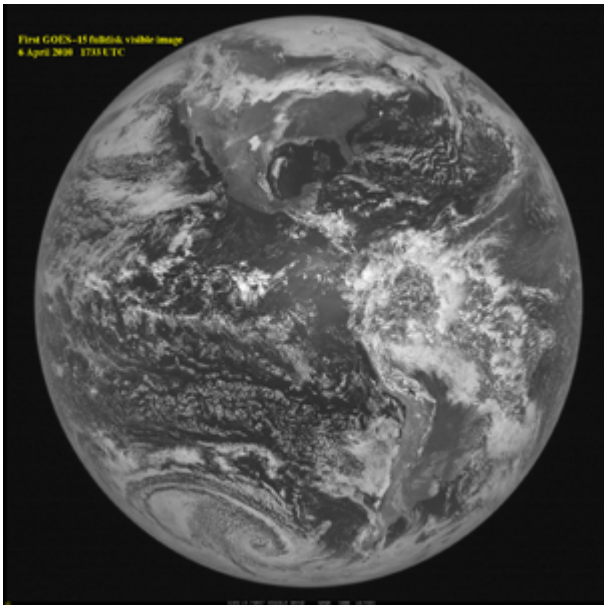
National Oceanic and Atmospheric Administration



Report to Congress
on
Environmental Data and Information Systems Management

2009

NOAA's Vision, Mission, Goals, Values

NOAA'S VISION

An informed society that uses a comprehensive understanding of the role of the oceans, coasts, and atmosphere in the global ecosystem to make the best social and economic decisions

NOAA'S MISSION

To understand and predict changes in the Earth's environment, and conserve and manage coastal and marine resources to meet our Nation's economic, social, and environmental needs

MISSION GOALS

- Protect, restore, and manage the use of coastal and ocean resources through an ecosystem approach to management
- Understand climate variability and change to enhance society's ability to plan and respond
- Serve society's needs for weather and water information
- Support the Nation's commerce with information for safe, efficient, and environmentally sound transportation
- Provide critical support for NOAA's mission

CORE VALUES

People, Integrity, Excellence, Teamwork, Ingenuity, Science, Service, and Stewardship

Foreword

I am pleased to present the National Oceanic and Atmospheric Administration's (NOAA's) Report to Congress on Environmental Data and Information Systems Management 2009. NOAA produced this report based on the requirements described in Public Law 102-567, Section 106.

NOAA's mission within the Department of Commerce (DOC) is to understand and predict changes in Earth's environment and conserve and manage coastal and marine resources to meet our Nation's economic, social, and environmental needs. In support of this mission, NOAA provides the critical end-to-end data management that includes archival, stewardship, and access to a large portion of the Nation's environmental data.

Environmental decisions require timely access to accurate data and data products. Often, decisions with far reaching economic consequences at local, regional, and global levels are made based on NOAA's environmental data. Data managed by NOAA get used every day by hundreds of thousands of users in government, commerce, industry, science, engineering, and national defense.

Over the past decade, the environmental data under NOAA's stewardship has grown exponentially. The exponential growth is expected to continue into the foreseeable future. This report details the progress made in developing and maintaining a data archiving and access system that will improve the quality and stewardship of NOAA's environmental data and information to support key operational and research challenges. It also describes the steps that will be taken to modernize and improve NOAA's environmental data and information systems in the coming decade.

[original signed]
Gary Locke
Secretary, U.S. Department of Commerce

Preface

Every day practically every sector of the Nation's economy uses the environmental data and information under the National Oceanic and Atmospheric Administration's (NOAA) stewardship. These data encompass weather, climate, geophysics, oceans, and coasts. NOAA ensures timely access to global environmental data, provides information services, and develops science products to facilitate NOAA's environmental stewardship and ecosystem management for the benefit of society.

NOAA is currently in the process of developing a Next Generation Strategic Plan to define NOAA's vision, mission, and goals to serve society for the next 25 years. NOAA's Plan will respond to the priorities of the Obama administration to include data availability to the public and open government. The plan supports the Department of Commerce's priority to strengthen information access and create new data products and services. An example of a key NOAA priority is climate change. Climate change is happening right now, and NOAA is an agency on the front line. Environmental data and information management are key parts to support our role in helping understand and predict changes in Earth's environment.

NOAA data users are more sophisticated and more dependent on us than ever before. In turn, delivering data to those users in a manner that meets their needs for time critical information is even more essential. This report describes our plans to address these challenges through 2009..

[Original signed]
Jane Lubchencho, Ph.D.
Under Secretary of Commerce for
Oceans and Atmosphere
and NOAA Administrator

Contents

Executive Summary

As directed under Public Law 102-567, Section 106, this report describes the state of environmental data and information management within the Department of Commerce's National Oceanic and Atmospheric Administration (NOAA). The Public Law requires NOAA to at least biennially 1) complete an assessment of its environmental data and information systems and 2) submit a modernization and improvement plan based upon that assessment. Effective data and information management by NOAA enables decision makers, scientists, and businesses to make more informed assessments based upon NOAA's ability to collect, preserve, and distribute data and information of the highest scientific quality and completeness.

NOAA assessed the adequacy of its environmental data and information systems using a similar methodology as the 2007 report. The four Mission Goals as defined in the NOAA Strategic Plan were independently assessed based upon twelve end-to-end data management components. Compared to the 2007 assessment, NOAA environmental data and information systems management shows consistency and a general trend of improvement. At the goal level, there were successes in the areas of Data Discovery, Planning, and Complete Metadata. Challenges exist in the area of Appropriate Formats and continue in the Collect and Rescue area.

Similar to previous reports, five broad themes categorize the successes and challenges of NOAA's environmental data and information systems management in 2009. These themes include managing increasing data volumes, filling gaps and extending the environmental record, improving metadata and processing steps, improving data access, and enabling data integration. Since the 2007 report, the National Research Council (NRC) has issued a report entitled "Environmental Data Management at NOAA." NOAA used that report to develop a "What to Archive" procedure that is now being used to meet the increasing data volume challenge by determining what data should be preserved indefinitely versus what data can be considered for a more limited lifetime. NOAA made advancements in extending the environmental record by migrating historical data to Web-accessible media. NOAA is adopting international metadata standards and improving metadata content, which will enable better data discovery capabilities for international activities such as the Global Earth Observing System of Systems (GEOSS) and national activities such as Geospatial One-Stop (GOS) and Data.gov. NOAA improved access and data integration capabilities through the development of Web portals and other systems tailored toward user needs.

Central to NOAA's plan to modernize and improve its environmental data and information systems is the Global Earth Observation - Integrated Data Environment (GEO-IDE) initiative. GEO-IDE proposes the use of a Service-Oriented Architecture (SOA) and a standards-based data and information infrastructure that will provide access to the full range of underlying data system capabilities. In the near term, NOAA is developing a strategy to ensure portal development is integrated. NOAA is also surveying its geospatial data requirements to determine high-priority geospatial needs and is beginning a strategic planning activity that will address model data issues such as data storage and the computational requirements needed for processing and access. NOAA has identified five longer-term projects to modernize its data and information management. These projects are the Comprehensive Large Array data Stewardship System (CLASS), Climate Data Modernization Program (CDMP), Advanced Weather Interactive Processing System (AWIPS), Integrated Ocean Observing System (IOOS), and Climate Data

Records (CDR) Program. An implementation schedule necessary to achieve modernization and improvement objectives is included for these projects.

1. Introduction

This report is the ninth in a series of publications entitled "NOAA Report to Congress on Data and Information Management" (hereafter referred to as the Report to Congress) since Public Law 102-567, Section 106 (hereafter referred to as the Public Law and listed in Appendix A) was enacted in 1992. The Public Law requires NOAA to at least biennially: 1) complete an assessment of its environmental data and information systems; and 2) submit a modernization and improvement plan based upon that assessment. The current Report to Congress follows a similar format as the previous two reports. NOAA has adopted a program-oriented structure consisting of four Mission Goals and a Mission Support Goal around which the agency plans and organizes its work. The four NOAA Mission Goals are Climate, Commerce and Transportation, Ecosystems, and Weather and Water. The 2009 Report to Congress is organized around these same Mission Goals, allowing for a clear-cut analysis and comparison with the 2007 report.

NOAA environmental data and information form the basis for making decisions that have far-reaching economic and political effects at local, regional, and global levels. These data are summarized into data products and used in scientific assessments that are distributed to hundreds of thousands of researchers in government, commerce, industry, science, engineering, and national defense. As an example of the impact of NOAA data and information, the Department of Commerce's Bureau of Economic Analysis estimates that at least one-third of the U.S. Gross Domestic Product is weather and climate sensitive ($4 trillion in 2005 dollars) ranging from finance, insurance, and real estate to services, retail and wholesale trade, and manufacturing. The data and information management activities NOAA provides are critical in enabling decision makers, scientists, and businesses to make more informed assessments and conclusions based on easily accessible and reliable data and information.

NOAA adopted an overarching business strategy for integrating observations and associated data and information management. This strategy is depicted in Figure 1 as a reporting structure for NOAA Environmental Data and Information Management. The NOAA Observing System Council (NOSC) is the focal point for coordinating observational and environmental data management activities across NOAA; for proposing priorities and investment strategies for observation related initiatives; and for identifying programs that might benefit most from integration. The Data Management Committee (DMC) is the NOSC agent responsible for coordinating the development and implementation of environmental data management policy across NOAA. The DMC meets monthly and consists of representatives from all NOAA Goals and NOAA Line Offices. The DMC also has representatives from major NOAA projects that have large data management requirements. An example is IOOS, which aims to simplify the ways in which customers find, obtain, and use information about the oceans.

Recognizing the role of information technology, NOAA plans to share the responsibility for environmental data and information management between the NOSC and the NOAA Chief Information Officer (CIO) Council in fiscal year (FY) 2010. Plans are underway to recast the DMC to the Environmental Data Management Committee, which will have an expanded role in data and information management to support the objectives of the two councils.

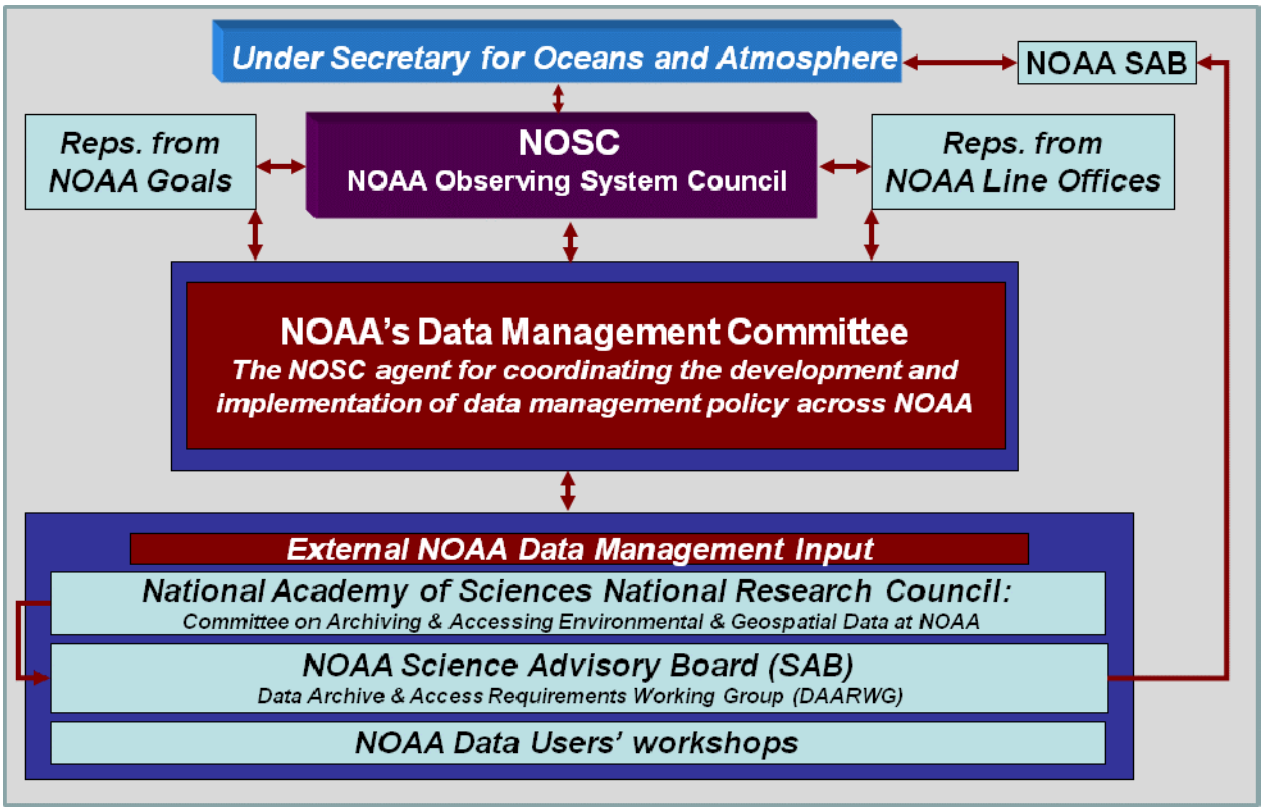


Figure 1: Reporting Structure for NOAA Environmental Data and Information Management

Establishing mechanisms to obtain input from external stakeholders and experts on NOAA environmental data management issues is a critical aspect of the NOAA strategy. The following are examples of this activity:

- A NRC Committee on Archiving and Accessing Environmental and Geospatial Data at NOAA was impaneled pursuant to a request by the DMC and that committee produced a final report in November 2007, entitled "Environmental Data Management at NOAA." The report described nine main principles and associated guidelines for NOAA data management.
- The NOAA Science Advisory Board established the Data Archive and Access Requirements Working Group (DAARWG) in November 2006, in response to a request from the NOSC and DMC. Its charter is to make specific recommendations to NOAA on data archive and access consistent with established principles and guidelines.
- In addition, NOAA held many data user workshops that allowed public input into NOAA environmental data management activities.

