4B.4 **"NOAA'S CLIMATE DATABASE MODERNIZATION PROGRAM PAVING THE ROAD FOR DATA STEWARDSHIP"**

William E. Angel*
NOAA National Climatic Data Center, Asheville, North Carolina

Thomas F. Ross
NOAA National Climatic Data Center, Asheville, North Carolina

Raymond T. Truesdell
IBM Corporation, Asheville, North Carolina

***Corresponding Author Address:** William E. Angel, NOAA/NCDC, 151 Patton Ave, Asheville, NC 28801-5001; email: William.E.Angel@noaa.gov.

ABSTRACT

The Climate Database Modernization Program (CDMP) supports NOAA's 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 NOAA's holdings were originally recorded on paper, film, and other fragile media, and stored at various facilities. CDMP's mission includes transforming these older observations to a more useful friendly and accessible digital media which will help meet the predicted demand for additional scientific baseline observations. Many significant decisions such as future energy use, climate, and infrastructure issues will depend on the accuracy and availability of this data for predictive modeling. The scientific community depends on the ongoing work of CDMP to ensure that the hard-earned, irreplaceable research is protected, preserved and made available on-line. CDMP partners with four private sector contractors and have placed online over 54 million weather and environmental images, available to researchers around the world via the Internet. The amount of data online has grown from 1.75 terabytes in 2001 to over 14 terabytes in late 2010 and millions of pieces of data are still waiting to be digitized. This presentation will highlight several of the 97 various NOAA national and international data rescue projects under the CDMP program.

4B.4 "NOAA'S CLIMATE DATABASE MODERNIZATION PROGRAM PAVING THE ROAD FOR DATA STEWARDSHIP"

William E. Angel*
NOAA National Climatic Data Center, Asheville, North Carolina

Thomas F. Ross
NOAA National Climatic Data Center, Asheville, North Carolina

Raymond T. Truesdell
IBM Corporation, Asheville, North Carolina

1. Project Goals and Background

The year 2010 marked the beginning of the Climate Database Modernization Program's (CDMP) second decade of service. CDMP continues to support every NOAA line office by preserving and enhancing the availability of valuable climate and environmental data. These data are used by researchers and others dealing with climate and environmental issues crucial to our planet and our global society.

2. NOAA's Stewardship Commitment

The Climate Database Modernization Program supports NOAA's 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 (Figure 1). Without proper preservation of the media, the information they contained was in danger of being lost forever.

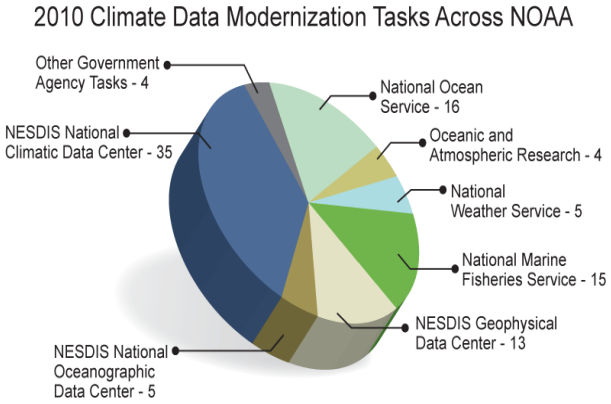

Figure 1: *The number of NOAA's CDMP Projects reached a record total of 97 in 2010.*

3. A New Decade Begins

The National Oceanic and Atmospheric Administration's Climate Database Modernization Program (NOAA's CDMP) has just completed its eleventh year. When the program began in the year 2000, the demand for rapid and complete access to the Nation's and world's climate data by researchers and global change scientists was a key motivation in the establishment of CDMP, which is managed through NOAA's National Climatic Data Center (NCDC) located in Asheville, NC. This program was initiated by Congress to assist NOAA in modernizing and improving access to the Nation's climate data and information.

As CDMP has entered its second decade, demand for climate data access has increased. In addition, CDMP has expanded its support across all of NOAA, preserving and making available environmental data ranging from the ocean floor to the top of the ionosphere. Partnering with four private sector contractors, CDMP has placed online over 54 million weather and environmental images, available to researchers around the world via the Internet. The amount of data online has grown

* Corresponding author address: William E. Angel, NOAA/NCDC, 151 Patton Ave, Asheville, NC 28801-5001; email: William.E.Angel@noaa.gov.

from 1.75 terabytes in 2001 to over 14 terabytes in 2010. Major progress continues in making these data available through a number of NOAA web sites (see URL list in section 8 and 9).

