1993, Preprints of the Eighth Conference on Applied Climatology, American Meteorological Society, Anaheim, CA, 264-269.

6.3

HISTORICAL DAILY CLIMATIC DATA FOR THE UNITED STATES

David A. Robinson Rutgers University New Brunswick, NJ 08903

1. INTRODUCTION

The need to have climatic data of a high quality has received increasing attention in recent years from scientists, United States federal agencies (i.e., Interagency Working Group on Data Management for Global Change) and international committees (i.e., International Geosphere/Biosphere Program Management Working Group) (Robinson, 1990). For most climatic purposes, especially in climatic change studies, these data should preferably be gathered for multiple decades at observing stations having relatively stable local surroundings. It is also best to have data from as many stations as possible, with a variety of climatic variables being observed. Optimally, these data should be available to the research community in a digital format on a daily basis once quality control measures have been applied.

Observations from United States first order weather observing sites meet the multiple variable criterion, but are only several hundred in number. Most significantly, station surroundings are unstable over this century due to moves from city to airport locations and urbanization trends at both locations. U.S. cooperative climate stations observing temperature and precipitation currently number over 5000, and many of these have been in operation in mostly nonurban locations from the early part of this century. While cooperative stations are the best for gaining a representative picture of the nation's climate, they are not without limitations. Observers, observation times, and instrumentation have changed over the years, and the surroundings of most stations have not remained completely stable. Having a large number of stations to study helps to minimize biases that may result from these changes. Most daily observations, cooperative and first order, are not available in digital form before 1948, and until now most of those that are

have not been subject to rigorous quality control.

In recognition of the need for lengthy and accurate daily observations of climatic variables throughout the United States, a unique digital set of data, the Historical Daily Climatic Data Set (HDCD) has been assembled. Data from approximately 1000 principally cooperative stations are included in the set. This paper will provide an overview of HDCD particulars and development, including quality control procedures. Examples of studies in progress that are dependent on the availability of the daily data are also presented.

2. HDCD PARTICULARS

The HDCD includes observations of precipitation, snowfall, snow on the ground and maximum and minimum temperature. Records extend from 1988 back to the turn of the century for about 350 of the 1000 stations and to the 1910s, 20s and 30s for most others. The quality of all data (each variable for each day of record) has been examined (to be discussed later) and results of the query incorporated into the set. The HDCD contains station observations from every contiguous U.S. state, with the exception of Delaware. The number of stations for a state range from one for Rhode Island to 72 for Texas. incorporated into this set were originally digitized in a piece-meal fashion at the National Climatic Data Center (NCDC) and elsewhere. The mosaic of station data available at the start of this project lacked observations from before the 1920s at stations east of the Appalachians and west of the Rockies. This has been remedied by the additional keying of approximately 1500 station years of data from some 50 stations in these regions. Daily data for approximately 33 stations keyed as part of NCDC's Historical Climate Network Daily (HCN/D) effort (Hughes et al., 1992) have also been incorporated in the set. The result is a spatially diverse network of several hundred stations from coast to coast with virtually temporally complete records extending back to the turn of the century. These stations are supplemented by numerous other stations with useful, yet less complete, observations for all or at least 50 years of the century to date.

HDCD DEVELOPMENT AND QUALITY CONTROL

The HDCD began as an effort to assemble a daily snow data base (Robinson, 1988). However, upon initial inspection it was apparent that extensive quality controlling of the daily snow data was required before their utilization in any climatic investigations, and for this it was best to also have coincident temperature and liquid (or liquid equivalent) precipitation data. Therefore, a routine was developed to assess the quality of all of these variables for each day of record. This involved checking for outliers by examining multiple variables from the day in question (today) and the previous day (yesterday) along with statespecific thresholds. Examples from the query

routine include identifying reports where: 1) today's minimum temperature exceeded yesterday's maximum, 2) precipitation exceeded the record 24 hour total for the state, 3) snowfall was reported, yet no liquid equivalent precipitation was reported, and 4) considerable snow fell yet the ground was reported to be snowfree.