2. Assessment of Current Capabilities

2.1 Methodology

NOAA assessed the adequacy of its environmental data and information systems as defined by the criteria in the Public Law (see Appendix A). The current assessment, shown in Figure 2, followed the same methodology as used in the 2005 and 2007 assessments. The four Mission Goals as defined in the 2008 NOAA Strategic Plan were independently assessed based upon twelve survey questions. These questions focused on an end-to-end environmental data management approach that addressed the five data management, archive, and distribution areas identified in the Public Law. These questions were grouped into three broad data management categories: (1) Observation Acquisition and Transmission, which includes the functions of Maintenance and Monitoring and Collect and Rescue, (2) Scientific Data Management, which includes the functions of Calibration and Validation, Appropriate Formats, and Complete Metadata, and (3) Archive and Access, which includes the functions of Long-term Preservation, Data Discovery, and Access and Dissemination. Within those groups, each must consider long-term planning. The final and twelfth function is Contingency Planning, which was also assessed in order to measure the adequacy of producing and delivering NOAA products and services during disruptions of normal business activities. An advantage of continuing the same assessment methodology is that it allows for clearer comparisons when identifying successes and challenges that occurred since the previous *Report to Congress* was issued. Additional details on the assessment methodology are provided in the *2005 Report to Congress*.

As in past assessments, the responses were not normalized across the NOAA Goals and no quantitative value was assigned to "substantial" or "incremental" increases in resources. To aid in comparing the results to the previous assessment, symbols are added in Figure 2 to indicate an improvement (>) or a decline (<) in capabilities by one level since 2007.

A successful NOAA Goal is defined as one that acquires high-quality data; analyzes, evaluates, documents, and preserves these data; and provides timely information and products to the end users of the services. These assessments, in combination with strategic planning at the enterprise level, provide NOAA with an integrated end-to-end approach for Earth environmental data and information management. In addition to the summary chart in Figure 2, more detailed assessments of the NOAA goals are provided in Appendix B where a chart for each NOAA goal is presented along with additional descriptive text. These charts offer an assessment at the individual program level of the goal.

		End-to-End Environmental Data Management Functions											
Doing with current resources		Observation Acquisition & Transmission			Scientific Data Management				Archive & Access				Contin-gency Planning
Need incremental increase													
Requires substantial additional resources		5-Year Plan	Maintain & Monitor	Collect & Rescue	5-Year Plan	Calibrate & Validate	Appropriate Formats	Complete Metadata	5-Year Plan	Long-term Preservation	Data Discovery	Access / Disseminate	
NOAA Mission Goals	Ecosystems								>		>		
	Climate							>			>		
	Weather & Water						<						
	Commerce & Transportation												
Symbols indicate change from last (2007) assessment: (>) = improvement; (<) = decline													

Figure 2: Assessment of 2009 NOAA environmental data management capabilities by Mission Goal

2.2 Successes and Challenges

Compared to the 2007 *Report to Congress*, the current overall assessment of NOAA environmental data and information management shows consistency and a general trend of improvement. As depicted in Figure 2, there were successes in the areas of Data Discovery, Five-Year Planning, and Complete Metadata. Challenges exist in the area of Appropriate Formats and continue for the Ecosystem and Commerce and Transportation Goals in the Collect and Rescue area as well as Climate Goal's ability to Calibrate and Validate and Commerce and Transportation's Long-term Preservation and Access and Dissemination. In order to provide a more quantitative analysis, these successes and challenges are described in more detail in the remainder of section 2 at the overall goal level and also at the individual program or project level.

At the goal level, the Climate Goal reports improvement in Complete Metadata. The job is not complete but significant progress has occurred as additional data sets become more fully described, utilizing the Federal Geographic Data Committee (FGDC) standard. Data Discovery also showed improvement in the Climate Goal as a number of activities occurred in the areas of search capability and data portal functionality, to enhance access to a worldwide set of observations and derived products. The Ecosystem Goal also showed improvement in the area of Data Discovery by developing capabilities that provide improved access to data and information through the cataloging, formatting, archiving, and disseminating of its data streams. The Weather and Water goal indicates a challenge in the area of Appropriate Formats in that work is needed in converting space weather satellite data products to more modern standard formats to support interoperability and accessibility. The Commerce and Transportation Goal showed consistency in its data management capabilities since 2007, with significant improvement in the Surface Weather Program and the Aviation Weather Program's establishment of the Next Generation Air Transportation System (NextGen) Net Enabled Weather Information Database (WIDB). Additional details at the goal program level are available in Appendix B.

The environmental data managed by NOAA advances, or supports, six of the eight National Essential Functions (NEF) identified by the White House, via the Homeland Security Council. The NEFs are the subset of Government functions necessary to lead and sustain the Nation during a catastrophic emergency, and are therefore the highest national priorities in the aftermath of major incidents of all

origins. Thus, contingency planning for NOAA data systems is vital to ensure uninterrupted delivery of NOAA's products and services during disruptions to normal business.

NOAA continues to perform annual reviews of the Certifications and Accreditations (C&A) documents supporting its major systems. The C&A documents include information on disaster recovery and contingency planning. At the Goal level, NOAA's contingency planning capability shows consistency with the 2007 assessment. Much of the environmental data available from the NOAA Data Centers (NDC) are available as online mirrored back-ups. A more detailed assessment of contingency planning for the individual programs within each NOAA Goal is provided in Appendix B.

Five areas again emerge as the prominent themes that categorize the successes and challenges of NOAA environmental data and information management since 2007:

- Managing the increasing volume and diversity of data
- Filling gaps and extending the environmental data records
- Improving descriptions of data, metadata, formats, and processing steps
- Improving access to the long-term archive
- Enabling integration of quality observations and products

In the remainder of section 2, each of these themes is examined in detail by highlighting recent successes and areas that present challenges.

2.2.1 Managing Increasing Data Volume and Diversity

Influx of new data challenge: NOAA archives are continuing to face rapid growth in the total volume of data that requires safe storage and stewardship. The following diverse large array data streams are expected to be produced over the next decade:

- New satellite sensing systems, including the Joint Polar Satellite System (JPSS) and NOAA responsibility under the restructured National Polar-orbiting Operational Environmental Satellite System (NPOESS) program, the NPOESS Preparatory Project (NPP), and the Geostationary Operational Environmental Satellite Series R (GOES-R), are expected to increase the incoming satellite-related data by an order of magnitude over the next decade.
- Weather Surveillance Radar 88 Doppler - Next Generation Weather Radar (NEXRAD) data volume will increase rapidly as higher resolution and improved technology are implemented across the National Weather Service (NWS) network over the next decade. Dual polarization upgrade to the Nation's Doppler radar network will be deployed in 2010 through 2012, resulting in annual data volumes increasing by a factor of 10. Phased array radar technology is being considered for implementation with annual data receipt expected to approach about five petabyes (PB) or 50 times what is received today.
- Large-array data streams from climate, weather, and ocean models are expanding rapidly. Current estimates predict substantial growth in the amount of model data, suggesting that models will soon rival satellites in the amount of data produced. For example, the estimate for weather and climate model output is on the order of PBs of data that will be submitted per year to the NDCs by 2010.
- New multi-beam and side-scan sonar systems are continuing to produce large volumes of data from NOAA's hydrographic surveys. Current individual hydrographic surveys are often

as large as a terabyte (TB), largely due to increased data density and resolution of modern surveying technology. The widespread use of multi-beam sonars in the NOAA fleet led to impressive full seafloor coverage, but created a challenge for those managing the data.

- High resolution video data include images of sea floor transects used for quantifying biological abundance, diversity, and habitat changes; images from stationary cameras that monitor changes in weather, shoreline, or other characteristics of weather, surf, or shore conditions; and images of marine creatures (e.g., whale flukes) used to track and monitor individual or groups of organisms. Video data can expand an archive by multiple TBs per day. Another challenge concerns the difficulties of generating useful metadata and search capabilities for users. Video data is a relatively new field in terms of archiving data, and would involve a rather intense effort to document what the video shows. NOAA plans to assess the feasibility of archiving these data using the "What to Archive" procedure.

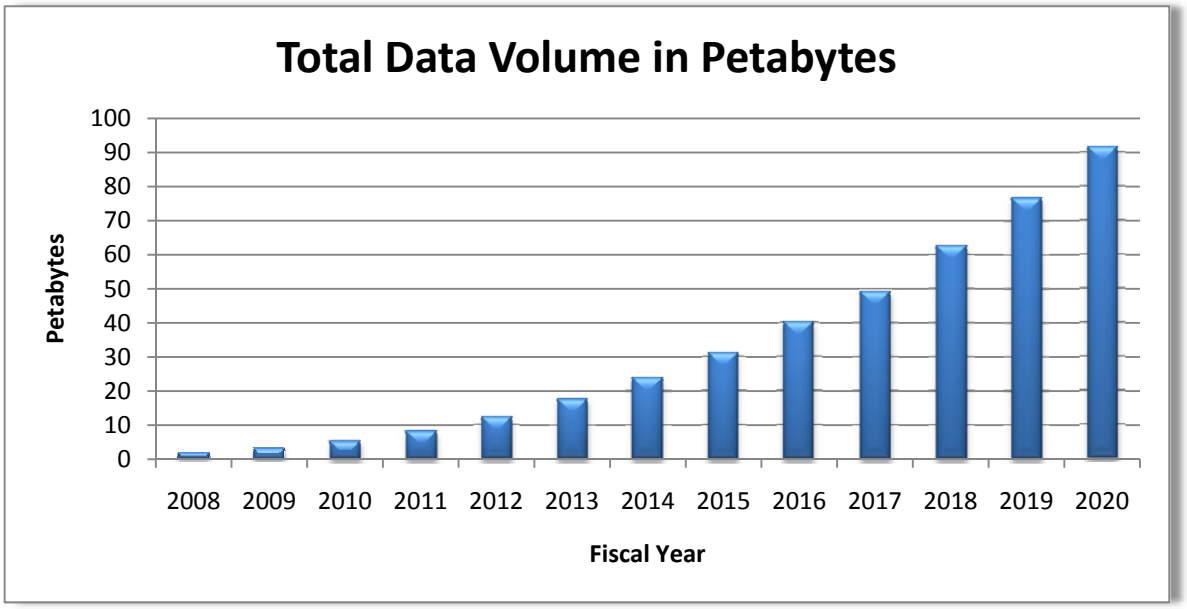


Figure 3: Total cumulative data volume projected for storage by the NOAA Data Centers (single copy, backup not included)

Figure 3 illustrates the 2009 projection for the growth in data volume that will be stored at the NDCs. The graph depicts the cumulative total single copy data volume expected through 2020. A back-up copy is required to provide safe storage and doubles this data volume. Data volume projections are very dynamic due to system development delays. As an example, a new satellite may have a delayed launch resulting in changes to when data actually enter the archive. In addition, reassessments are part of the planning process in regard to what data is required in order to keep pace with the advancements that occur in science.

As described below, a procedure to determine what data should be archived is critical in order to keep pace with this enormous growth in data volume. Efforts are underway to build the infrastructure necessary to provide proper and complete stewardship of the data archived at NOAA.