In addition, hourly weather records keyed through CDMP continue to be integrated into NCDC's digital database holdings extending the period of record for many stations back into the 1800's. The increase in the quality and quantity of historical data is helping researchers worldwide to improve real-time monitoring and forecasting of environmental, solar and geophysical events.

4. NOAA's CDMP Project Partners

Figure 2: *CDMP Support Contractors.*

CDMP could not exist without the extraordinary efforts of people within NOAA and those in the private sector who do the keying, imaging, and database development. CDMP projects have created hundreds of new private sector data entry and information management jobs in several economically challenged areas in West Virginia, Kentucky, and Maryland. The project tasks supported by CDMP are well suited for the private sector. Many of these tasks have been shifted from government employees to CDMP contractors in the above mentioned states. Tasks performed by these contractors include the printing and distributing of the NCDC serial climate publications, managing accounts receivable, imaging and keying incoming records, hosting and maintaining online images, and providing expert personnel in support of various projects.

The three prime contractors (Figure 2) for CDMP are: National Interest Security Company, an IBM Company, Rocket Center, West Virginia; SourceCorp, Mount Vernon, Kentucky; and HOV Services, Beltsville, Maryland. Excellent support is also provided by the NCDC on-site contractor, STG Corporation, whose staff prepares many of the data for shipment and performs extensive quality control on the returning data products. The CDMP program employs nearly 300 highly skilled government and contractor personnel dealing with climate and environmental data rescue tasks. With over 97 projects ongoing, the contractors must remain focused and flexible to meet each project's requirements.

5. CDMP's National Highlights

The task of modernizing NCDC's vast archive of weather and climate data continues. While millions of records have been imaged, about a third of the paper documents held by NCDC, and much of its microfilm and microfiche archive, remains unscanned. Accompanying this imaging task, keying of many of the scanned images is ongoing, further supplementing NCDC's digital data resources. Currently, more than 54 million images are available to researchers and educators thru the Environmental Document Access and Display System (EDADS) (Figure 3).

NCDC's digital data for hourly surface observations generally began around 1948. CDMP has imaged over 2 million forms and keyed over 405 million records. As a result, this has added over 50 years of hourly and synoptic data for over 225 stations and is being integrated into NOAA NCDC's Integrated Surface Hourly (ISH) database (completed by the end of 2010). Another example of NOAA's NCDC data integration is from the CDMP "Forts" Project. Data from the historic Signal Service, Smithsonian, and other weather observations from 18th and 19th centuries are being saved. Data for over 350 stations have already been keyed.

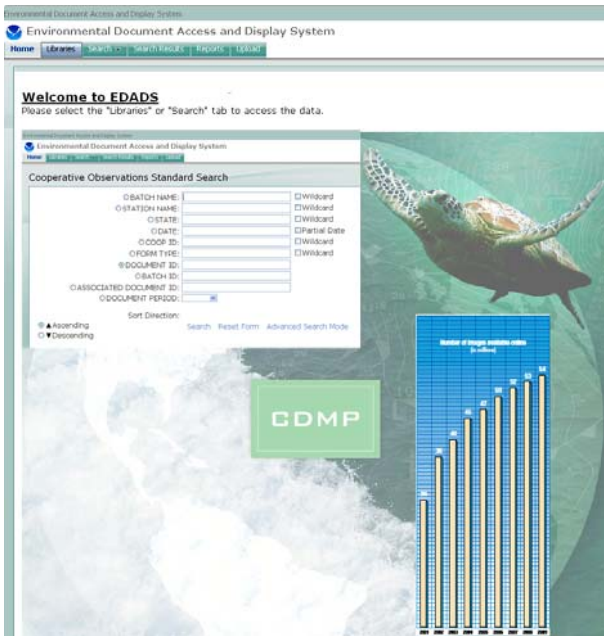

Figure 3: *A screen shot of the EDADS system.*

Work continues with NOAA's National Hurricane Center and Atlantic Oceanographic and Meteorological Laboratory to digitize historical hurricane reconnaissance data and "storm wallet" collections; the first of these datasets are already being utilized by hurricane specialists and storm researchers. Several projects with the NOAA's National Ocean Service and NOAA's National Marine Fisheries Service are making available rare historical nautical charts, plankton databases, shoreline data, fish surveys in Alaska, and much more. CDMP produced images of historical glacier photographs via the NOAA's National Geophysical Data Center have already made global headlines, with more photos still to be digitized.