An example of the latter error is shown in figure 1, where for most of the early decades of this century the observers at Grand Island, NE failed to record snow on the ground, and when the station records were keyed these missing observations were eroneously keyed as zeros. In total, 36 queries were posed to each day's data. Each datum in the HDCD is flagged based on the outcome of the query and this information is stored with the datum. The station files are organized so that more detailed evaluations (i.e., comparing observations from multiple stations) may be conducted when selecting data for a specific study.

No data from an HCDC station were comitted from the set unless the suspicious data were hand-checked against original records andfound to be in obvious error (i.e., keyed incorrectly or a clear error on the observer's

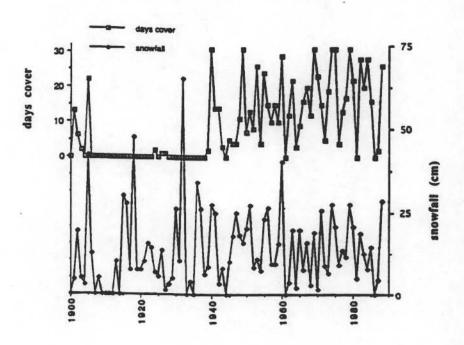


Figure 1. Grand Island, NE snowfall and number of days with ≥2.5 cm of snow cover for Januaries from 1901-1988 as found in the raw digital station file.

part). This permits those using the set to be as open or restrictive as they wish regarding the quality of the data they choose to use in their particular study. Also, it must be noted that some of the data flagged as suspicious are actually accurate observations. Hand checking was conducted on approximately 250 of the most temporally complete stations, and only keying errors were corrected. No attempt was made to correct other suspicious data, other than in some cases to mark them as missing. Also, no gaps in a station record were filled with estimated values. However, to create temporally complete files, gaps were keyed as missing on a daily basis.

Station histories are of considerable use in assessing the quality of a station's observations. Such metadata are available in hardcopy format at NCDC for all HDCD stations. The best histories are available for the approximately 650 stations in the HCDC that are also part of NCDC's Historic Climate Network (HCN) (Karl et al., 1990). Table 1 provides an example of the changes in observer, site location and observation time for the cooperative station at Napoleon, ND. While numerous changes have occurred, they have been relatively minor and with one exception do not appear to have influenced the secular continuity of the Napoleon records. The exception has to do with the observation of shallow snow depths. Charles Hoof tended to quickly report the loss of snow cover once depths fell below 7.5 cm (3 inches), while other observers reported snow at shallow depths for longer periods. This information is vital when assessing potential climatic change, particularly when looking at a time series for only one station.

4. EARLY USE OF THE HDCD

Several ongoing studies are taking advantage of the lengthy and daily nature of the HDCD. One of these is comparing daily means of temperature from different intervals over the past century to examine differences that might be less apparent or even masked by only using monthly averages. An example is shown in figure 2, where Grand Island, NE maximum daily temperatures for 1930-39 are compared to those for the entire 1900-1988 period. The anomalous warmth of the 1930s is apparent throughout the year, and is seen to peak from

Table 1. Napoleon, ND cooperative climate station history. Included are listings of; a) observers, b) station moves (D&D=distance and direction, Δ elev=change in elevation) and c) time of day of measurement. Note that the station began recording data in 1889, while digitized records only extend back to 1901. (1 ft = 0.30 m, 1 mi = 1.6 km)

a) OBSERVERS

1889-1899	Julius Hoof
1899-1944	Charles Hoof
1944-1945	Gladys Peterson
1945-1946	Joseph Hoof
1946-1954	Gladys Peterson
1954-1956	Ted Frank
1956-1957	Alvin Schuchard
1957	Warren Wentz
1957-1968	Gladys Peterson
1968-1973	Warren Wentz
1973-1974	Terry Wentz
1974-present	Warren Wentz

b) STATION MOVES

Year	D&D	△ elev
1942	30 ft S	none
1948	100 ft?	none?
1949	70 ft SE	none
1954	1.5 mi NW	incr 1 ft.
1957	4 blocks SE	decr 1 ft.
1958	0.6 mi W	none
1969	0.5 mi E	incr 9 ft.
1976	0.4 mi NE	incr 16 ft.
1985	400 ft SW	same