The NDCs are utilizing CLASS to ensure the long-term, secure storage of data, information, and metadata, particularly for large data sets. In early 2008, NOAA separated CLASS operations and development activities in order to better define the roles and responsibilities of the data centers and the CLASS project manager. The NDCs manage CLASS operations through the CLASS Operations and Planning Board (COPB), which consists of the directors of the three NDCs. The COPB also provides requirements and guidance to the CLASS project manager, who is housed in the NOAA National Environmental Satellite, Data, and Information Service (NESDIS) Office of Systems Development. The CLASS project manager will continue to focus on the information technology required to ingest, store, and maintain the submitted data sets. The NDCs will be responsible for information preservation and end-to-end stewardship of the archived data, as well as maintaining access interfaces that are responsive to users. Additional information on CLASS and its future plans are included in the Comprehensive Improvement Plan section of this report. Figure 3 reflects the data volume projections given current schedules.

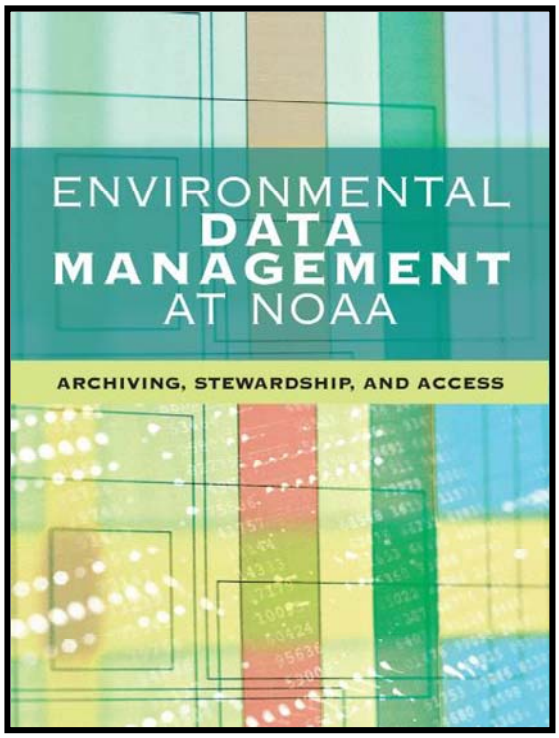


Figure 4: Cover of the 2007 National Research Council Report

NRC Report on Data Management Published: At the request of the DMC, NRC impaneled a committee in 2006 to provide recommendations on archiving and accessing the broad range of environmental and geospatial data collected by NOAA and its partners. With limited resources and enormous growth in data volumes, NOAA sought input on how to identify the data and derived products that should be preserved in perpetuity and made readily accessible versus those that could be considered for more limited access and storage lifetimes. Figure 4 shows the cover of the final NRC report produced in November 2007, entitled "Environmental Data Management at NOAA: Archiving, Stewardship, and Access."

The NRC report provided NOAA with recommendations on what types of data should be archived indefinitely (for at least 75 years); what types of data could be stored for shorter durations under budgetary constraints; and how best to provide access to different variables, data sets, and derived products. A total of nine broad principles and associated guidelines on data management were developed. The principles and guidelines provided NOAA with a comprehensive long-term approach for managing data and data products from existing, legacy, and future observing systems as well as environmental models. NOAA used the recommendations presented in the NRC report as a basis for developing a new "What to Archive" Procedure.

NOAA "What to Archive" Procedure Developed: A major achievement since the last Report to Congress was the development of a universal procedure now used across NOAA to identify, appraise, and decide what scientific records are preserved in a NOAA archive. To develop the new "What to Archive" procedure, the DMC assigned a team in December 2007 comprised of data management

experts across all NOAA Line Offices and Goals. The development involved an extensive literature review of past work in this area, coordinating reviewer comments from inside and outside NOAA, and presentations at various forums. The procedure was discussed at a June 2008 joint workshop of NOAA data managers and DAARWG. The DAARWG commended NOAA on the procedure and recommended the procedure be produced as a published document. A document entitled "NOAA Procedure for Scientific Records Appraisal and Archive Approval" was approved by NOSC in December 2008. As shown in Figure 5, there are four broad steps that apply to accepting or rejecting new scientific records into a NOAA archive and to retaining or disposing of existing records held in a NOAA archive. Pursuant to another DAARWG recommendation, the procedure was incorporated into a December 2008 revision of NOAA Administrative Order (NAO) 212-15, Management of Environmental and Geospatial Data and Information. In May 2009, the NOAA procedure was accepted by the National Archives and Records Administration (NARA) for inclusion as a best practice and added to the NARA Toolkit for Managing Electronic Record web portal. The "What to Archive" procedure is available at: http://www.nosc.noaa.gov/docs/products.html. NAO 212-15 is available at: http://www.corporateservices.noaa.gov/~ames/NAOs/Chap_212/naos_212_15.html.

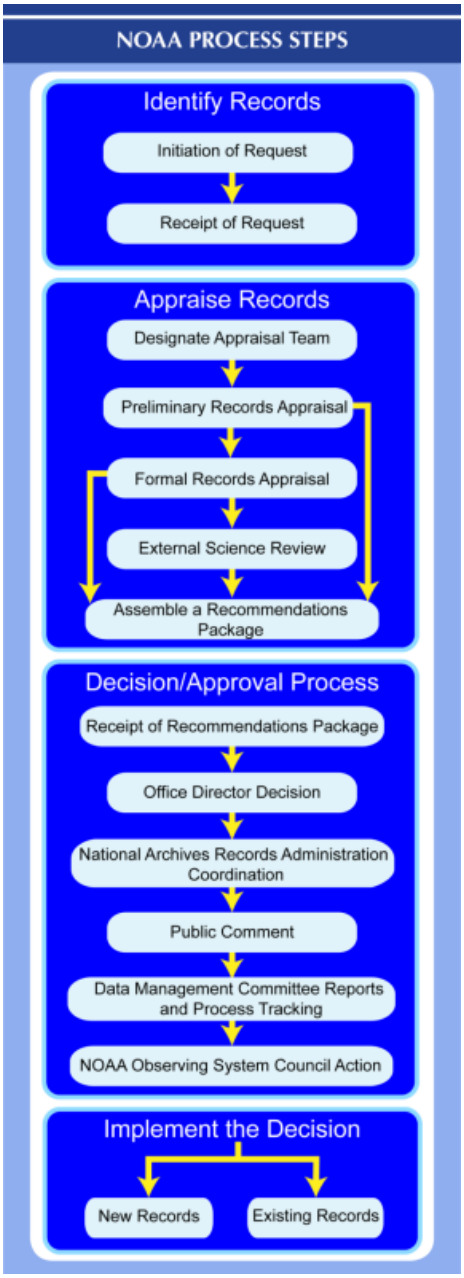


Figure 5: Steps in the "What to Archive" process

Computing Capability and Bandwidth Challenge: In addition to the investment in the capability to provide storage for the multi-PBs of data at the NDCs, a complementary investment in computing capability is needed in order to use and properly steward these data. Increases in computing capability and associated bandwidth needed to access the data among systems will enable robust and efficient data processing and reprocessing, analysis, and assessment, essential for studying societal issues involving integration of multi-disciplinary data. As an example, super-computers with petaflop computing capacity will be needed to assimilate archived data and provide state-of-the-science climate model simulations at the scales useful for local and regional climate change adaptation.

2.2.2 Filling Gaps and Extending the Environmental Records

The Climate Database Modernization Program: CDMP supports the NOAA mission to collect, integrate, assimilate, and effectively manage Earth observations on a global scale ranging from atmospheric, weather, and climate observations to oceanic, coastal, and marine life observations. Many

of these holdings, which are part of the U.S. National Archives, were originally recorded on paper, film, and other fragile media, and stored at various NOAA Centers. Prior to CDMP, not only were these valuable data sources mostly unavailable to the scientific community, but storage technology for the archive was becoming obsolete. Without proper preservation of the media, the information they contained was in danger of being lost forever. CDMP greatly improved the preservation and access to NOAA's holdings by migrating many of these resources to new digital media.

Partnering with four private-sector contractors, CDMP has placed over 53 million weather and environmental images online that are now available to researchers around the world. The amount of data online grew from 1.75 TBs in 2001 to over 11 TBs in 2009. Hourly weather records captured and entered through CDMP continue to be integrated into NOAA's digital data base holdings, extending the period of record for many stations back into the 1890s. Additional daily data records that are captured and entered through the CDMP "Forts" project will soon extend this data period back to the 18th century for several stations. Major progress continues in making these data available through a number of NOAA Web sites, as indicated by the task distribution shown in Figure 6.

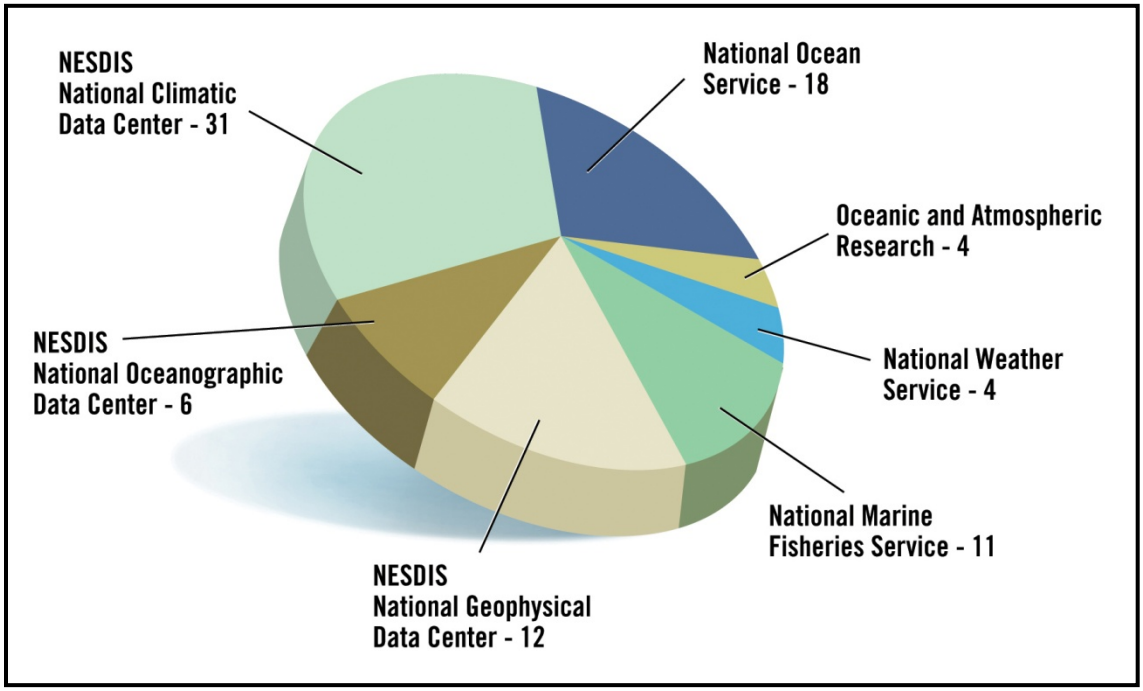


Figure 6: 2009 Climate Data Modernization Tasks Across NOAA

CDMP also supports preservation of other important NOAA environmental data, ranging from ocean cores below the seabed floor to the top of the ionosphere. For example, the Tsunami Program had more than 25,000 photos, slides, negatives, microfiche, and hardcopy images electronically scanned. The Tsunami Warning Centers used these historical data as tools in decision making for hazard assessment, and in validating tsunami propagation and inundation models. The increase in the accessibility of quality historical data is helping researchers worldwide to improve real-time monitoring and forecasting of environmental, solar, and geophysical events.

World Ocean Database: An example of extending and filling gaps in the environmental records is the World Ocean Database, which is a global, comprehensive, quality-controlled compilation of oceanographic vertical profile data. Data are submitted to NOAA from ocean researchers and programs from around the world. These data are processed and then used in developing ocean climatologies and in studies of inter-decadal variability, which is very important for studying global climate change. Examples of the oceanographic variables included are temperature, salinity, nutrients, oxygen, and pH data as a function of depth.

The amount of historical and modern profile data increased dramatically in the 2009 version of the World Ocean Database. Approximately 200,000 Ocean Station Data casts and approximately 200,000 Conductivity-Temperature Depth casts were processed in the database. These data were used to document that the world ocean has warmed during the past 50 years. The data are frequently used as internal and external boundary conditions in ocean general circulation models and in assimilation studies.

2.2.3 *Improving Descriptions of Data, Metadata, Formats, and Processing Steps*

Metadata Standards: The metadata world is changing fast with convergence across NOAA in many areas and rapid evolution towards new, very powerful international standards. Migration towards the International Organization for Standardization (ISO) 19115 geographic information metadata standard in NOAA is being led by the NDCs, National Ocean Service (NOS) Hydrographic Survey, the GOES-R Program, the CDR project, and the NOAA IOOS Program. That migration will accelerate as these programs come up to speed and share their experiences and examples with NOAA data providers. Another outstanding area of convergence is reflected in NOAA content available through GOS, which includes more than 11,000 NOS Hydrographic Surveys, almost 1,000 data sets from NDCs, and more than 7,000 coastal data sets. These metadata records will be included in GEOSS registries and portals as those systems mature over the next few years, as well as in the new Data.gov initiative. This will increase public access to high value, machine readable data sets generated by the Federal Government's Executive Branch.