One project started in 2009 is expanding the digital database of northwest Atlantic fisheries, extending hydrographic and plankton databases for this area back in time and providing invaluable data on how environmental change and extensive fishing have impacted fish stocks in the Atlantic. Two other tasks are digitizing rare surface weather observations from data-sparse areas in the south-western U.S. in Navajo Nation and from San Cristobal in the Galapagos Islands, providing more data to help initialize and verify vital climate change models. Another ongoing project is to digitize rare aerial photography of our Nation's coastlines from as far back as the 1920's, helping coastal scientists better understand how erosion and other factors have impacted our beaches and shorelines.

6. CDMP's International Highlights

International cooperation through the NOAA's National Weather Service International Activities (IA) is bringing valuable African, Asian, and South American climate data into global databases, and CDMP is actively involved with the World Meteorological Organization's (WMO) Commission for Climatology. NOAA's CDMP International partners are the International Environmental Data Rescue Organization (IEDRO) and the Atmospheric Circulation Reconstructions over the Earth (ACRE).

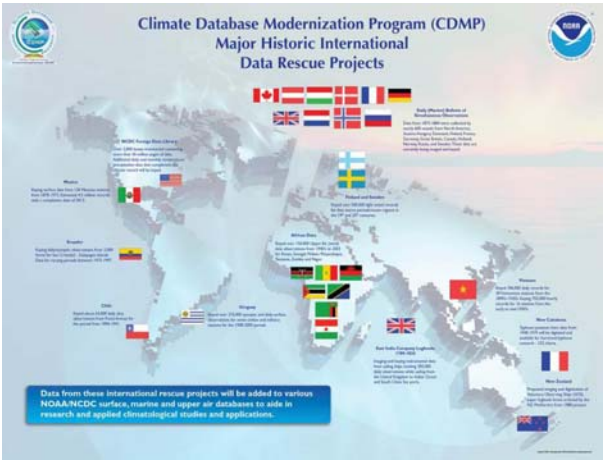

Figure 4: *Map showing the locations of CDMP's International Projects*.

Data from more than two dozen international data rescue projects will be added to NOAA's baseline (surface, marine, upper air, biological and ecological) databases to aid in research and applied climate and environmental studies and applications (Figure 4). Internationally to date, over 850,000 surface observations and over 200,000 upper-air have been rescued. Other examples are as follows:

East India Company Logbooks provide needed coverage for the Atlantic and Indian Ocean: Nearly 900 log books provide marine observations during 1788-1834 where existing marine data is very limited. This collection will support regional and global climate research and reanalysis efforts to the late 18th century and provide clues of the earth's climate during this period. One such recently keyed observation revealed a Tropical Cyclone encounter in 1827. On a return voyage from Whampoa, China, to England, there is strong evidence that the Lady Melville of the English East India Company encountered a tropical cyclone in the Southern Indian Ocean off the coast of

Madagascar in April, 1827. This tropical cyclone was not known prior to the time of keying.

Data from Tanzania keyed: CDMP continues to receive African "pilot balloon" (pibal winds and temperature) upper-air data in cooperation with the Tanzania Meteorological Agency. These data from the 1960's and 1970's have then been keyed for three of the thirteen pibal stations in Tanzania. The digitized data will be incorporated into NOAA's Integrated Global Radiosonde Archive database, filling gaps and holes in a data-sparse region, and leading to better analysis of upper air trends and patterns.

Historic Alaska Observations Added to Database: Daily weather data from several historic Alaskan stations, dating back to the 1860's, were keyed as part of the CDMP "Forts" project. Five weather stations, with data periods ranging from the 1860's to the 1890's, were indentified and keyed. The longest complete period of daily data was for Sitka, AK, from 1867-1877. This set dove-tails with prior records from Sitka uncovered by NOAA's Pacific Marine Environmental Laboratory. This is a very valuable find for climate scientists, as observations from data-sparse polar regions during this period are rare.

