

**Climate Database
Modernization
Program**

**Annual Report
2000**

The Climate Database Modernization Program (CDMP) was established in late 1999 with the goal of modernizing the Nation's ability to access climatological data. In the past, utilization of climatological data was severely limited due to the manually intensive methods required for data retrieval, and the lack of digitized records. Under the CDMP, data has become more accessible to researchers, the public, and economic and political decision makers. As these data become fully accessible, they will provide important information to help planners mitigate the effects of hurricanes, tropical storms, flooding and other severe phenomena. Access to these data will also enable decision makers to improve economic and political decisions by allowing them to base their decisions on fact rather than speculation.

Under the CDMP, data has become more accessible to researchers, the public, and economic and political decision makers

The CDMP focus during the past year has been on increasing the utilization of climate records by improving and facilitating the quality, completeness, and access to these climatological data archives. Many tasks were undertaken to further the digital climate database modernization efforts. The following report outlines the accomplishments made during the year under the CDMP, and how these accomplishments helped to increase the modernization and utilization of this Nation's climate databases.

Government and Private Industry as Partners

The National Climatic Data Center (NCDC), part of the National Oceanic and Atmospheric Administration's (NOAA) National Environmental Satellite, Data, and Information Service, formed contractual partnerships with three private sector companies under CDMP: Information Manufacturing Corporation (IMC) located in Rocket Center, WV; Image Entry, Inc., (IE) headquartered in London, KY; and DoxSys, Inc., of Bethesda, MD, whose subcontractor, Lason, performed CDMP tasks in their Beltsville, MD, facility. These companies are leaders in data conversion, database building and data utilization services. The CDMP work performed by these private sector industries provided significant employment opportunities in several impoverished geographical areas of the country. Tasks accomplished during the year provided opportunities for a vast pool of workers to develop the skills needed to move into the workforce of the future by introducing them to an information technology environment. This pairing of government and private industry provided increased private sector employment, and helped to further the modernization of America's climate databases, providing a multitude of benefits to the Nation.

The pairing of government & private industry increased private sector employment and helped modernize America's climate databases

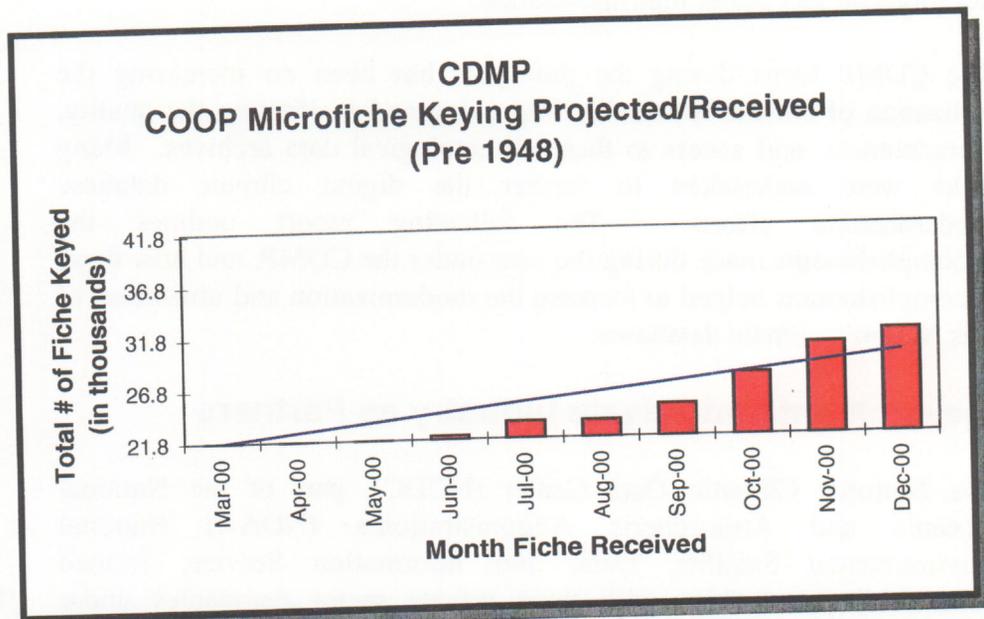
Modernization and Utilization Accomplishments

Database Modernization

IE worked toward completion of a digital database containing daily cooperative observations

The climate databases were substantially increased this year due to the CDMP efforts. IE developed and implemented a production system to work toward completion of a digital database containing daily cooperative observations. More than .5 billion characters from 9,342 microfiche were keyed into digital form by CDMP contractors during the year. These data have been entered into NCDC's mass storage system.