Habitat Restoration and Coral Reef Metadata: The capabilities of the National Estuaries Restoration Inventory (NERI) and the NOAA Restoration Center Data Base (RCDB) were successfully improved in the areas of metadata and data discovery. The RCDB and NERI are inventories of habitat restoration projects that contain site location information and associated attribute data (abstracts, project status, photos, etc.). The geographic data and associated attributes provided by these two data products are now fully described in FGDC-compliant metadata records and were published to multiple online data catalogs, including GOS. Similarly, the Coral Reef Conservation Program remains successful in providing improved metadata via the Coral Reef Information System, which is a partnership that leverages efforts of the National Oceanographic Data Center (NODC) to assist with data description and data discovery.

2.2.4 *Improving Access to the Long-Term Archive*

Plans to Phase Out User Fees: The NDCs, National Climatic Data Center, National Geophysical Data Center, and NODC are authorized under 15 U.S.C. § 1534 to assess fees for access to environmental

data, information, and products. Biennial cost reviews are performed to set prices for recovering the cost of dissemination, as mandated by OMB Circular A-130. The increase in the online accessibility of data resulted in a 20 percent decrease in access fees for the first eight months of FY 2009, compared to the same period in FY 2008. For the future, the NDCs plan to align their current customer on-line access and retrieval to become compatible with the "Open Government Initiative" and compliant with the GOS and Data.gov requirements and operating procedures.

NOAA Climate Services Portal: During the 20th century, private industry and regulatory bodies made decisions for future investments (e.g., airports, seaports, power plants) under the assumption of a stationary climate (i.e., that past climate is a good indicator of future climate). In the 21st century, this business approach is no longer tenable as a large body of scientific evidence (i.e., Intergovernmental Panel on Climate Change, Climate Change Science Program, academic research) shows that climate conditions are exhibiting non-stationary characteristics. As a focal point for improving the communication of climate information, NOAA has been developing a Web-based Climate Services Portal. The climate portal will enhance and ease access to a wide variety of climate information through a Web interface that enables individuals to search by sector, region, problem/issue, climate variables, or climate program element (e.g., monitoring, predictions, modeling, research, assessment, decision support tools). The portal will be developed over a multi-year timescale with developmental input from a vast array of NOAA data users to optimize its data flow architecture. The phase 1 prototype was released to the public in calendar year 2009.

National Integrated Drought Information System: Drought events have far-reaching impacts on many aspects of our daily lives, from water management to health to energy consumption and conservation. The National Integrated Drought Information System (NIDIS) is a dynamic and accessible drought-risk information system that was created by NOAA, other Federal and state Agencies, partners, and countries in response to extended drought conditions, especially in the Western United States, over the past decade. In 2007 NIDIS unveiled the U.S. Drought Portal (USDP), which allowed the public and civic managers to monitor U.S. drought conditions, get forecasts, assess the impacts of drought on their communities, and learn about possible mitigation measures. This Web site, located at http://drought.gov, is useful internationally as nations work to coordinate drought preparedness, response, mitigation, and recovery activities. This fits in well with drought related bilateral activities with U.S. partners in Canada and Mexico. The 2007 Group on Earth Observations' Plenary meeting featured NIDIS as a major contribution to GEOSS. In 2008 and 2009 NOAA and its partners, including the U.S. Department of Agriculture, began to institute geographic information system (GIS) mapping capabilities into the USDP.

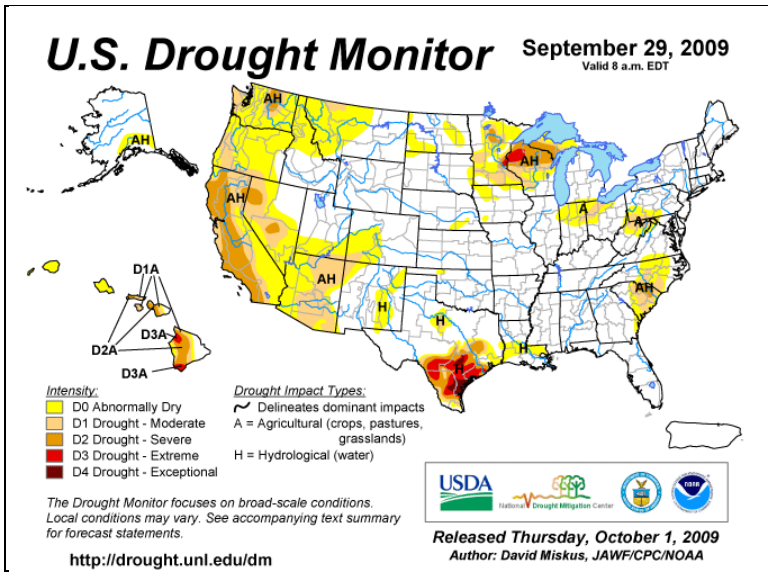


Figure 7 Weekly Drought Monitor product, available at Drought.gov portal

Simple NOAA Archive Access Portal: Simple NOAA Archive Access Portal (SNAAP) is a transitional interface and Web portal that facilitates user access to NOAA environmental data sets hosted in diverse systems, including those in CLASS. SNAAP will provide a capability to integrate diverse data systems via a standards-based interface while maintaining a clear functional separation between data archive and access tasks. SNAAP will provide improved Web-services based archive and access consistent with GEOSS objectives. When complete, this system will allow users to uniformly access a vast array of NOAA data. This project will make it easier to get an integrated look at NOAA's distributed data holdings. The potential challenges are delivering an operational system in a timeline to support the Climate Services Portal, making sure the system is compatible with the GEOSS gateway, and making sure the functionality captured in the prototype meets the needs of users.

External Space Weather Database Store: In 2007 and 2008 NOAA held Space Weather Workshops with private industry where they determined that real-time, unfettered access to space weather data could better serve external users such as the emergency management community, other government partners, and commercial service providers. As a result, NOAA developed External Space Weather Database Store (E-SWDS) to facilitate the transfer of near real-time space weather data from NOAA's Space Weather Prediction Center (SWPC) to private industry. This greatly aids in the SWPC mission of providing real-time monitoring and forecasting of solar and geophysical events, which can impact satellites, power grids, communications, navigation, and many other technological systems.
Using E-SWDS, information is exchanged from NOAA's SWPC automated systems and personnel to external parties via an Open Data Base Connectivity interface or similar technology, such as Java Database Connectivity. E-SWDS allows for the establishment of connections based on individual user IDs and passwords for the access of database information via predefined views or stored procedures. This service greatly improves the understanding of the evolving space weather conditions and aids in the private industry decision-making process by providing improved access to NOAA's data.

NOAA's Enhanced Online Positioning User Service: NOAA's National Geodetic Survey (NGS) released a new version of its popular Online Positioning User Service (OPUS) in February 2008. OPUS-Rapid Static (OPUS-RS), designed in partnership with the Ohio State University, enables users to receive an OPUS Solution by submitting as little as 15 minutes worth of dual frequency GPS data to NGS via the Web, where the data are processed to determine a position using NGS computers and software. OPUS-RS saves time and money as compared to standard OPUS, which requires at least two hours worth of GPS data; however, OPUS-RS results are less accurate than standard OPUS. NGS expects that OPUS-RS usage will exceed OPUS usage by a factor of 10.

NOAA's efforts with Google Earth, Ocean and Map products: In 2008 NOAA renewed its enterprise licensing contract with Google, allowing for unlimited deployment of the Google Earth Enterprise Client to all NOAA desktops. User-friendly tools and applications can be built in the Google environment, helping NOAA offer its data, products, and services through an increasingly common interface. An example is NOAA's Coastal Water Temperature Guide, which was developed in response to numerous daily requests for water temperatures of the U.S. coastal areas including the Great Lakes. NOAA created this timely and user-friendly tool, which has now become the primary gateway for providing useful information for the public in planning beach activities, such as swimming and fishing.

NOAA will continue to provide geospatial data to Google for Google Earth, Google Ocean, and Google Maps products. This is a developing relationship that will improve access to NOAA geospatial data and improve Google products that are based on high-quality data. In 2008 a NOAA staff member was on a four-month detail to Google headquarters to work with Google's Ocean Team. This facilitated the partnership of incorporating NOAA data into the Google Ocean product that was released in 2009. Discussions on how to continue this productive relationship are ongoing. Another example is depicted in Figure 8. NOAA's Community Model Interface for Tsunami provides tools for the construction of tsunami inundation maps under different scenarios and can also be used for real-time tsunami forecast applications.

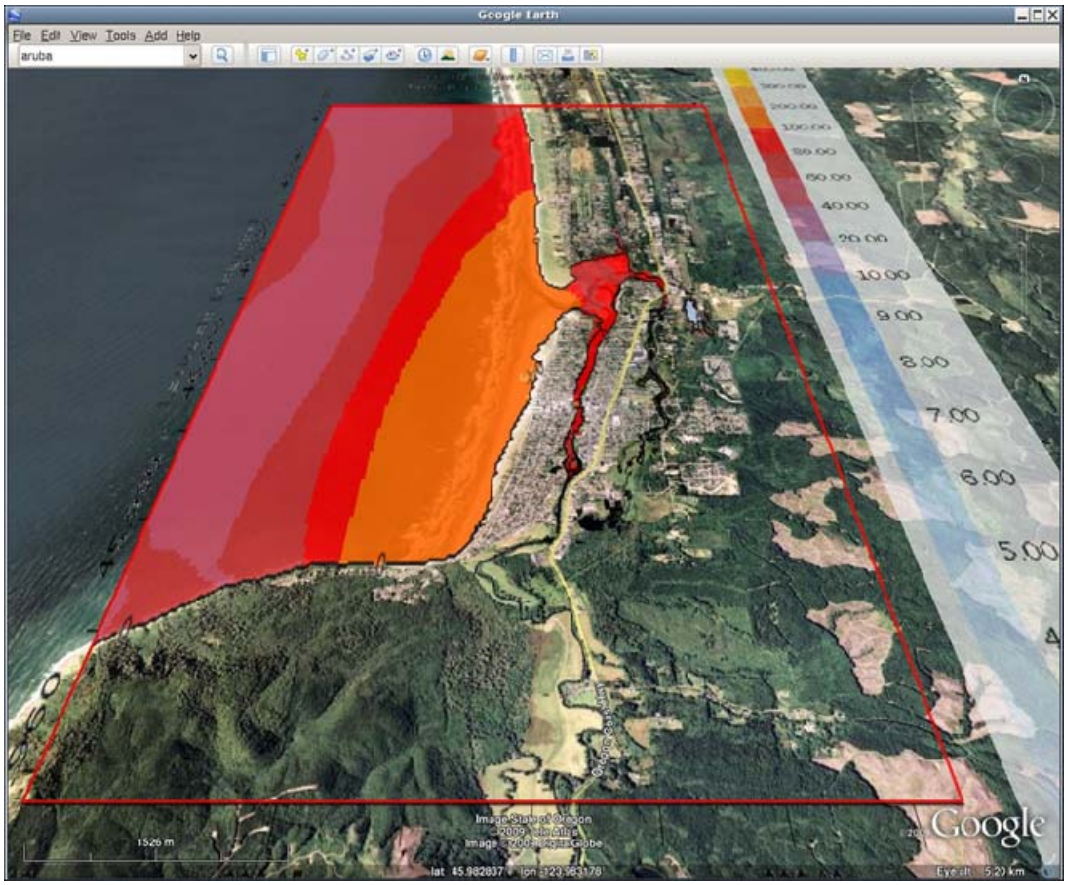


Figure 8: Example of NOAA Tsunami geospatial data access using Google Earth applications

2.2.5 Enabling Integration of Quality Observations and Products

The U.S. Integrated Ocean Observing System: IOOS leverages the Nation's existing capacity in ocean observations and forecasts by linking those outputs to end users via standardized Data Management and Communications (DMAC) services. Programmatically, IOOS is the ocean portion of the broader national Integrated Earth Observing System (IEOS) and the U.S. contribution to related international efforts—the Global Ocean Observing System (GOOS) and GEOSS. NOAA is the lead Agency for IOOS and chairs the Interagency Working Group on Ocean Observations. IOOS also provides the governance structure for participation and standardization activities.

