

## **CLIMATE DATABASE MODERNIZATION PROGRAM**

### **Making Climate Data Accessible**

The Climate Database Modernization Program (CDMP) was established by the U.S. Congress in fiscal year 2000. The goal of the program is to make major climate databases available via the World Wide Web thus increasing the utilization of the climate information. The program also has as its goal the creating of information technology jobs in several economically depressed areas of the country. Therefore, the vast majority of the CDMP work is being accomplished by several private sector contractors.

There are two basic methods of improving the accessibility of climate information. One is to convert paper or microforms to digital images, i.e., an image of the original record that can be displayed on a computer screen. The second method is to actually key or digitize the observations contained on the original forms, thus making the observations part of a larger database that can be manipulated to produce statistics or other climate information.

The National Climatic Data Center (NCDC) has chosen several major data sets to concentrate on first. The largest data set is the Surface Weather Observations which contains hourly data from about 1900 to the present. It consists of over 35 million paper records of which 25 million have now been imaged, indexed, and loaded to a web accessible system called Web Store Search Retrieve Display (WSSRD).

NCDC has also begun imaging the Cooperative Observations which consist of paper records of monthly forms which begin in 1895. To date the forms from 1995 to present have been imaged and loaded to WSSRD. In the coming year the period 1980-1994 will be completed and then in the following years the remainder of the forms will be imaged.

Not only are the forms being imaged but the daily data from 1895-1948 are being keyed. Some 171,850 station months have been keyed and integrated into the existing database of daily data. The earliest data begin in 1852 but the bulk of the stations were not reporting until the early 1890's.

Pushing the climatological record back further in time with more climate parameters are the Climate Record Books. These data, beginning in 1860, were originally stored on microfilm. The CDMP has created digital images from the film and has keyed the values as well. Over 60 different data summaries were keyed.

To be able to use and interpret these newly digitized data it became necessary to make station histories available. This data set contains information on station location, instruments, and observing schedules. The first station network to be addressed was the Cooperative Observation network and its station history. Over 600,000 pages of many different sizes, weights (onion skin to cardboard to photographs) and colors of ink are being imaged.

A unique data set being digitized, is the National Shoreline Maps. These historical maps present the ever-changing coastal areas. The large paper maps were first imaged but to make them suitable for use within a Geographical Information System (GIS) the maps are being "vectorized". Using curve following techniques, the actual shoreline boundaries are being converted to vectors which are used in GIS mapping systems.