

## MRCC Cooperative Observer Station Annual Temperature Record Data Set

### *Description*

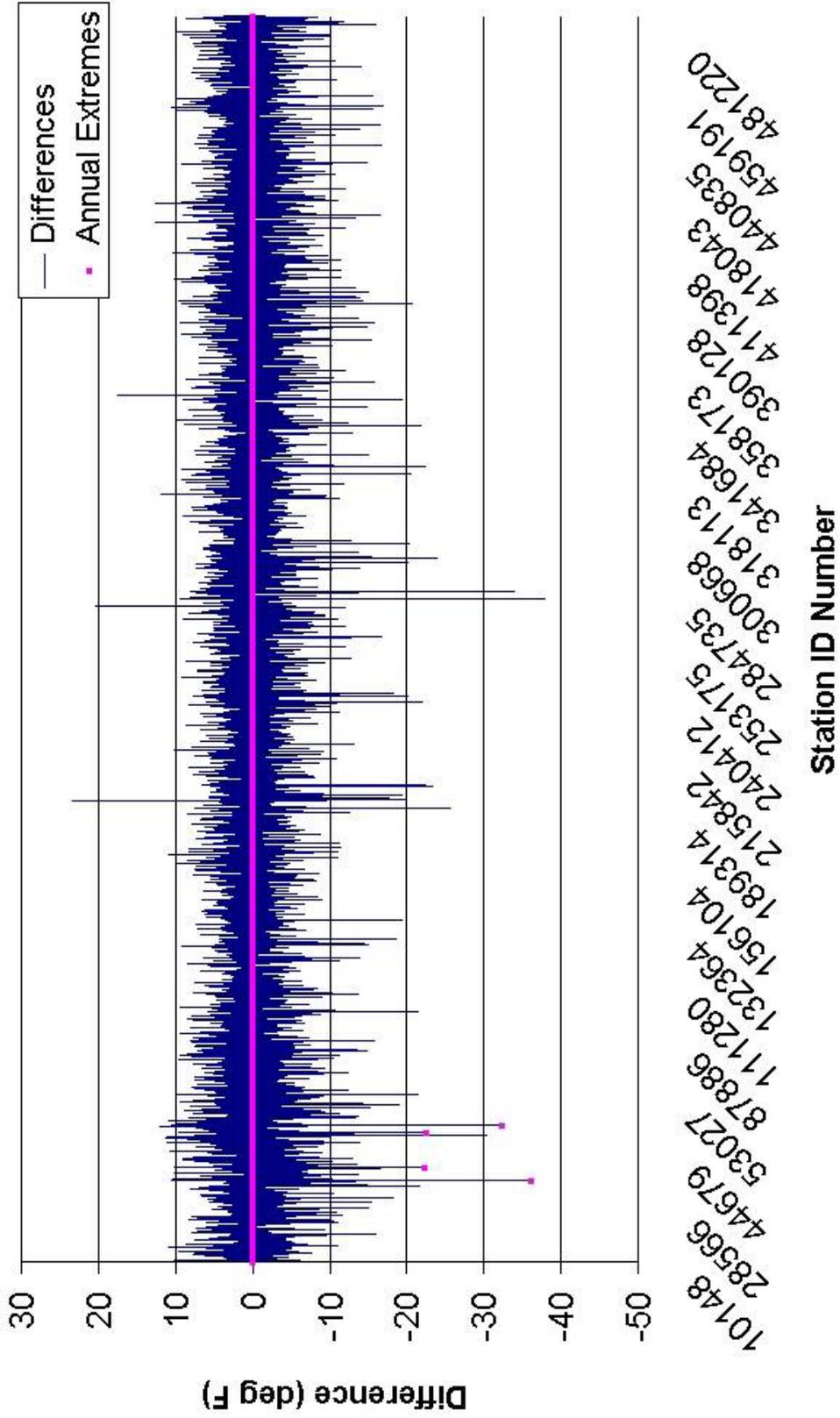
This data set contains the annual maximum and minimum temperature records for stations in the conterminous U.S. that are archived in NCDC Data Sets DSI-3200, 3205, and 3206 (these data sets are described at the end of the document). The version of the DSI-3206 used here includes the additional quality control of extreme outliers performed by Kunkel et al. under the auspices of a NOAA Office of Global Programs grant. Each station with a unique cooperative observer system i.d. number (coop number) and a running length of greater than 5 years was examined. After the absolute maximum and minimum of the time series were determined, additional quality control was applied. Those records that initially exceeded the established NCDC state records were examined first. In every case, extreme records were first compared to images of paper records available on the Web Search Store Retrieve Display (WSSRD) system at Information Manufacturing Corporation (IMC), and NCDC contractor. If the values were correctly digitized, and were in general agreement with surrounding stations, the values were accepted. If the particular station record image was not available as a cooperative observer form or in an image of a Climatological Data (CD) publication, the location of the target station was established. Nearby stations listed in the CD were examined to establish if the station record was plausible. If the electronic record was incorrect, or implausible, it was removed from the list. At this time, overlaps of records between NCDC data sets that individually contained the same coop records were also resolved. In completing this stage of the quality control, it was found that none of the maximum and minimum temperatures exceeding state records were valid, and they were set to missing and the process re-run.

