasts, is used to calibrate NEXRAD Doppler radar data, and is the basis for the emerging Weather Derivatives industry discussed in Section II of this report. The Network is based on 1950's technology, and its instruments are failing. The volunteers manually record instrument readings on paper and send them to NOAA. Then, NOAA spends more than $1 million each year to digitize these records to safeguard the data, and to make the data usable for NOAA's many users.

One part of this network—the COOP 15-Minute Precipitation Network—collects rainfall data in 15-minute increments and is key

The volunteer-based Cooperative Network stations have changed little over the years. Left: A cooperative weather station at Granger, Utah, circa 1930s. Right: Today's COOP Network observer.

to identifying extreme rainfall events. Instrument readings are recorded on paper-punch tapes that are collected by volunteers. Then, these tapes are sent to NOAA where they are "read" by a device that subsequently outputs the data into digital records. NOAA spends substantial sums each year processing these paper tapes. The device that reads the paper tapes (produced by the only high-resolution, precipitation-measuring gauge) are old, unreliable, and require frequent maintenance. The transcription is labor-intensive and difficult to learn.

The COOP problem is a good example of how the data management system seriously impacts the need to modernize, upgrade, and refresh both the network and the observing system. Sites with long, continuous observations and good documentation—such as those in the COOP Network—provide meaningful data continuity records for climate change monitoring, agricultural and business planning, disaster planning and avoidance, and long-term trend analysis. Therefore, the establishment of a climate reference network that is selected from the COOP Network must be a high priority.

Increasing Data Complexities

The management of data sets is a function of data complexity, as well as volume. For example, a well-defined data stream from a single source, such as NOAA AVHRR sensors, produces high volumes of data that require little effort to make them accessible once processing algorithms have been established. On the other hand, oceanographic data sets are produced by many different types of sampling instruments—hundreds of which are placed in the water to record observations. The accumulation and assembly of thousands of data sets from these instruments is labor-intensive work that produces an invaluable resource. However, based solely on volume, this may not appear to be significant when compared to another data set.

Data complexity is substantially increased for remotely sensed data sets, in part because these data require *in situ* observing systems' data for calibration and validation. Satellite instruments observe Earth's surface and atmosphere through the entire depth of the atmosphere; therefore, any atmospheric contamination (e.g., clouds, tiny particles from a volcanic eruption, etc.) limits the instruments' clear vision of the atmosphere near Earth's surface. Furthermore, satellite sensors and their platforms tend to drift over time. *In situ* surface observations are imperative for the calibration of the satellite sensors, to filter and adjust for space-borne contaminants and satellite drift.

NOAA Data Assists FEMA in Planning

NOAA develops statistics on snowfall extremes for thousands of sites across the U.S. The data are used by the Federal Emergency Management Agency (FEMA) during severe winter storms. Weather statistics, along with NOAA satellite data and in situ observations, also provide guidance in FEMA's declaration of areas targeted for emergency relief funding.

As the amount, diversity, and complexity of data archived by NOAA continue to grow, the task of maintaining these data sets becomes more difficult.

In addition, data from remote sensing systems must be referenced to specific geographic locations on Earth's surface. Remotely sensed data from surface-based sites—such as radars or wind profilers—also require data from *in situ* observing systems for data calibration and validation. For instance, Doppler Radar data require NWS Cooperative Observing Program (COOP) Network data to calibrate rainfall relationships. Any occurrence that impacts satellite readings must be documented so that users will not misinterpret the data.

As the amount, diversity, and complexity of data archived by NOAA continue to grow, the task of maintaining these data sets becomes more difficult. As the complexity of the archived data grows, the skill sets required for the personnel who maintain these data sets are changing.

Data Availability and Cost Issues

There is no uniform policy guiding the use and re-use of NOAA environmental data. OMB Circular No. A-130, *Management of Federal Information Resources*, provides guidelines for data held by the Federal government. However, these guidelines are so flexible they allow for different interpretations, resulting in a variety of implementations throughout the government.

Within the various components of NOAA, data policy has been based upon the interpretation of these guidelines and a body of precedents developed over time. For example, catch data that are voluntarily provided by members of the commercial fishing industry are considered proprietary—therefore, they may be disseminated only in aggregate form or not at all. Also, the sharing of satellite data is governed by a variety of agreements with other commercial and international satellite operators. Some meteorological data obtained from non-NOAA sources may be used by NOAA in research, but these data may not be redistributed for commercial use.

The issue of data sales affects many aspects of NOAA's mission to archive and disseminate environmental data. NOAA has been mandated through the appropriations process to charge users for the dissemination of data. The collection of these fees help offset some costs of services.