7. Future Work of CDMP

In 2010, the number of NOAA projects supported through CDMP reached an all-time high of 97 tasks (Figure 5). Most of these are multi-year projects and without continued support through CDMP, these projects would be left incomplete and valuable climate and environmental data would be left undiscovered and unavailable to the scientific community.

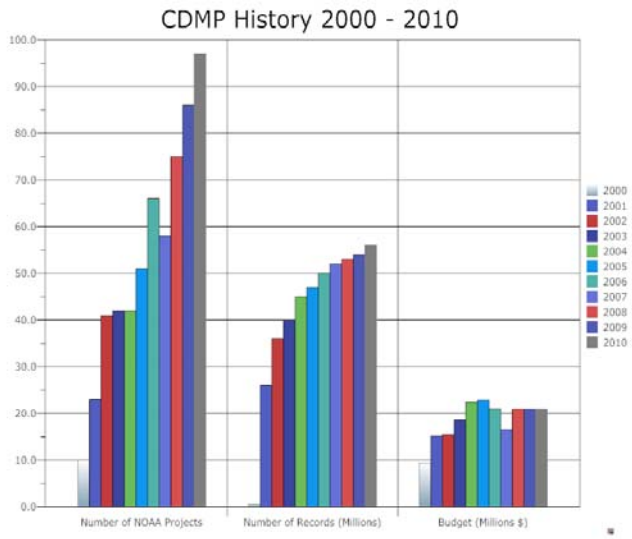

Figure 5: *Chart showing CDMP's eleven year history of projects, records, and budget*.

The task of modernizing NCDC's vast archive of weather and climate data continues. While millions of records have been imaged and keyed, and at the current pace, there are still years of work ahead to image and key the remaining archived NCDC data.

Ongoing CDMP projects involving other agencies within NOAA are also expected to extend well into the program's second decade. To further extend the program's mission, additional proposals are solicited and accepted each year as funding permits. Just this year, eleven new tasks were added to the docket, with still more proposals recently submitted for 2011 at CDMP's annual proposal workshop.

8. Web Addresses for NOAA Organizations

Climate Database Modernization Program (CDMP)
www.ncdc.noaa.gov/cdmp.html

National Oceanic and Atmospheric Administration (NOAA)
www.noaa.gov

NOAA's National Environmental Satellite, Data, and Information Service (NESDIS)
www.nesdis.noaa.gov

NOAA's National Climatic Data Center (NCDC)
www.ncdc.noaa.gov

NOAA's National Geophysical Data Center (NGDC)
www.ngdc.noaa.gov

NOAA's National Oceanographic Data Center (NODC)
www.nodc.noaa.gov

NOAA's National Ocean Service (NOS)
http://oceanservice.noaa.gov

NOAA's National Marine Fisheries Service (NMFS)
www.nmfs.noaa.gov

NOAA's National Weather Service (NWS)
www.nws.noaa.gov

NOAA's Office of Oceanic and Atmospheric Research (OAR)
www.oar.noaa.gov

9. Web Addresses for Selected CDMP Projects

Oceanographic and Meteorological Laboratory's Atlantic Hurricane Research Division
www.aoml.noaa.gov/hrd

National Hurricane Center's "Storm Wallet" Tropical Cyclone Archive
www.nhc.noaa.gov/pastall.shtml#wallet

National Marine Fisheries Service COPEPOD Global Plankton Database
www.st.nmfs.noaa.gov/plankton

Multi-Lens Air Photos Retrieval System (prototype)
www.ngdc.noaa.gov/dmsp/multilens

NOAA Central Library's TIROS & Satellite Meteorology website
www.lib.noaa.gov/collections/TIROS/tiros.html

National Snow and Ice Data Center Glacier Pairs
http://nsidc.org/data/glacier_photo/index.html

10. Conclusion

The Climate Database Modernization Program – or more precisely, the outstanding group of people that make it work – has clearly accomplished a great deal in the last eleven years. But the mission of CDMP is far from accomplished. In many ways, the work has just begun.

The people associated with CDMP are proud of the program's accomplishments over the last eleven years. But the accomplishments of the past years only serve to underscore the importance of the program. The CDMP mission to make climate and environmental data more accessible is as vital now as it was over a decade ago.

List of Figures