To further quality control the results, the difference between a station extreme minimum for a month and the station mean minimum for that month was examined graphically. These differences were made comparable between states by subtracting the state average of this metric. For the record minimum temperatures, this monthly relationship is given as:

$$\text{Difference} = \left( \begin{array}{c} \text{Station} \\ \text{Extreme} \\ \text{Minimum} \end{array} - \begin{array}{c} \text{Station} \\ \text{Mean} \\ \text{Minimum} \end{array} \right) - \text{State Average} \left( \begin{array}{c} \text{Station} \\ \text{Extreme} \\ \text{Minimum} \end{array} - \begin{array}{c} \text{Station} \\ \text{Mean} \\ \text{Minimum} \end{array} \right)$$

The resulting differences for all stations could then be plotted on a bar chart for each month. The annual record values falling in each month could then be plotted as a dot at the end of the appropriate bar, and visually inspected. This method allowed the identification of further cases of annual maximum and minimum temperature records that had a magnitude not corresponding well to records of surrounding stations in the same state. In this case, however, a moderate percentage of the annual records were found to be valid, with the error detection rate greatest in the shoulder seasons between winter and summer. Errors were also likely when the data for the warm season indicated a record minimum. The figure on the following page displays the difference data for approximately 8900 stations for the month of June. The four points marked with squares were selected as the annual minimum record for those stations. All four were bad data points, and were removed from consideration. Due to resource limitation for conducting this project, only annual minimum and maximum records are provided for users at this time.

# Minimum Record Differences for June



## Caveats

These station records were created by the MRCC at the request of a user. Since they were needed quickly, only a minimum amount of quality control has been possible at this point. Here are some potential problems with these data.

1. Stations with more than five years of data are included in this file. However, the values of all-time records for stations with brief histories are limited in accuracy and may vary from nearby stations with longer records.
2. While these data have been through NCDC quality control and NOAA OGP Kunkel et al. project quality control for extreme outliers, there is still a need for more resources to quality control extremes. Some additional objective and subjective quality control techniques were applied to the records in this data set. However, some of the difference graphics indicate a substantial number of suspect annual maximum and minimum temperature records. Time and resources do not allow further examination presently.
3. These are records strictly for single stations in the cooperative observer network, and are limited to the time of operation of each station under one coop number. The records for a place may need to be constructed from several individual station histories. In the case of NWS records for locations, these often consist of multiple stations and additional non-digitized data sources such as record books. Therefore, the records in this resource may not correspond to NWS records.
4. Some cooperative observer stations have missing place names in this list; there are some metadata gaps in the DSI-3206 dataset.

## Format

The MRCC Cooperative Observer Station Annual Temperature Record Data Set is provided as a text file in the following format:

Columns	Variable
1-6	Cooperative station identification number (integer)
7-38	Cooperative observer station name (characters)
39-43	Beginning year of station record (integer)
44-48	Ending year of station record (integer)
49-54	Percentage of record missing (floating point 6.1)
55-60	Annual daily maximum temperature record in °F (integer)
61-71	Date of record in MM-DD-YYYY format (characters)
72-77	Annual daily minimum temperature record in °F (integer)
78-88	Date of record in MM-DD-YYYY format (characters)
89-96	Annual average of the daily maximum temperature (floating point 8.1)
97-104	Annual average of the daily minimum temperature (floating point 8.1)

## Data Sets

All three data sets used in this study consist of daily measurements taken by individuals or institutions, the majority of whom are volunteers. Maximum and minimum temperature measurements are taken at a standard height with standard instruments maintained by the National Weather Service, so that results can be compared nation-wide. Observers enter their measurements on monthly forms which are submitted to the local NWS office, and then forwarded to NCDC. The data on the forms are digitized by a contractor for NCDC. A subset of the observers submits these data electronically in near-real time, for rapid access. The data undergo several layers of quality control as they are processed. The data used in this study extend to the end of 2003, and so incorporate final quality-controlled NCDC data. A complete description of the network can be found at the National Weather Service Web Site at: <http://www.nws.noaa.gov/om/coop/index.htm>

DSI-3200 - Most of the records in this dataset encompass the period of 1948 to present. The systematic digitization of daily data began in 1948, when data were transferred to punch cards. In recent decades, selected high quality stations were digitized from archived cooperative observer forms or U.S. Weather Bureau Climatological Data publications for various research purposes, and included in the DSI-3200 data set. For more details, please consult: <http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td3200.pdf>

DSI-3205 - One of the largest archival digitization projects was performed by the MRCC during the 1990s. Most of the archival Climatological Data publications for the 1896-1947 period were digitized through projects at the State Climate Offices of the 9 states in the Midwestern Region (IL, IN, IA, KY, MI, MN, MO, OH, WI) and New Mexico. These data have been quality controlled and used for a variety of research projects. For more details, please consult: <http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td3205.pdf>

DSI - 3206 - In 1998, Congress made available economic development money for several states bordering the Appalachian Mountains. Some of these resources were directed towards a massive archival digitization project managed by NCDC. This project started by digitizing all available microfilm and paper records for cooperative observer stations from the mid-1890s to 1948. These data are contained in the 3206 data set, and, except for some modest overlap with both the 3200 and 3205 records, are unique additions to the digitized daily data holdings of NCDC. For more details, please consult: <http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td3206.pdf>