Several attempts have been made to revise and standardize pricing policies within NOAA. In 1981, pricing guidelines were established to formalize and document the process of pricing data. Beginning in 1991, there was discussion about the possibility of charging Fair Market Value for data disseminated from NOAA's Data Centers.

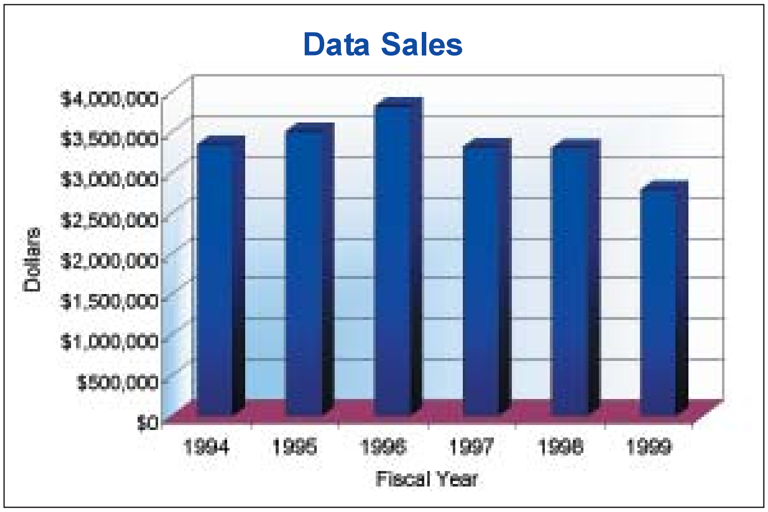
In 1996, 15 U.S.C. 1534 authorized the Secretary of Commerce to charge Fair Market Value for data. The aim of this legislation was to enable NOAA to realize additional income to cover budgetary deficiencies and to authorize standard prices across the three Data Centers. Congress reduced NNDC funding in anticipation of increased data sale revenues that never developed (and actually diminished).

The three NOAA Data Centers began to charge the same prices for similar products in 1997. In addition, two-tiered pricing was introduced, resulting in a discounted price for educational organizations and orders placed through NNDC's on-line store, and a higher price for commercial entities and orders not placed through the on-line store.

In some instances, the cost-recovery data policy is affecting the completeness of the archive. It has alienated a few Principal Investigators—the lead scientist for a research project or program—and the university research community on whom NOAA relies (and with whom NOAA partners) to develop new scientific applications of the data for the benefit of the Nation. These are the users who bring value to data sets by interpreting them in context with other data sets to solve a problem. These users are also concerned that they will not be able to access data given to NOAA for archive, as NOAA gets farther behind in making new data sets accessible.

Data sales, as a source of revenue, have been disappointing, and revenues continue to fall on an annual basis by approximately 6 percent. In addition, NOAA now finds that data for which it is charging are being distributed throughout the world by a multitude of Web sites, including those of other Federal government agencies. There are many instances in which NOAA is required to charge for data and information that are being provided free-of-charge by other Federally-supported projects. Furthermore, the cost of maintaining a

In 1996, Congress authorized NNDC to charge Fair Market Value for data to cover budgetary deficiencies. Congress reduced NNDC funding in anticipation of increased data sales that never developed (and actually diminished).



Gross NNDC data sales have declined by one-third in the past four years.

fee-based service is comparable to funds received; as data sales revenues continue to decrease, soon it will cost more to recover fees than to provide a free service. These issues are requiring NOAA to re-examine its data sales policy.

Providing End-to-End Data Management

Basic functions are necessary for every data set that is developed or acquired by NOAA. These functions include various quality control procedures; data verification and metadata creation; cataloging; storage; and provision of access. Today, many of these processes are manual, requiring extensive staff resources. A new approach is required to manage the immense volumes of data that are being acquired throughout NOAA—and particularly those data that are to be archived and made accessible for dissemination to users.

The concept of "end-to-end data management" is accepted as a critical approach to ensure the viability of data for long-term use. All too often, however, the value of scientific information beyond its immediate real-time use has been overlooked in the budget process. The result—as shown in the table (next page) *Status of NOAA Environmental Data Management*—is that NOAA needs capital investments and improved procedures to provide end-to-end data management functions. Systems need to process massive amounts of data that are now flooding NOAA for archive and dissemination to users. Budget considerations must also address data that will be coming to NOAA from the new remote sensing systems, and historical data that are not yet available in digital form.