Technically, IOOS aims to harmonize or simplify the ways in which customers find, obtain, and use information about the ocean. Data providers have traditionally offered data in disparate ways (different formats, different methods) that make it difficult for users to gather a unified view of the oceans, coastal regions, and the Great Lakes from several sources. IOOS data providers are augmenting or replacing existing access services and formats with standardized approaches. The standards to be used are open, non-proprietary, vendor-neutral standards adopted by GEOSS and developed by bodies such as the Open Geospatial Consortium and the ISO. Providers will register their new services with an IOOS catalog to enable users to find available data. Users include decision-makers using highly specialized analysis

tools, or simulation models that produce forecasts of weather or ocean conditions, or the general public via common Web and desktop software. The complete IOOS DMAC will include additional services such as data visualization and format conversion.

Regional Ecosystem Data Management: A data-integration framework was developed to support NOAA's Ecosystem Based Management efforts, including Integrated Ecosystem Assessments. This framework consists of a collection of mutually supportive components providing data discovery, access, fusion, and long-term preservation, known collectively as Regional Ecosystem Data Management (REDM). The overall approach is to link disparate databases and other web information resources by building and maintaining a comprehensive metadata catalog and associated search application, and employing standard Web services and access protocols to deliver data to clients (either human or machine) following GEO-IDE guidelines. The cataloguing and description of data sets is being pursued on a regional basis, tracking NOAA's regional approaches to ecosystem management. In addition, REDM metadata and data services are designed to support interoperability so that this data-integration framework can be leveraged by other targeted data-integration efforts.

NOAA Tsunami Program Integration: NOAA's Tsunami Mission is to provide reliable tsunami forecasts and warnings, and to promote community resiliency. In 2007, to accomplish this more effectively, the Tsunami Program assessed current data stewardship, identified gaps, and made recommendations to address the gaps. Recent successes in addressing the identified gaps include: improved transmission of essential seismic data to the Tsunami Warning Centers utilizing NOAANet to deliver data; the stewarding of formerly unarchived data streams such as coastal Tide Gauge data from the National Ocean Service's (NOS) Center for Operational Oceanographic Products and Services and from the Tsunami Warning Centers; rescuing non-digital tsunami event data; improving the completeness and accuracy of the Global Historical Tsunami data bases; and implementing off-line access to these data bases for areas without robust Internet capability. These new collaborative efforts improved NOAA's ability to produce coastal Digital Elevation Models (DEM) for a host of coastal management and emergency response applications. NOAA created a geospatial online catalog or "DEM Discovery Portal" that links to Web-published DEM's using an Arc Internet Map Server (ArcIMS) map interface to locate, preview, and download the DEMs. The Portal also supports the National Tsunami Hazard Mitigation Program effort to prioritize national development of inundation maps for emergency preparedness.

Pacific Climate Information System: Pacific Climate Information System (PaCIS) provides: a programmatic framework to integrate ongoing and future climate observations, operational forecasting services, and climate projection; research; assessment; data management; and outreach and education to address the needs of Pacific Island communities. The regional vision for PaCIS is "resilient and sustainable communities using climate information to manage risks and support practical decision-making in the context of climate variability and change." One of the highest priorities for PaCIS was the completion in 2009 of a prototype Web-based portal that will facilitate discovery of and access to climate data, products, and services. A demonstration of PaCIS' ability to bring multiple assets together towards addressing critical environmental issues such as coastal inundation is the Pacific Region Integrated Climatology Information Products (PRICIP) project. PRICIP (see http://www.pricip.org/) is an effort to improve understanding of patterns and trends of storminess (e.g., strong winds, heavy rains and high seas) in the Pacific and develop a suite of integrated information products that can be used by

emergency managers, coastal resource managers, mitigation planners, government agencies, and decision-makers in key sectors including water and natural resource management, agriculture and fisheries, transportation, recreation and tourism, and communications. Core support for PaCIS is provided by NOAA's Integrated Data and Environmental Applications Center, which was established in 2005 in order to advance NOAA's mission objectives in the Pacific region and help meet critical needs for climate, ocean, and ecosystem information to protect lives and property, support economic development, and enhance the resilience of Pacific Island and other coastal communities in the face of changing environmental conditions.

3. Comprehensive Improvement Plan

3.1 NOAA's Data Management Plan

NOAA continues to move forward with the implementation of GEO-IDE. The GEO-IDE vision is the effective and efficient integration of NOAA's many existing environmental data management systems with future development efforts to enhance the ability of NOAA's stakeholders to discover and access NOAA's environmental information. The GEO-IDE Concept of Operations proposes the use of a service-orienting architecture (SOA) and a standards-based data and information infrastructure that will provide access to the full range of underlying data system capabilities. GEO-IDE implementation focused on coordinating and communicating various guidelines and standards across major NOAA projects.

GEO-IDE is being guided by a small project office and the Data Management Integration Team (DMIT), which comprises NOAA data management specialists with representation from all NOAA Line Offices, Goal Teams, and the CIO office. DMIT regularly reports to the DMC. DMIT will: define and implement the processes required to establish NOAA-wide data management standards and guidelines; evaluate new and emerging data management technologies; and sponsor or oversee the coordination, evaluation, and development of infrastructure components.

The objectives and approach of GEO-IDE are fully consistent with and support the U.S. Group on Earth Observations (USGEO) Strategic Plan for the IEOS, which is the U.S. contribution to GEOSS. All three complementary initiatives are attempting to remove barriers that prevent or inhibit the most productive use of Earth observation data applied to a wide range of societal benefit areas. Where GEO-IDE is focused on improving access to NOAA's data and information systems and services, USGEO and GEOSS are addressing the same challenges faced across Federal agencies and the nations of the world. In each case, the initiatives are leveraging the technical innovations and best practices of their members to develop common, standards-based components and an interoperable set of data and information services.

3.2 Implementing the Plan

The following are several efforts that highlight NOAA's work to implement its data management plan responsibilities using the GEO-IDE concept.

Improved NOAA Climate Services: In FY 2009, NOAA began planning the integration of its portfolio of climate activities. Because the climate challenge is real, the Nation needs targeted climate services at scales from local to global, to help users better understand, adapt to, and mitigate climate change. From a data management perspective this will involve strengthening observations, standards, and data stewardship, which will result in more effective data management and delivery systems. NOAA's current climate services generate significant social, economic, and environmental benefits to the Nation. As NOAA further defines its role in a National Climate Service, the process will require an unprecedented level of coordination among Federal Agencies, as well as with NOAA's nongovernmental partners to accomplish the goal of providing high quality climate information and services that are user-friendly, responsive, and relevant.

Next Generation Air Transportation System Net Enabled Weather Information Database: Approximately 70% of today's aviation delays are weather related, with an annual cost of some $29 billion. Studies indicate that as much as two thirds of these delays can be prevented with more accurate and timely weather information effectively integrated into air traffic management operations. The Federal Aviation Administration is leading the multi-agency NextGen initiative to modernize the U.S. Air Transportation System in order to increase capacity and reliability, improve safety and security, and minimize the environmental impact of aviation. The NextGen weather concept calls for a consistent, common weather picture— a "network-enabled" virtual repository of weather information to be available to the aviation community and integrated into aviation decision-making. NOAA is the lead agency for building and populating this virtual repository, which is referred to as WIDB or the Four Dimensional (4-D) Weather Data Cube. The NextGen Weather capability will be rolled out incrementally between 2013 and 2022. Concept exploration and prototype development activities began in 2008. The effort will address such data management challenges as establishing metadata standards, making weather data discoverable for real time decision support, implementing the infrastructure for handling the large volume of data that must be accessed to support aviation operations, and enabling exploitation of weather data by GIS users.

Assessing NOAA's Geospatial Requirements: In 2009, NOAA's GIS Committee surveyed NOAA's geospatial professionals to determine the agency's high priority geospatial needs. The survey covered various aspects of geospatial use within NOAA and assessed the ability of users, analysts, managers, data providers, and software developers to reach their work goals with existing geospatial data, tools, and information. The survey focused on data, existing capacity, information technology, training, and outreach. It will provide the GIS Committee with an understanding of strengths and weaknesses in each of these areas within the geospatial community. With close to 50% of NOAA offices operating as authoritative geospatial data sources, it is essential to understand the needs of NOAA's geospatial community in order for them to better serve the public. Results are currently being analyzed and distributed to the geospatial community, and a work plan based on these results is being finalized.

NOAA Corporate Portal Strategy: NOAA has a number of agency-wide portals serving a variety of customers and is developing several new GIS-based portals for enhanced data and information discovery and delivery. These new Web portals are based on either a theme, such as the NOAA Climate Services Portal or a region, such as the Southeast and Caribbean Data Explorer, and they provide customers with easy access to desired data. To guide the evolution of these portals and the development of future portals, NOAA is in the process of defining a corporate portal strategy to:

- Enable a consistent user experience and integrated access to NOAA's data, regardless of which NOAA portal serves as the user's entry point;
- Ensure coordination and maximum reuse of portal infrastructure components across NOAA;
- Promote consistent usage of industry best practices for portal development and operations

Development of the corporate portal strategy will entail a facilitated workshop planned for FY 2010 to coordinate requirements across existing portal initiatives and validation of the strategy from an independent review group, such as DAARWG.

Data Management for the NOAA Modeling Enterprise: As described above, environmental modeling poses a data management challenge for NOAA in the areas of data storage and the computational requirements needed for processing and access. However, recent planning activity addressed model-related concerns. NOAA environmental modeling activities reach across most of the organization including such diverse areas as weather and climate, oceans, space weather, hydrology, and ecosystems, and including both research and operational environments. A formal effort to organize the NOAA environmental modeling enterprise is underway, led by the Environmental Modeling Program with the assistance of NOAA's Office of Program Planning and Integration. The resulting 10-year strategic plan will include an inventory of environmental modeling efforts, which will represent a very significant first step in accurately scoping the data management problem. Moreover, the strategic plan will explicitly address modeling infrastructure issues, to include data management, as well as the resources required for infrastructure maintenance**.**

Marine Spatial Planning: President Obama proclaimed June 2009 as National Oceans Month and called on Americans to learn more about the key role the oceans play in our Nation's economy, transportation, food, recreation, and energy resources. This declaration was followed by the establishment of an Interagency Ocean Policy Task Force tasked with developing a national policy for oceans, coasts, and the Great Lakes, and with providing a framework for coastal and marine spatial planning. Marine spatial planning is a comprehensive, ecosystem-based process through which compatible human uses are objectively and transparently allocated to appropriate ocean areas to sustain critical ecological, economic, and cultural services for future generations. From a data management viewpoint, marine spatial planning will require coordinated data management for coastal and ocean data along with strong partnerships with coastal communities in order to meet the needs of our Nation. The planning for this effort began in late FY 2009. NOAA expects to participate and support its ocean stewardship mission through the implementation of a data integration framework in order to provide access to the Nation's coastal and ocean data resources.

Fisheries Enterprise Data Management Planning: The National Marine Fisheries Service (NMFS) is moving forward with an enterprise data management strategy to address information management

challenges over the next decade. The Enterprise Data Management (EDM) strategy includes a Governance Structure consisting of:

- an Information Architect
- Information Management Coordinators in each Regional Office and Science Center
- the Fisheries Information Management Advisory Committee
- policies and procedures addressing data stewardship, metadata collection and maintenance, security, confidentiality, and data quality.

An Authoritative Data Catalog is being planned that will include an inventory and appropriate metadata for all enterprise NMFS information assets including databases; photograph collections; administrative records; and other digital, analog, and hard copy information assets. This will provide data discovery of high quality, reliable, and timely data that will be continually assessed for sufficiency and improvement. One or more Web portals will provide easy access to all NMFS enterprise data based on commercial transaction portal models to enable NMFS internal and external constituents to effectively discover, access, integrate, and use NMFS information resources to answer key current and future questions that may or may not be related to the original purpose for collecting the data.

Integrated Ocean and Coastal Mapping: NOAA will continue promoting the development and implementation of an Integrated Ocean and Coastal Mapping (IOCM) capacity. IOCM is defined as the practice of acquiring, managing, integrating, and disseminating ocean and coastal geospatial mapping data. An ocean and coastal mapping training center and an initial operating capability for an IOCM data processing center will be established in FY 2011 at the NOAA-University of New Hampshire Joint Hydrographic Center to support the tenets of IOCM. The training center will train and certify personnel from across NOAA and other Federal mapping agencies in ocean (seafloor and water column) geospatial data acquisition techniques; standards and specifications; data processing and management protocols; and in the use of acoustic and Light Detection and Ranging data acquisition and technologies. The IOCM data processing center will be staffed to process and manipulate seafloor mapping data, which were acquired from a variety of sources using differing standards and in a variety of formats.

The Ocean and Coastal Mapping Inventory is an inter-agency effort to improve the efficiency of the Nation's ocean and coastal mapping activities. The inventory, which is being built within the GOS Web Portal, will offer a clearinghouse for ocean and coastal geospatial data and interpretative information and a registry of planned, current, and completed mapping activities. To date, metadata for more than 12,000 NOAA hydrographic surveys was published in GOS by the NDCs.