More than .5 billion characters keyed from over 9,000 microfiche were added to the existing database



48-hour turnaround from data receipt to image access via web-based system

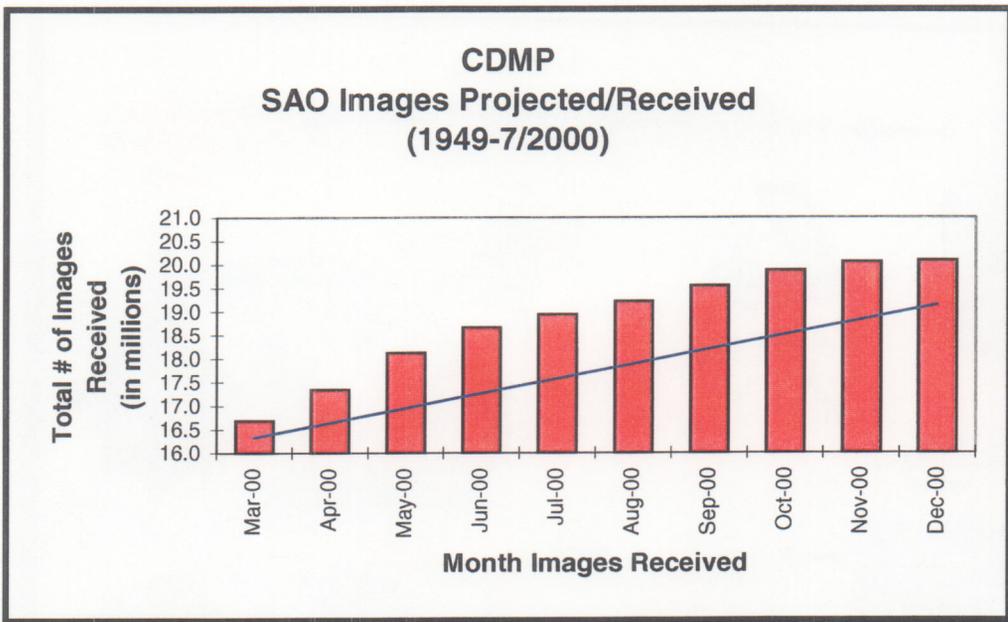
IE, in conjunction with IMC, also developed a records ingest system that provides timely digitizing and access to climatological data through a web-based, on-line system. Climate records are sent from various National Weather Service and Federal Aviation Administration observing sites throughout the United States to the London, KY, site where they are sorted, digitally imaged, and data from selected stations are keyed into digital form. Over 8,000 daily, and over 5,000 hourly, surface observation forms are sent to IE each month for digitizing and keying. The digital images are sent to IMC in West Virginia, where they are placed on their web-based, on-line system. Working together, IE and IMC are able to provide a 48-hour turnaround time from receipt of paper-based climate data to digital image accessibility. This quick turnaround time has substantially increased access to current climatological data. Concurrently, select keyed data are sent to NCDC, where the data are quality controlled and integrated into existing

(Continued on page 5)