The following is a brief description of the activities involved in each phase of the end-to-end data management approach:

- **Planning.** The initial phase of all environmental data-gathering activities includes the analysis and design of sensors to observe the physical characteristics of properties of interest. The historical record of these physical parameters is key to identifying areas for improvement in the sensors. Additionally, the need for efficient data management and accessibility contributes to the design of observation formats and to spatial and temporal resolution requirements. All aspects of data management must be considered in the initial planning—and in ongoing planning for modifications or improvements—for any instrument or sizable environmental data-gathering project. The interests of the project, the archive, and the secondary user community (i.e., those users who will come to NOAA for access to the data) must be considered and balanced in order to reduce the overall cost

Under the Endangered Species Act (ESA), NOAA is responsible for marine fish, marine mammals, and sea turtles. End-to-end data management is critical in accomplishing this mission. NOAA collects and analyzes data to determine whether a species qualifies for threatened or endangered status. For qualifying species, NOAA coordinates the development and implementation of a recovery plan. NOAA also works to ensure that the recovery of the threatened or endangered species is not hampered.

A sheet of water streams off the rounded tail of a humpback whale, a well known endangered species. It is estimated that the present worldwide population of the humpback whale is only about 10,000—only about 8 percent of the species' initial population. (Photo credit: Dann Blackwood, USGS.)

Status of NOAA Environmental Data Management

	Data Sets and Observations[1]	End-to-End Environmental Data Management Functions: Planning	Collect or Rescue	Ingest	Metadata & Cataloging	Calibrate & Validate	Store	Access	Migrate
HISTORICAL	*In Situ* [2]- Centers of Data	√	×	√	×	×	⊕	⊕	⊕
	NOAA National Data Centers	√	√	√	√	√	√	⊕	⊕
	COOP / USHCN	√	√	√	√	√	√	⊕	⊕
	GHCN	√	√	√	√	⊕	⊕	⊕	⊕
	CARDS / COADS	√	√	√	√	⊕	⊕	⊕	⊕
« MODERNIZATION »	DMSP	√	√	√	√	⊕	⊕	⊕	×
	POES	√	√	√	⊕	⊕	⊕	×	×
	ASOS	√	⊕	⊕	⊕	⊕	⊕	×	×
	NEXRAD	√	⊕	⊕	⊕	×	×	×	×
	GOES	√	⊕	⊕	⊕	×	×	×	×
FUTURE	New *In Situ* Land & Ocean Observing Systems [3]	√	×	×	×	×	×	×	×
	NPP	⊕	×	×	×	×	×	×	×
	NPOESS	⊕	×	×	×	×	×	×	×
	EOS [4]	⊕	×	×	×	×	×	×	×

√ = *Can Do With Current Resources*

⊕ = *Need Incremental Resources*

× = *Requires Substantial Additional Resources*

NOTES:

1. See Appendix D for description of these data sets.
2. *In Situ* Observations of the Environment. NOAA has hundreds of multi-thematic data sets that broadly cover the spectrum of climatology, oceanography, geophysics, and biology. These data are observed on land (e.g. surface readings and cores), at sea (e.g., on-board ships and buoys), and in the air (e.g., balloons and aircraft). In general, these collections are relatively small (less than 300 gigabytes) with poor-to-excellent access.
3. New *In Situ* Land & Ocean Observing Systems. New integrated observing systems have been identified as high NOAA, Commerce, and Executive priorities. These systems are currently in the initial planning phases.
4. The NASA Distributed Active Archive Centers (DAAC) will collect and ingest the data, provide some metadata, perform initial calibration and validation (including some reprocessing). However, NOAA must still ingest the data from the DAAC, update and maintain the catalogs, and facilitate or perform calibration and reprocessing, particularly with respect to other NOAA environmental data sets.

NOAA needs modern distributed processes and additional capital investments to provide end-to-end data management functions. Massive amounts of data are now flooding its Data Centers for archive and distribution to users.

of data management throughout the life-cycle of the data sets.

- **Collect or Rescue.** Data collection occurs in many modes, ranging from satellite communications (consisting of a high-volume stream of observations) to the rescue of observational notes recorded centuries ago. Many data sets are thematic in nature—encompassing numerous techniques, such as proxy (or surrogate) data for paleoclimatological research. Automated delivery systems that provide data to NOAA in real-time or near-real-time are needed for large, remotely sensed data sets or when there is a need for near-real-time data by the secondary user community.

- **Ingest.** When data are received, they must be processed into formats that lend themselves to preservation and access. This may require procedures as simple as tape copying, data transcription, or some form of analog-to-digital image conversion. In the case of remotely-sensed data, some level of calibration and validation may also be required, necessitating substantial computational power. For these data sets, the ingest function must be integrated with the system utilized by the program to deliver data, and the ingest processes must be automated to the extent possible in order to reduce the time and resources required to complete this function. A critical aspect of ingest is the assurance of adequate documentation and validation that incoming data formats are as described.