3.3 NOAA's Data Management Responsibilities with other Agencies

NOAA continues to be a premier Earth observation data agency, working with both national agencies and international organizations such as the World Meteorological Organization, International Hydrographic Organization, and World Data Center system. The list of these activities is long, many of which were described in previous Reports to Congress. The following are four highlighted activities for this current Report to Congress.

U.S. Group on Earth Observations: The USGEO is a unifying context for defining data management responsibilities across NOAA and other Federal agencies. USGEO comprises representatives from 15

Federal agencies and three White House offices. NOAA served in the role of co-chair of the Architecture and Data Management working group of USGEO, which in December 2008 produced and published "Exchanging Data for Societal Benefit: An IEOS Web Services Architecture." This document describes an Internet-based SOA as an ideal approach for developing interfaces to both new and legacy data and information systems to promote interoperability and allow users seamless access to USGEO data and services from multiple sources.

NOAA is also a major participant in the corollary international effort to develop GEOSS. The U.S. contribution to GEOSS is IEOS. GEOSS and IEOS will facilitate the sharing and applied usage of global, regional, and local data from satellites, ocean buoys, weather stations, and other surface and airborne Earth-observing instruments. The end result will be access to an unprecedented amount of environmental information integrated into new data products benefiting societies and economies worldwide. GEO-IDE provides NOAA with a mechanism to both contribute to and benefit from the efforts underway at the national USGEO level and international GEOSS level.

NOAA-NASA Collaboration: Due to the Nunn-McCurdy Certification in 2006, several National Polar-orbiting Operational Environmental Satellite System climate and space weather sensors were demanifested to reduce overall program cost. Responding to the jeopardy in which the overall climate program was placed by the potential gap in several key satellite data records, NOAA and NASA, with advice from the NRC, developed a cost-effective strategy to mitigate or recover the lost climate observing capabilities. As a result, several climate sensors were remanifested onto the JPSS program or other satellites (e.g., Jason-3 for sea level height) and new partnerships with international agencies were developed to close the climate data gap. As part of this effort, NOAA and NASA developed a collaborative framework to transition mature research satellite products at NASA into operational CDRs at NOAA. Coming on the heels of NASA's major Earth Observing System (EOS) investment, the CDR Program allows NOAA to capture the revolutionary developments and community knowledge from NASA/EOS and integrate them into a sustained climate service from NOAA's operational satellite programs. The resulting CDRs provide authoritative multi-decadal global data sets and information (e.g., Arctic Ice Sheet Extent, Sea Level Rise, and Global Hurricane Intensity Trends) that are required to initiate and evaluate climate prediction models and to assess climate change mitigation and adaptation approaches.

NOAA Leadership in Federal Geographic Data Committee Efforts: The FGDC is an interagency committee that promotes the coordinated development, use, sharing, and dissemination of geospatial data on a national basis. This nationwide data publishing effort is known as the National Spatial Data Infrastructure (NSDI). NOAA is a leader within the FGDC membership. The NOAA Office of the Chief Information Officer serves as the Senior Agency Official for Geospatial Information within the Department of Commerce. NOAA staff chair two FGDC Subcommittees: The Geodetic Control, Marine and Coastal Spatial Data Subcommittee; and the Marine Boundary Working Group. NOAA has either the lead or co-lead responsibilities for the following NSDI spatial data themes: Baseline (Maritime); Climate; Elevation (Bathymetric); Geodetic Control; Marine Boundaries; and Shoreline. In addition, NOAA activities have resulted in the adoption of numerous standards and best practices within the FGDC community.

Global Observing Systems Information Center: The Global Observing Systems Information Center (GOSIC) (http://gosic.org) is a developmental activity begun in 1997 at the University of Delaware in order to aid the data management needs of the various global observing system efforts. Its transition to an operational global data facility at NOAA was completed on behalf of, and with the concurrence of, the global observing community in 2007. GOSIC provides information, and facilitates easier access to data and information produced by the Global Climate Observing System (GCOS), GOOS, and Terrestrial Observing Systems and their partner programs. GOSIC provides explanations of the various global data systems, as well as an integrated overview of the myriad global observing programs, which includes on-line access to their data, information, and services. GOSIC offers a search capability and data portal functionality across international data centers to enhance access to a worldwide set of observations and derived products, and partners with data centers around the world to aid in facilitating data access for all users.

3.4 Implementation Schedule

NOAA identified five projects that modernize its data and information management and address the themes of managing increasing data volumes, filling gaps and extending the environmental record, improving metadata and processing steps, improving data access, and enabling data integration.

Figure 9 shows the major milestones and implementation dates for these modernization projects.

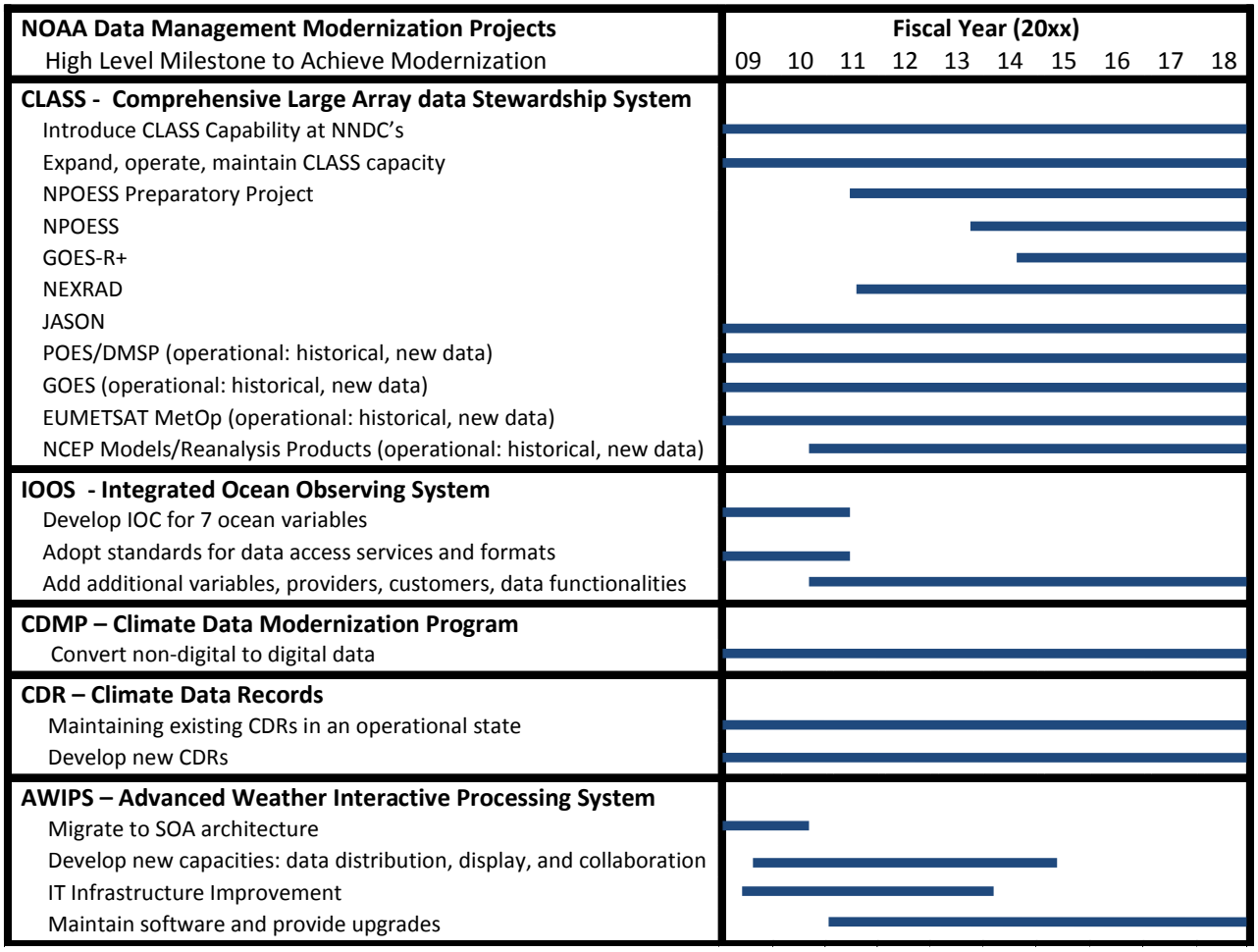


Figure 9: Implementation Schedule for NOAA Data Management Modernization Activities

NOAA Data Management Modernization Projects	FY 2008 Enacted ($K)	FY 2009 President's Budget ($K)	FY 2009 Enacted ($K)	FY 2009 ARRA* ($K)	FY 2010 President's Budget ($K)	Modernization Activity
CLASS	6,315	6,476	16,467	0	6,476	Managing increasing data volume
IOOS	26,360	21,055	26,500	0	21,110	Improving data access, metadata, integration
CDMP	20,942	4,063	20,980	0	4,063	Filling gaps in the historical environmental record
CDR	0	0	0	4,950	7,000	Extending the environmental record
AWIPS	36,863	38,065	38,065	0	39,346	Improving processing steps
* FY 2009 American Recovery and Reinvestment Act funding						

Figure 10: Funding for NOAA Data Management Modernization Activities

Figure 10 shows the associated budget table for FY 2008, FY 2009, and FY 2010 for these five projects.

Comprehensive Large Array data Stewardship System: CLASS is the enterprise Information Technology (IT) capability for NDCs, acting as a Web-based data storage and distribution system for NOAA's environmental data. It is currently utilized by NDCs for the distribution of operational environmental satellite data from NOAA's GOES and Polar-orbiting Operational Environmental Satellites and derived data products. CLASS is under development to support additional satellite data streams, such as GOES-R, NPP, and NPOESS. In addition, NEXRAD and modeled data are planned for inclusion in CLASS and the system is being evolved to provide a configurable set of tools for data ingest to allow rapid response to new requirements, additional tools for data management and stewardship by data center experts, and generalized access interfaces to allow tailored tools for data access. In the near term, efforts will focus upon operations and maintenance of CLASS components that have transitioned from development to operational status. Longer term plans for CLASS include expanding the safe storage/access capacity to meet the data influx expected from the operational introduction of data from radar, models, and new satellites. The current CLASS configuration can provide services for approximately four PBs of data so significant hardware investments will be necessary in the coming years to accommodate the increase in observational data and model outputs.

Integrated Ocean Observing System: In 2008, the IOOS Program developed an Initial Operating Capability for its portion of IOOS DMAC focusing on a limited subset of five ocean variables: temperature, salinity, sea level, surface currents, and ocean color. This effort is now underway with FY 2008 and FY 2009 resources and the list of variables was augmented to include wind and wave measurements. IOOS has targeted NOAA data providers who have observations of relevance to four customer focus areas of national importance: Hurricane Intensity forecasting, Coastal Inundation modeling, Harmful Algal Bloom prediction, and Integrated Ecosystem Assessments. The initial set of three data providers comprises the NWS National Data Buoy Center, the NOS Center for Operational Oceanographic Products and Services, and the NESDIS CoastWatch program. The providers adopted standardized data access services and formats, and the customers are now beginning to obtain some of their data from these services. This initial development project includes a plan to test and evaluate the resulting product enhancements and the standards used. As shown in Figure 11, IOOS will use these

evaluations to help form the expansion of the framework to encompass additional ocean variables, data providers, customers, and data management functionalities.