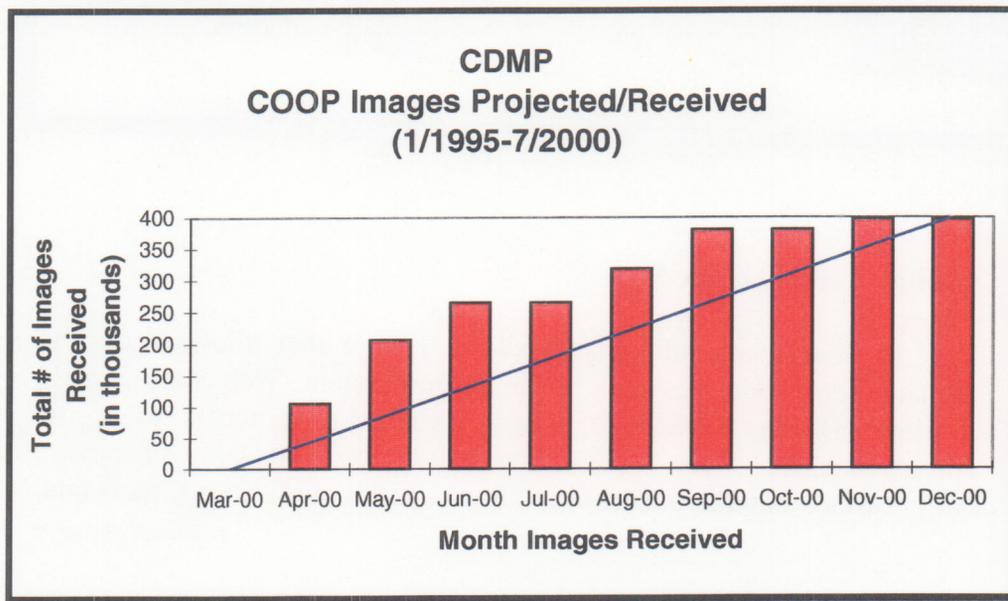
(Continued from page 4)

databases. These data are used to compile a variety of meteorological products and publications, and provide near-real-time accessibility to researchers, government agencies, and the public sector.

During the year, IMC imaged 3,732,676 pages of historical hourly surface airway forms from the years 1949 through 1998. They also produced digital images of daily cooperative observer forms from 1995 to the present. More than .5 million images of historical cooperative observer forms have been delivered in digital form. This addition of over 4 million climatological images greatly increased the existing climate databases.



More than 3,732,676 images of historical hourly surface airway forms were added to the existing database under CDMP



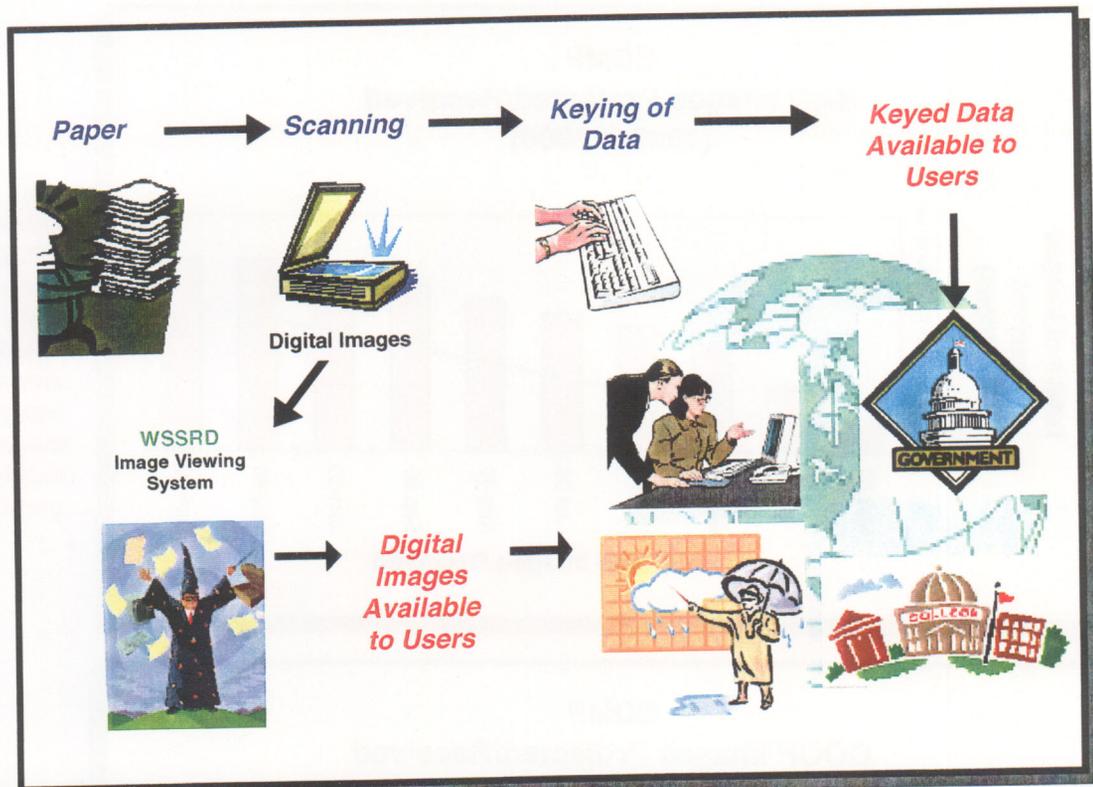
.5 million images of historical cooperative observer forms have been delivered in digital form