c) TIME OF DAY OF MEASUREMENT

1889-1899	twice/day, times?
1899-May 1911	1800
Jun 1911-Aug 14, 1918-	0700
Aug 15, 1918-Jul 1919	evening
Aug 1919-Sep 1919	0700
Oct 1919-Apr 1921	sunset
May 1921-Jun 1921	0700
Jul 1921	evening
Aug 1921-Feb 1923	0700
Mar 1923	1900
Apr 1923-Oct 23, 1942	0700
Oct 24, 1942-Nov 1942	1900
Dec 1942-May 1943	sunset
Jun 1943-Jul 1943	1900
Aug 1943-Jan 1944	sunset
Feb 1944-May 6, 1970	1730-1830
May 7, 1970-present	0700-0800

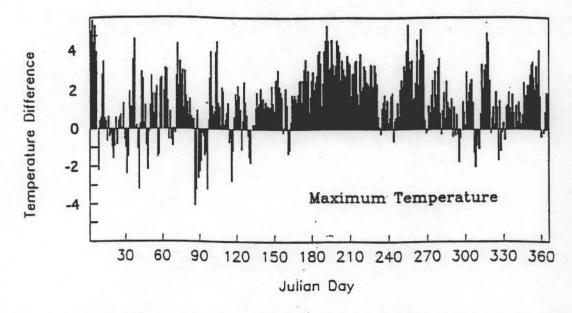


Figure 2. Mean daily maximum temperatures (°C) at Grand Island, NE for the 1930-39 period minus the mean maximum temperatures for the entire 1900-88 interval.

late June to late August, with other maxima in late September and in the late fall and early winter. Anomalously cold maximum temperatures were rare in the 1930s and mostly scattered from mid winter to mid spring.

No monthly or seasonal reports of snow cover frequency are available from NCDC or elsewhere, thus daily observations are required to compile any type of temporal record. Two studies are underway using the HDCD to examine snow cover across the nation. One uses daily snow depth and temperature data to generate regional algorithms that provide information on snowpack settling or melting as a function of temperature (Hughes and Robinson, 1992). For instance, figure 3 shows mean changes in the depth of snow on a winter day for 0.56°C increments of mean temperature for four states on the Great Plains. These values were generated using approximately 180 cases (days with snow on the ground) for each temperature increment. Input was drawn from all stations within a given state.

Daily decreases in depth are larger as temperatures increase, and exceed 2 cm/day when mean temperatures are near or above freezing. Daily changes are smaller at a given temperature as one continues northward in the Plains, probably due to differences in the availability of solar radiation. Winter values were found to be approximately 1.5 times lower than spring values at a given temperature.

Using these algorithms in conjunction with snowfall data, snow depths on days with missing observations have been estimated with considerable success. This is a particularly useful methodology, as quite often snow cover is the only variable failing to be consistently reported at a station (Robinson, 1989).

Another snow study is examining time series of cover over the past century to gain insight into the natural variability of snow. Only stations with little missing data are being used, and such gaps are filled where possible (snowfall and temperature data must be available) using regional models such as described above. An example shows the number of days with ≥7.5 cm of snow on the ground between September and the following May at Fairbury, NE for the winters of 1899/1900 through 1986/87 (figure 4). Only seven years during this period had insufficient data, thus prohibiting a seasonal summation. The remaining years show considerable year to year variability in snow cover duration. The 1920s and 1930s had a number of years with

infrequent cover, while years in the late 1960s to middle 1980s frequently had extended cover.