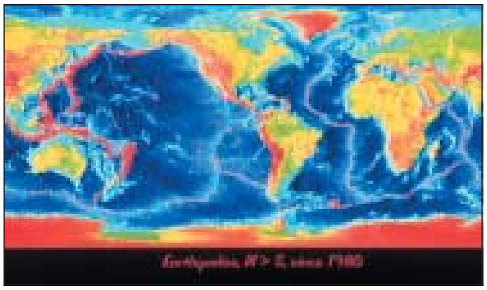
- **Metadata and Cataloging.** Data sets must be documented with metadata descriptions, and cataloged so that the data sets can be located and understood by users. A popular criterion, emphasized by the National Science Foundation, is that scientists 50 years from now should be able to completely understand the processing and use-constraints of the data. Metadata is the information about the data that makes the data understandable and usable. Preparation of standardized metadata (such as those endorsed by the National Spatial Data Infrastructure) for a data set can take anywhere from hours to years, and requires dedicated staff members who have the right skill sets. Once compiled, the metadata must be maintained and updated as instruments and methods change, stations relocate, processing algorithms change, or other factors affect the validity and consistency of the data contained in the data set. Catalogs are developed to inform potential users of the existence of the data set, its spatial and temporal extent, and to advise data managers of techniques to subset the collection in a manner conducive to end-user requests. These activities are labor-intensive processes that require personnel who have knowledge of both the data set and the observing system. Personnel are also required to document the program's history, or to develop the special-purpose computer scripts (programs and utilities) that read the data format and build the catalog.

- **Calibrate and Validate.** Often, data collections need to be adjusted to facilitate accurate understanding of the measurements. Examples include: orbit corrections for satellite data (ensuring accurate geographical registrations); reinterpretation of data to a common geographic reference field; or, reprocessing of data to incorporate improved noise-reduction techniques (such as the removal of ionospheric and water vapor signals from Global Positioning System data). Reprocessing activities include the incorporation of new data into existing data streams. Since these are computation-bound tasks for large-volume, remotely-sensed data sets, significant increases in computing power will be required to perform these tasks efficiently.

- **Store.** Inventory systems are maintained to keep accurate controls on the physical location of files and the lineage of corrections. Robust, large-volume, rapid-access data archival systems coupled with the use of modern, automated data management systems are needed to accommodate the massive volumes of environmental data being generated by the current and planned observing systems. Efficiencies in storage are realized by utilizing current technologies, formats and interfaces based upon accepted industry standards, and compatible platforms.

- **Access.** The accessibility of data is key to the value of data. A data set has no value if no one knows it exists, if it is on a media or in a format that makes it unusable, or if the cost associated with being given access to the data is higher than the user can afford to pay. In order to adequately respond to the growing user demand for access to environmental data, large-volume, fast-access systems must be utilized and integrated with data archive systems to provide access to the massive amounts of data. In addition, NOAA Data Directory nodes on the NSDI FGDC Clearinghouse must be updated and maintained to facilitate user search capabilities.

- **Migrate.** To preserve data for long-term future use, it is critical to continue to migrate data to modern "archival quality" media. This is particularly important due to the rapid advances in computer technology that occur in just two to three years. Otherwise, the data will be lost because of deteriorating materials or obsolete reader technology.



Earthquake History Basis for Major U.S. Mortgage Loan

A large U.S. insurance company was shopping for a mortgage loan to construct a new headquarters building. The most favorable rates were offered by a Japanese-based bank with the caveat that, if an earthquake of greater than a specified magnitude occurred within a 25-mile radius of the bank's headquarters, the loan would have to be repaid immediately. Before deciding whether to take this loan, the insurance company used NOAA data detailing historic earthquake events in Japan to determine the statistical probability of such an earthquake occurring.

Implementation of an end-to-end data management approach for data managed by NOAA is critical in maintaining the viability of the environmental data archived for the benefit of the Nation, for current and future use.