Utilization and Access

After NCDC quality assurance procedures are completed, climate data are integrated into current digital archives

To improve the utilization of these climate records, NCDC has strengthened its quality assurance procedures to ensure that images and keyed data received from CDMP contractors are of the highest quality. After NCDC quality assurance procedures are completed, climate data are integrated into current digital archives. Due to CDMP efforts, vast amounts of previously inaccessible data are now in digital form and are available to users around the world.

Vast amounts of previously inaccessible data are now available to users around the world



On-line Access via the Web

IMC's on-line system allows near-real-time access to digital climate images

IMC has developed an image viewing system that allows access to climate records via the web. Their on-line system, Web Store Search, Retrieve, Display (WSSRD), allows near-real-time access to digital climate images. Over 2 million digital images consisting of surface daily, surface monthly, and marine forms, serial publications, metadata,

(Continued on page 7)

(Continued from page 6)

handbooks and circulars are now web accessible. Various climate records that once resided only in the basement archives of NCDC are now seeing the "light of the world wide web." NCDC is working to integrate its On-Line Store system with WSSRD to provide seamless and totally integrated web access to customers.

Subscription Services

IMC completed the development and implementation of a redesigned subscription services database. This comprehensive database contains over 20,000 customer records and approximately 60,000 subscription orders. By creating, maintaining, and utilizing this updated database system, IMC employees can perform all functions needed to provide complete subscription services to customers. These services include invoicing, subscription renewals, and mailing of NCDC climate publications. In the process of building the subscription database, a massive address cleanup was undertaken. By having addresses conform to U.S. Postal Service standards, the cost of mailing will be substantially reduced. The new subscription database also allows IMC to provide comprehensive monthly reports on subscription activities.

IMC developed and implemented a redesigned subscription services database

To better provide services to customers, IMC established a toll free number dedicated solely to subscription services customers. Subscription services staff are available 12 hours per day to respond to customer inquiries, or to refer customers to professional assistance when needed.

Historical Publications

The DoxSys/Lason team is developing an image database consisting of the five premier climate publications in the United States: *Climatological Data*; *Hourly Precipitation Data*; *Storm Data*; *Monthly Climatic Data for the World*; and *Local Climatological Data*. Imaging of the *Climatological Data* publication, which consists of publications from the 1890's to the 1990's, has been completed. More than 800,000 paper documents were scanned and the digital images have been received at NCDC. Scanning has begun on the *Hourly Precipitation Data*, with all five publications expected to be completed by spring 2001. All image files are being sent to NCDC for quality assessment and eventual loading on the WSSRD on-line web system.

The DoxSys/Lason team is developing an image database consisting of the five premier climate publications in the United States

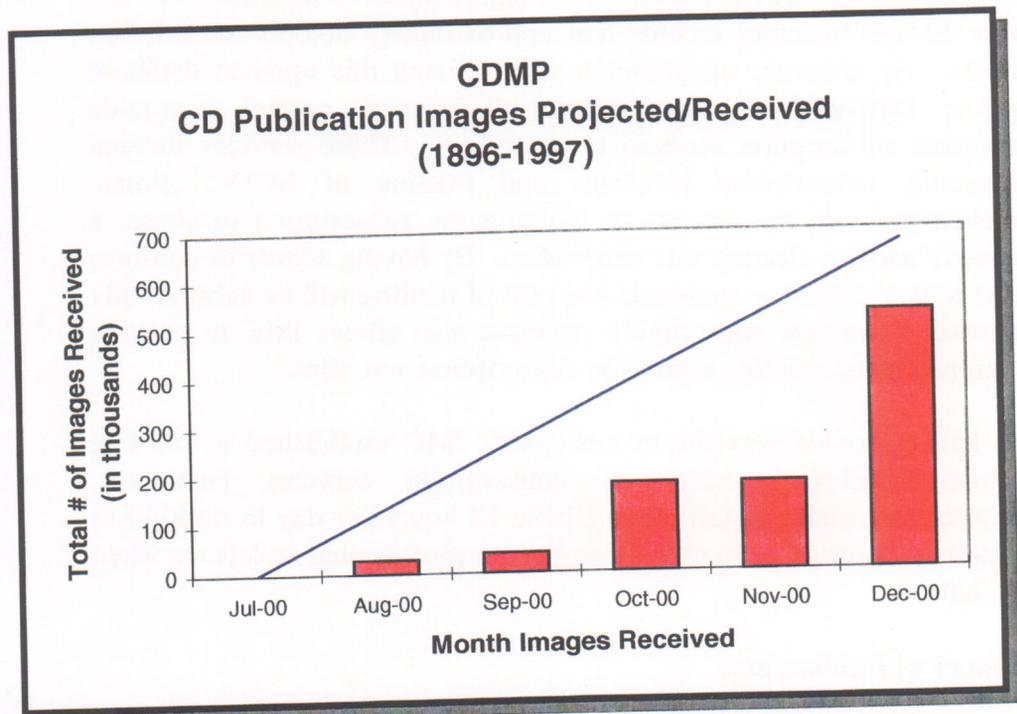
The team is also developing a production system to create a climate database consisting of the unique set of observations as found in the