CONCLUSIONS

The Historical Daily Climate Data Set is beginning to be employed in several studies that take advantage of its unique features. These attributes include the half to full century length of digital data from about 1000 stations across the contiguous U.S., the daily frequency of the data, the multivariate nature of the set. especially the snowfall and snow cover data. and the rigorous quality control that has been applied to the set. Experience gained in the generation of the HDCD is or will soon be applied to the organization and inspection of climatic data from the Commonwealth of Independent States, China and Canada. Also, an effort by a number of state climate offices to key daily data from a number of currently unavailable stations having 100 years or more of virtually continuous observations is being coordinated and compiled by NCDC (NCDC, 1989) and these data will be added to the HDCD. In addition, a NCDC project has begun that will create a nearly error-free digital set of daily temperature and precipitation for U.S. cooperative stations, most of which have been

digitized from 1948 onwards (Reek et al., 1992). All of these efforts will contribute greatly to present and future studies of global climatic change

Acknowledgments: Many thanks to those at NCDC who in some measure assisted in the generation of this data set, including director K. Hadeen, S. Doty, G. Goodge, P. Hughes T. Karl, E. Mason, P. Steurer, and especially R. Ezell, and to students at the University of North Carolina - Asheville, including W. Capehart, E. Ezell, T. Speece and especially R. Henry. Thanks also to T. Baker at the Lamont-Doherty Geological Observatory of Columbia University and to A. Frei, M. Hughes and J. Wright at Rutgers University for contributing to the project, and to D. Leathers at the University of Delaware and K. Dewey at the University of Nebraska for their cooperation on the mean daily temperature study. This work has been supported by the NSF Office of Climate Dynamics under grants ATM87-19865, ATM88-22803 & ATM89-96113, the NSF Geography and Regional Science Program under grant SES-9011869, and NOAA under grant NA90AA-D-AC518.

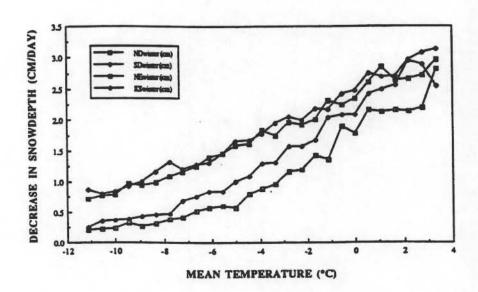


Figure 3. Daily snow depth change at selected mean daily temperatures during winter (Nov - mid Feb) in North Dakota, South Dakota, Nebraska and Kansas.

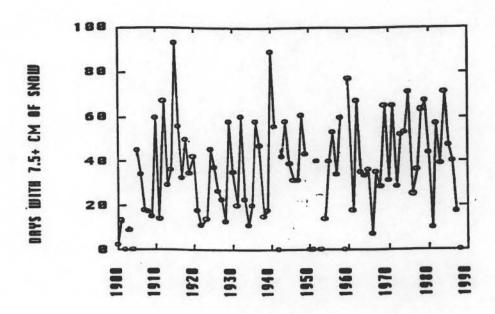


Figure 4. Days with ≥7.5 cm of snow cover at Fairbury, NE during the winters (Sep - May) of 1899/1900 through 1986/87. Missing years plotted along the bottom axis, no year with data had a complete absence of snow cover days.

REFERENCES

Hughes, M.G., and D.A. Robinson, 1992: Creating temporally complete snow cover records using a new method for modeling snow depth changes. Snow Watch 1992 Proceedings, Niagara-on-the-Lake, Ontario. (in press)

Hughes, P.Y., E.H. Mason, T.R. Karl, and W. A. Brown, 1991: United States Historical Climatology Network Daily (HCN/D) temperature and precipitation data, ORNL/CDIAC-50, NDP-042, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN.

Kari, T.R., C.N. Williams, Jr., and F.T. Quinlan, 1990: United States Historical Climatology Network (HCN) serial temperature and precipitation data. ORNL/CDIAC-30, NDP-019/R1. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN. NCDC, 1989: Executive Briefing Material: United States Centennial Cooperative Weather Station Program 1890-1990, NOAA, NESDIS, 67 pp.

Reek, T., S.R. Doty and T.W. Owen, 1992: A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *Bull. Amer. Meteor. Soc.*, 73, 753-762.

Robinson, D.A., 1988: Construction of a United States historical snow data base. Proceedings 1988 Eastern Snow Conference, Lake Placid, NY, 50-59.

______, 1989: Evaluation of the collection, archiving and publication of daily snow data in the United States. *Phys. Geog.*, 10, 120-130.

______, 1990: The U.S. cooperative climate observing system: reflections and recommendations. *Bull. Amer. Meteor. Soc.*, 71, 826-831.