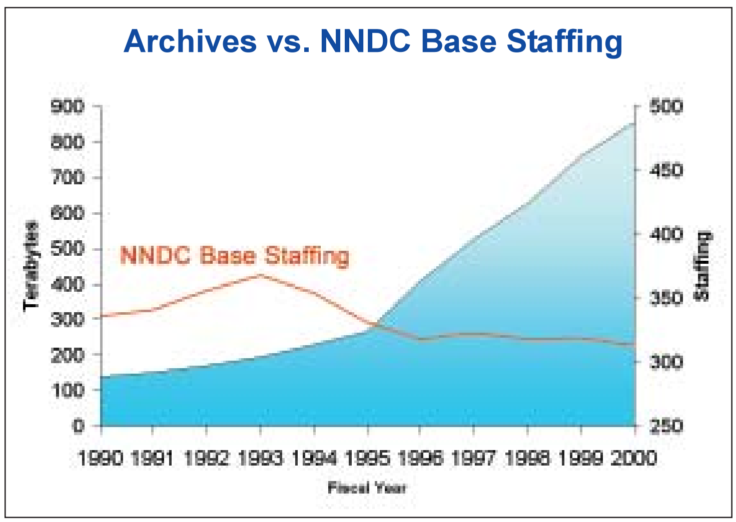
Maximizing Efficiencies

During the 1990s, NOAA made great strides in streamlining data management operations and in using technology to support ongoing operations. The development of interactive data access systems — such as the NOAAServer, the NOAA Virtual Data System (NVDS), and the National Geodetic Survey Products and Services Web site — have enabled many data sets to be made available to users via the Internet. They have enabled more user requests to be handled by a declining staff. Many archived data sets that were in danger of being lost due to aging storage media were "rescued" through migration to modern media in a highly successful Data Rescue Program. Significant resources were also expended in resolving Y2K issues in preparation for the Millennium change.

Although spending on environmental research programs increased over the years, NOAA took on the archive and management of more and more data without comparable increases in funding for basic operations—salaries and infrastructure. While the size of the NOAA archives quadrupled and data requests increased exponentially, a declining NOAA staff managed to support basic operations on existing data sets, as well as service most user requests.

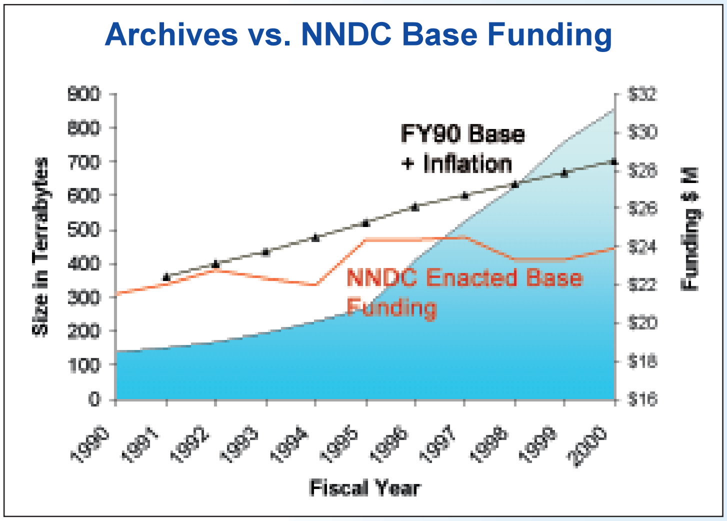
NOAA has been developing and storing many new experimental data sets with commensurate data management programs that are tailored to those data sets. New initiatives, spearheaded by NOAA's Coastal Services Center and the Coastal Data Development Center, depend

During the last half of the 1990s, NNDC staffing has declined by 14% while archives tripled in size.



upon easy access and manipulation of data by numerous State and local constituencies.

Temporary funding sources have been used to employ contract staff in an attempt to compensate for decreasing Federal government staffing, but the quality of service and the stewardship of data are beginning to suffer due to insufficient staffing levels. These sources are not adequate to provide the ongoing resources that are necessary for the ingest and maintenance of the traditional data sets, much less for the massive new data streams.



NOAA National Data Center (NNDC) funding has remained relatively level at a time of drastic archive growth.

Reimbursable work has been undertaken by the Data Centers for other NOAA components and non-NOAA agencies. However, rather than augmenting basic ongoing Data Center operations, these activities have frequently resulted in NOAA resources being redirected away from base services and operations in support of the special projects. Consequently, basic Data Center operations are falling even farther behind.

Inflation has taken its toll on NOAA's data management budget, and work loads at reduced staffing levels have reached a point of saturation. Many more data sets need to be brought on-line or near-line to meet user demands, and more data sets need to be migrated to new storage technologies to ensure their availability to support future user demands. The ever-growing influx of new data and the increasing demands from the user community are overwhelming NOAA's data management capabilities.

The demands of NOAA's environmental data management are stressing NOAA's capabilities and resources, threatening NOAA's ability to be successful in its mission.

NNDC Comparisons, 1979 vs. 1999

	1979	1999
Archives	20 TB	760 TB
User Requests (Accesses)	95,400	4,200,000
FTE Staff	582	321

In FY2000, NNDC had 5.3 million on-line users.