(Continued on page 8)

(Continued from page 7)

Climate Records Books. The records document, in great detail, the climatological history of the United States since the Civil War. Digital images have been produced from approximately 400 rolls of microfilm. Keying of selected information will be done from these images. It is estimated that when completed, over 500 million characters will be keyed from 150,000 pages. The images and digital data will be provided to NCDC for placement on WSSRD and ingested into NCDC's archival system.

Imaging of the Climatological Data publication, which consists of publications from the 1890's to the 1990's, has been completed



Development

Station History

During the year, IMC worked to design a new, comprehensive metadata database. This newly designed Oracle database will greatly improve access to station history records currently stored in NCDC's station history processing system. Implementation is tentatively planned for early in 2001. This new station history system will enable NCDC to digitally ingest station history updates utilizing the National Weather Service's new Cooperative Station Service Accountability (CSSA) system. It will eliminate the costly, time-consuming, and inherently error-prone manual input of station history information from paper forms.

The newly designed Oracle database will greatly improve access to station history records

Observational Practices

With the new accessibility to historical records comes the need to understand the “rules” of observing as they have grown from infancy in the late 1800’s to today’s sophisticated, automated observing systems. CDMP research done during the year has resulted in creation of a database consisting of Federal Meteorological Handbooks and other observational practice “circulars,” as they were called at the turn of the century. An inventory of these handbooks and circulars was compiled. These publications were found in a variety of libraries, in state climate offices, and in the archives of NCDC. Handbooks are now in the process of being digitally imaged by IMC and placed on-line, greatly increasing access to this essential information.

A database consisting of Federal Meteorological Handbooks and circulars is being created

Summary

The first year of the CDMP provided many challenges, but resulted in enormous benefits. CDMP endeavors provided employment within the private sector, while significantly increasing the access and utilization of NOAA’s climate databases. Over .5 billion characters were keyed from over 9,000 microfiche, and more than 4 million digital images of climatological data were produced by CDMP contractors during the first year of this program. These data have been used to update existing climate databases, greatly advancing the modernization of the climatological data archives. This modernization effort strengthened the access and utilization of the vast store of climate information that, in the past, was accessible to only a limited number of individuals.

More than 4 million digital images of climatological data were produced

Challenges for the Future

Although access and utilization of the extensive holdings of climate data has greatly improved over the past year due to the CDMP, there is still a vast collection of climatological information that is inaccessible to most researchers, academia, and the public. Challenges for the future include continuing the efforts to increasing climate databases for the period before the Civil War. In addition, the detailed, multi-parameter observations taken to support aviation from its inception need to be imaged and digitized. Upper air observations from the dust bowl years must be made accessible so research can be conducted on this devastating period of U.S. history.

There is still a vast collection of climatological information that is inaccessible to most researchers, academia, and the public

Across NOAA, a monumental amount of other climatological data awaits modernization efforts. Charts detailing shoreline changes, i.e., sea level changes, must be digitized. Oceanographic and geophysical databases, such as the “climate” of the ocean and the solar-geophysical environment, are also currently inaccessible and thus in need of CDMP attention.