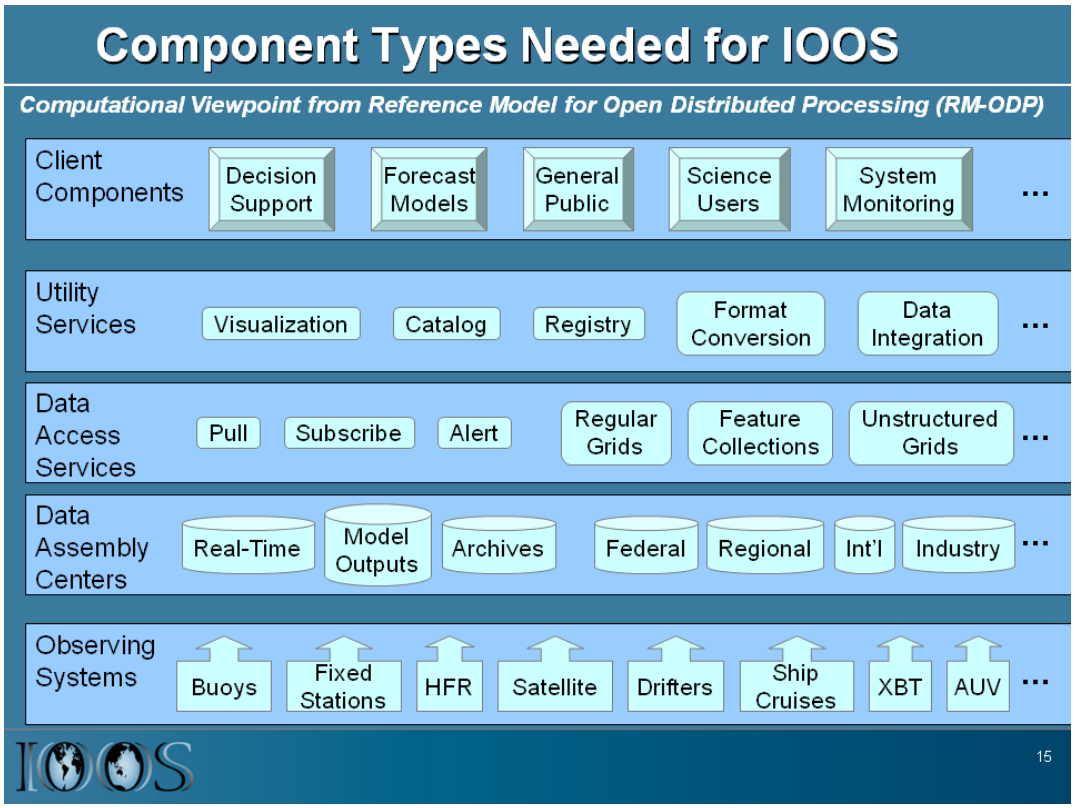


Figure 11: Data Management components that will be developed for IOOS

Climate Data Modernization Program: CDMP will continue to image and key in paper and microfilm records and make them available on the Web to researchers. These observations collected from numerous platforms and sensors range from space weather down through the Earth's biosphere and to sediment data of the ocean floor. CDMP will continue transforming these older observations to a more useful and accessible digital media, which will help meet the predicted demand for additional scientific baseline observations. Many significant decisions on future energy use, climate, and infrastructure issues will depend on the accuracy and availability of these data for predictive modeling. Although many tasks are complete, the work is far from done and over the next decade millions of pieces of data will be prioritized and then preserved to meet the needs of science and the scientific community.

Climate Data Records Program: CDR's goal is the production of high quality, multi-decadal time series data describing the global atmosphere, oceans, and land surface with a focus on essential climate variables as identified within GCOS. Building upon the initial successes of the Scientific Data Stewardship program, it supports the regular, operational production of CDRs for the atmosphere, oceans, and land surface. This production requires collaboration between experts in the climate community and experts in data management. It also must be informed by scientific application and associated user feedback on the accessibility and usability of the produced CDRs.

The program is developing three areas for operational and regular production of high-quality CDRs:

- Operational quality assurance of ingested data and regular monitoring of data quality and provenance
- Generation of authoritative, long-term records through rigorous data analysis and research that will validate and improve these CDRs
- Configuration management to ensure the documentation of the product and preservation of the information context to allow future users to understand, modify, and use the CDR.

The CDR Program activities encompass the full range of institutional diversity within the climate community. The program adopted approaches that allow organizations producing and archiving CDRs to maintain local autonomy within a context that encourages responsible participation in federations that foster increased data sharing, interdisciplinary data understanding, and improved assessments of data quality. An Announcement of Opportunity was published in FY 2009 to begin a CDR grants program. The CDR Program Office awarded the FY 2009 grants and is preparing to incorporate the results of the projects into the CDR framework. The authoritative nature and vitality of the CDRs will be maintained through peer reviews, user recommendations, and independent processing of data describing essential climate variables.

Advanced Weather Interactive Processing System: AWIPS is an advanced information processing, display, and telecommunications system that is the cornerstone of NWS modernization. AWIPS workstations and servers are installed at all 150 NWS Weather Forecast Offices and River Forecast Centers. AWIPS integrates meteorological and hydrological data, including satellite and radar data, enabling forecasters to prepare and issue more accurate and timely forecasts and warnings. NWS is migrating AWIPS to an improved, open system with new capacities for display, visualization, product generation, communications, and access to data. The updated AWIPS will employ a SOA that will modularize it to make AWIPS more flexible, adaptable, and scalable to accommodate new requirements. New features will be adopted once the new architecture is complete. The AWIPS migration to a SOA is being completed by the AWIPS Prime contractor who will complete development in 2010. Following an intensive acceptance test program, the software will deploy nationally in 2011.

AWIPS and other NOAA systems are also preparing for the arrival of new data described in section 2.2.1 of this document. NPP data will begin flowing in 2011 and GOES-R data begin flowing for testing in 2013 (launch in 2015). NOAA's IT infrastructure must be robust to exploit GOES-R, JPSS,, dual polarity radar data from NEXRAD, and high resolution modeling capabilities and to accept data from other major NOAA programs. NOAA seeks resources for these infrastructure requirements through its Planning, Programming, Budgeting and Execution System. The infrastructure changes required include upgrades to AWIPS, National Centers for Environmental Prediction, the Telecommunications Gateway, and NOAAnet.

4. Conclusion

NOAA assessed the adequacy of environmental data and information management for its four Mission Goals. Compared to the 2007 *Report to Congress*, the overall assessment of NOAA data and information management shows consistency and a general trend of improvement. At the goal level, there have been successes in the area of Data Discovery, Planning, and Complete Metadata. Challenges exist in the area of Appropriate Formats and continue in the Collect and Rescue area.

Five areas again emerge as the prominent themes that categorize the successes and challenges of NOAA environmental data and information management: managing increasing data volumes; filling gaps and extending the environmental record; improving metadata and processing steps; improving data access; and enabling data integration. Using a NRC report entitled "Environmental Data Management at NOAA," as a guide, NOAA made progress in addressing the increasing data volume challenge by developing a procedure to determine what data should be preserved versus those data that are deemed to have a limited lifetime. Advancements continue to be made by NOAA in addressing the challenges presented by the other themes.

The central strategy to modernize and improve NOAA's environmental data and information systems revolves around the unifying concepts of GEO-IDE. By using a SOA and a standards-based data and information infrastructure, the vision of GEO-IDE is enabling the integration of NOAA's existing data management systems with future development efforts that are enhancing discovery and access to NOAA's environmental information. Through GEO-IDE, NOAA is participating in efforts at the national USGEO level and international GEOSS level to facilitate the sharing and usage of global, regional, and local data from satellites, ocean buoys, weather stations, and other surface and airborne Earth-observing instruments.

In order to implement the modernization and improvement of the environmental data and information systems across NOAA, several near-term activities are identified and are underway. Five longer-term projects—CLASS, CDMP, AWIPS, IOOS, and CDR—are identified as necessary efforts to achieve modernization and improvement objectives. An implementation schedule and funding levels for FY 2008 and FY 2009 are included for these projects. With this strategy, NOAA remains committed to collecting, preserving, and providing the stewardship that will result in the highest quality data and information delivered securely and efficiently to our users.

Abbreviations and Acronyms List

AWIPS	Advanced Weather Interactive Processing System
C&A	Certifications and Accreditations
CDMP	Climate Database Modernization Program
CDR	Climate Data Records
CLASS	Comprehensive Large Array-data Stewardship System
COM	Climate Observations and Monitoring
COPB	CLASS Operations Planning Board
CRM	Climate Research and Monitoring
CSD	Climate Services Development
DAARWG	Data Archive and Access Requirements Working Group
DEM	Digital Elevation Model
DMAC	Data Management and Communications for IOOS
DMC	Data Management Committee
DMIT	Data Management Integration Team
DOC	Department of Commerce
E-SWDS	External Space Weather Database Store
EDM	Fisheries Enterprise Data Management
FGDC	Federal Geographic Data Committee
FY	Fiscal Year
GCOS	Global Climate Observing System
GEO-IDE	Global Earth Observation – Integrated Data Environment
GEOSS	Global Earth Observation System of Systems
GIS	Geographic Information System
GOES	Geostationary Operational Environmental Satellite
GOOS	Global Ocean Observing System
GOS	Geospatial One-Stop
GOSIC	Global Observing Systems Information Center
IEOS	Integrated Earth Observation System
IOCM	Integrated Ocean and Coastal Mapping
IOOS	Integrated Ocean Observing System
ISO	International Organization for Standardization
IT	Information Technology
MADIS	Meteorological Assimilation Data Ingest System
NAO	NOAA Administrative Order
NARA	National Archives and Records Administration
NASA	National Aeronautics and Space Administration
NDBC	NOAA's National Data Buoy Center
NDC	NOAA Data Centers
NEF	National Essential Functions
NERI	National Estuaries Restoration Inventory
NESDIS	National Environmental Satellite, Data, and Information Service
NextGen	Next Generation Air Transportation System

NEXRAD	Next Generation Weather Radar
NGS	National Geodetic Survey
NIDIS	National Integrated Drought Information System
NMFS	National Marine Fisheries Service
NOAA	National Oceanic and Atmospheric Administration
NODC	National Oceanographic Data Center
NOS	National Ocean Service
NOSC	NOAA Observing Systems Council
NPOESS	National Polar-orbiting Operational Environmental Satellite System
NPP	NPOESS Preparatory Project
NRC	National Research Council
NSDI	National Spatial Data Infrastructure
NWS	National Weather Service
OFCM	Office of the Federal Coordinator for Meteorology
OPUS-RS	Online Positioning User Service- Rapid Static
PaCIS	Pacific Climate Information System
PB	Petabyte (one quadrillion bytes or 1,000,000,000,000,000 bytes)
PRICIP	Pacific Region Integrated Climatology Information Products
RCDB	NOAA Restoration Center Data Base
REDM	Regional Ecosystem Data Management
RWIS	Road Weather Information System
SDS	Scientific Data Stewardship
SNAAP	Simple NOAA Archive Access Portal
SOA	Service-Oriented Architecture
SWPC	Space Weather Prediction Center
TB	Terabyte (one trillion bytes or 1,000,000,000,000 bytes)
USDP	U.S. Drought Portal
USGEO	United States Group on Earth Observations
WIDB	Weather Information Data Base

Appendix A - Public Law 102-567, Section 106: Data and Information Systems

(c) NEEDS ASSESSMENT FOR DATA MANAGEMENT, ARCHIVAL, AND DISTRIBUTION

(1) Not later than 12 months after October 29, 1992, and at least biennially thereafter, the Secretary of Commerce shall complete an assessment of the adequacy of the environmental data and information systems of the National Oceanic and Atmospheric Administration. In conducting such an assessment, the Secretary shall take into consideration the need to –

A) provide adequate capacity to manage, archive, and disseminate environmental data and information collected and processed, or expected to be collected and processed, by the National Oceanic and Atmospheric Administration and other appropriate departments and agencies;
B) establish, develop, and maintain information bases, including necessary management systems, which will promote consistent, efficient, and compatible transfer and use of data;
C) develop effective interfaces among the environmental data and information systems of the National Oceanic and Atmospheric Administration and other appropriate departments and agencies;
D) develop and use nationally accepted formats and standards for data collected by various national and international sources; and
E) integrate and interpret data from different sources to produce information that can be used by decision makers in developing policies that effectively respond to national and global environmental concerns.

(2) Not later than 12 months after October 29, 1992, and biennially thereafter, the Secretary of Commerce shall develop and submit to the Committee on Commerce, Science, and Transportation of the Senate and the Committee on Science, Space, and Technology of the House of Representatives a comprehensive plan, based on the assessment under paragraph (1), to modernize and improve the environmental data and information systems of the National Oceanic and Atmospheric Administration. The report shall –

A) set forth modernization and improvement objectives for the 10-year period beginning with the year in which the plan is submitted, including facility requirements and critical new technological components that would be necessary to meet the objectives set forth;
B) propose specific agency programs and activities for implementing the plan;
C) identify the data and information management, archival, and distribution responsibilities of the National Oceanic and Atmospheric Administration with respect to other federal departments and agencies and international organizations, including the role of the National Oceanic and Atmospheric Administration with respect to large data systems like the Earth Observing System Data and Information System; and

provide an implementation schedule and estimate funding levels necessary to achieve modernization and improvement objectives.

Appendix B–Goal Team Assessment Summaries by NOAA Program

A more detailed assessment of end-to-end data management functions is provided below for the NOAA Mission Goals. A single chart is presented for each NOAA Goal and these charts provide an assessment at the individual program level of that Goal. A green cell in the chart indicates the program is accomplishing 100 percent of the required activity or is able to do so with current resources. A yellow cell indicates the activity is partially accomplished or an incremental increase in resources is required to reach the 100 percent level. A red cell indicates the activity is not accomplished or that significant new resources are required. A gray cell indicates that the activity was considered to be not applicable to the associated program. Symbols are added in each chart to indicate an improvement (>) or a decline (<) in capabilities by one level that has occurred since 2007.

With the exception of the Climate Goal, the programs within the other NOAA Goals have remained the same as those described in the 2007 *Report to Congress*. The Climate Goal was restructured in 2008 from five programs down to three programs with the intent of more efficiently and effectively managing the program. The assessment of data management activities in this *Report to Congress* is based on this new three program structure consisting of: (1) Climate Observations and Monitoring (COM); (2) Climate Research and Modeling (CRM); and (3) Climate Services Development (CSD). The CRM and CSD programs are new programs combining different sets of activities from the former Climate Forcing, Regional Decision Support, and Climate and Ecosystems programs. The COM program essentially covers what was in the former Climate Observations and Assessment program. The CRM and CSD programs should be considered as new programs, and as such the stoplight chart for those programs forms a new baseline from which to report progress for future reports.

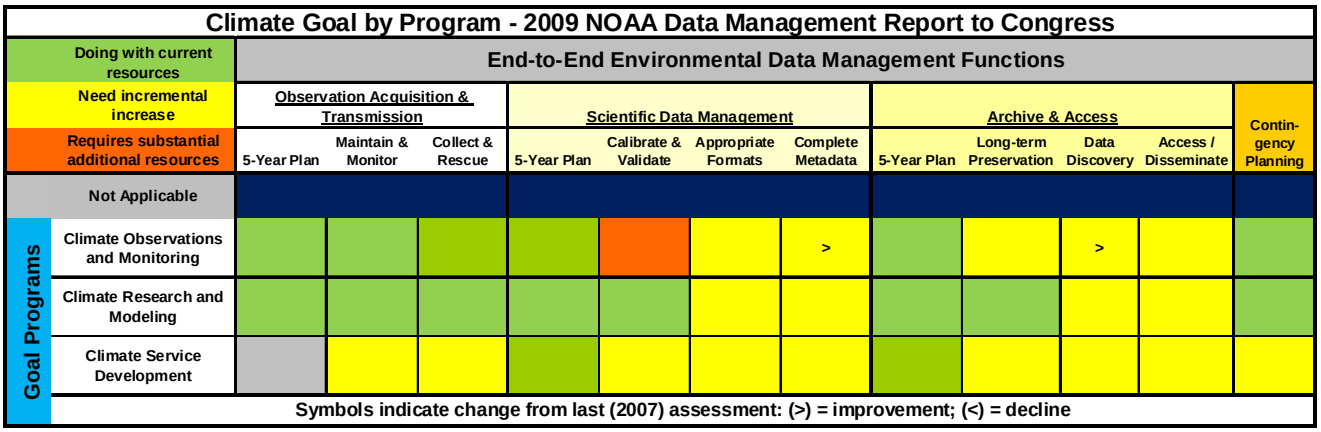
As compared to the 2007 *Report to Congress*, almost all NOAA programs showed consistency or changes in two or less of the surveyed data management components. The exceptions are the Commerce & Transportation Goal's Surface Weather Program and Aviation Weather Program, the Ecosystem Goal's Habitat Program, and the Weather & Water Goal's Space Weather Program.

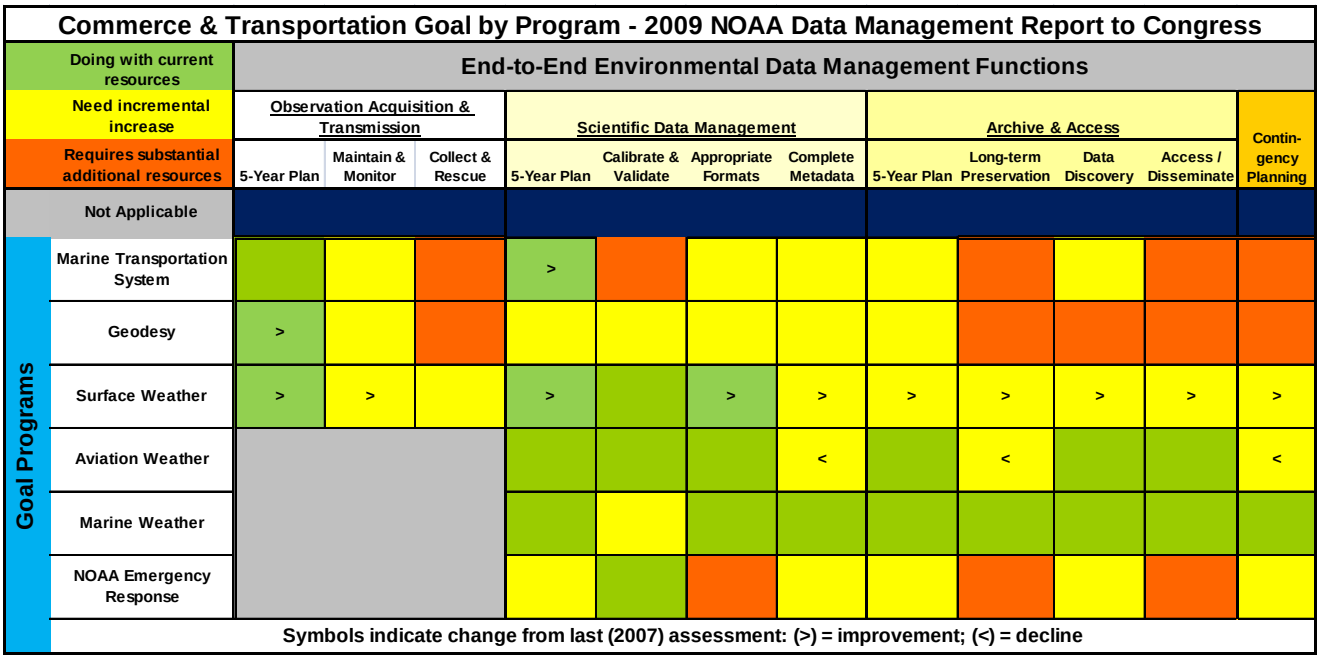
The Commerce & Transportation Goal's Surface Weather Program showed improvement in 10 of the 12 data management components. This has been attributed to the implementation of the Meteorological Assimilation Data Ingest System (MADIS) which has allowed NOAA to make great strides over the past two years in acquiring and disseminating surface transportation observations to internal and external users. MADIS now collects weather observations from 26 state-owned (and one private) Road Weather Information System networks. These observations are distributed to NWS Forecast Offices via AWIPS, and to the public via the web. The Weather and Water Goal is concerned about properly supporting MADIS to develop more comprehensive metadata standards for storage and management of disparate data sets. MADIS is currently being transitioned into full operational capability scheduled for FY 2011. MADIS will become an integral component of National Mesonet initiative of NOAA, to leverage existing observations from non-Federal networks which are sited, operated, and documented in accordance with national and international standards. The objective of building a National Mesonet is to monitor and predict meteorological and hydrological phenomena at temporal resolution less than 15 minutes and spatial resolution less than 5 km, focusing on high priority regions (e.g., mountains, coastal, urban). Beginning in FY 2009, NOAA is funding activities designed to enable the integration of mobile, vehicle observations, and to collect information on established networks' critical metadata needs.

The Commerce & Transportation Goal's Aviation Weather Program is another exception due to the NextGen project which showed a decline in three of nine surveyed data management components. This change in status is the result of new requirements for data management identified in the planning process for the NextGen WIDB. The development and implementation of this virtual database will require a gradual increase of funding above the FY 2010 level to support the metadata development and contingency planning activities and to provide for the long term archival of the significant quantities of weather information available through the WIDB. The NextGen program will heavily leverage the data management advances implemented by the MADIS and AWIPS programs and be fully compliant with GEOSS and GEO-IDE standards.

The Ecosystem Goal's Habitat Program showed a decline in the Maintain and Monitor component of Observation Acquisition and Transmission. This was due to the maintenance and monitoring needs expanding as the Habitat Program scope has grown and as new data sources have been incorporated. Improvements were noted in the components of Complete Metadata and Data Discovery.

The final exception is the Weather & Water Goal's Space Weather Program which showed a decline in 5 of the 12 surveyed data management components. This program reports that adequate back-up systems do not exist in regard to an Alternate Space Weather Operations Center and an Alternate Processing Facility. In 2008, SWPC entered into an agreement between the NWS and the Air Force Weather Agency to create an alternate processing facility for products and forecaster displays at Offutt Air Force Base in Nebraska. Full build out of this facility is dependent on budget augmentation currently forecasted to occur in FY 2011. In 2010, SWPC plans on replicating the critical space weather database at both Air Force Weather Agency and NWS backup facilities. In an effort to ensure a secure and compliant computing environment, the Space Weather Program completed a comprehensive mitigation plan to address all deficiencies within the next three years, assuming additional resources are available in FY 2010 and FY 2011. Based on a late FY2009 funding augmentation, the Program anticipates having a contract in place to hire programmers and system administrators to mitigate critical Controls and Accessories findings. In addition, the Space Weather Program relies on solar wind data from a NASA satellite which has now exceeded its operational life, with no backup system to replace it when it fails. NOAA, NASA, and the Department of Defense are currently working together with the Office of the Federal Coordinator for Meteorology (OFCM) to identify options for a replacement. OFCM leads the Committee for Space Environment Sensor Mitigation Options to provide recommendations in September 2009 that will include resources required for each option. There is also a need to migrate to more standard data formats, the need to transition to more modern storage media, and to increase the availability of on-line access.


Climate Goal by Program - 2009 NOAA Data Management Report to Congress
Doing with current resources
Need incremental increase
Requires substantial additional resources
Not Applicable
End-to-End Environmental Data Management Functions
Observation Acquisition & Transmission
Scientific Data Management
Archive & Access
Contin-gency Planning
5-Year Plan
Maintain & Monitor
Collect & Rescue
5-Year Plan
Calibrate & Validate
Appropriate Formats
Complete Metadata
5-Year Plan
Long-term Preservation
Data Discovery
Access / Disseminate
Goal Programs
Climate Observations and Monitoring
Climate Research and Modeling
Climate Service Development
>
>
Symbols indicate change from last (2007) assessment: (>) = improvement; (<) = decline


Commerce & Transportation Goal by Program - 2009 NOAA Data Management Report to Congress
Doing with current resources
Need incremental increase
Requires substantial additional resources
Not Applicable
End-to-End Environmental Data Management Functions
Observation Acquisition & Transmission
Scientific Data Management
Archive & Access
Contin-gency Planning
5-Year Plan
Maintain & Monitor
Collect & Rescue
5-Year Plan
Calibrate & Validate
Appropriate Formats
Complete Metadata
5-Year Plan
Long-term Preservation
Data Discovery
Access / Disseminate
Goal Programs
Marine Transportation System
Geodesy
Surface Weather
Aviation Weather
Marine Weather
NOAA Emergency Response
>
>
>
>
>
>
>
>
>
>
>
>
>
<
<
<
Symbols indicate change from last (2007) assessment: (>) = improvement; (<) = decline

Ecosystem Goal by Program - 2009 NOAA Data Management Report to Congress

		End-to-End Environmental Data Management Functions											
Doing with current resources		Observation Acquisition & Transmission			Scientific Data Management				Archive & Access				Contin-gency Planning
Need incremental increase													
Requires substantial additional resources		5-Year Plan	Maintain & Monitor	Collect & Rescue	5-Year Plan	Calibrate & Validate	Appropriate Formats	Complete Metadata	5-Year Plan	Long-term Preservation	Data Discovery	Access / Disseminate	
Not Applicable													
Goal Programs	Habitat		<					>			>		
	Corals												
	Coastal and Marine Resources												
	Protected Species												
	Fisheries Management												
	Aquaculture					>	>						
	Enforcement												
	Ecosystem Observations							>					
	Ecosystem Research												

Symbols indicate change from last (2007) assessment: (>) = improvement; (<) = decline

Weather & Water Goal by Program - 2009 NOAA Data Management Report to Congress

		End-to-End Environmental Data Management Functions											
Doing with current resources		Observation Acquisition & Transmission			Scientific Data Management				Archive & Access				Contin-gency Planning
Need incremental increase													
Requires substantial add. Resources		5-Year Plan	Maintain & Monitor	Collect & Rescue	5-Year Plan	Calibrate & Validate	Appropriate Formats	Complete Metadata	5-Year Plan	Long-term Preservation	Data Discovery	Access / Disseminate	
Not Applicable													
Goal Programs	Air Quality												
	Coast, Estuaries and Ocean												
	Hydro												
	Local Forecasts									>			
	Space Weather		<	<			<					<	<
	Science and Tech Infusion												
	Tsunami			>									

Symbols indicate change from last (2007) assessment: (>) = improvement; (<) = decline