Status and Climate Applications of the 19th Century Forts and Volunteer Observer Database

Nancy E. Westcott[1], Jason Cooper[2], Karen Andsager, Leslie A. Stoecker[3], and Karsten Shein[1]

Affiliation:

1. Midwestern Regional Climate Center, Illinois State Water Survey, Prairie Research Institute, University of Illinois

2. National Centers for Environmental Information, National Oceanic and Atmospheric Administration

3. Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois

Corresponding Author: Nancy Westcott, nan@illinois.edu

ABSTRACT

The Climate Data Modernization Program Forts and Volunteer Observer Database (CDMP-Forts) currently consists of 450 keyed and 355 quality-controlled stations for the period 1788–1892, reaching across the United States. In conjunction with the Global Historical Climate Network (GHCN) daily data, this resource is invaluable for examining 19th century weather and climate in the United States. CDMP-Forts is incomplete, however, with a considerable amount of data remaining to be digitally transcribed and quality controlled. It is the intent of this paper to provide an overview of the processes involved in rescuing these data and to show important ways these data can be used and the considerations that may have to be taken to create meaningful analyses. Finally, the dataset is placed in the context of other global datasets and efforts to rescue historical weather data.

1. Introduction

The United States (U.S.) has an extensive history of recorded weather observations (Miller, 1931) beginning with diaries from the 1700s. The first organized network was established in 1820 when the U.S. Army surgeons began regular weather observations at military forts as part of the Surgeon General Network. By mid-century, Smithsonian Institution volunteers began taking observations as well. On 9 February 1870, the majority of official observations shifted to the U.S. Army Signal Services Division in the War Department. In 1890, the Weather Bureau was created within the U.S. Department of Agriculture and by 1893 had assumed management of observing network operations. Hopkins and Moran (2010) provide a detailed history of these networks, which evolved into today's National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) Cooperative Observer Program. Although automated weather networks have gained prominence in the past few decades, the manual observations from the late 19th century to the present are essential for reconstructing and analyzing the climate of the U.S. However, preserving and making these manual observations available requires substantial effort.

In the 1940s, the National Archives and Records Administration had the foresight to copy the handwritten paper observation forms onto microfilm. From 2000 to 2011, the National Climatic Data Center (now the National Centers for Environmental Information, NCEI) digitally imaged that microfilm via the Climate Database Modernization Program (CDMP). These images, plus newly found data, cover approximately 4800 stations and are officially archived at NCEI but are also available at the Midwestern Regional Climate Center (MRCC).

The CDMP-Forts and Volunteer Observer Database (referred to here as CDMP-Forts) was built as part of the overall CDMP program (Dupigny-Giroux et al., 2007). Data for 39 variables were digitally transcribed (keyed) from copies

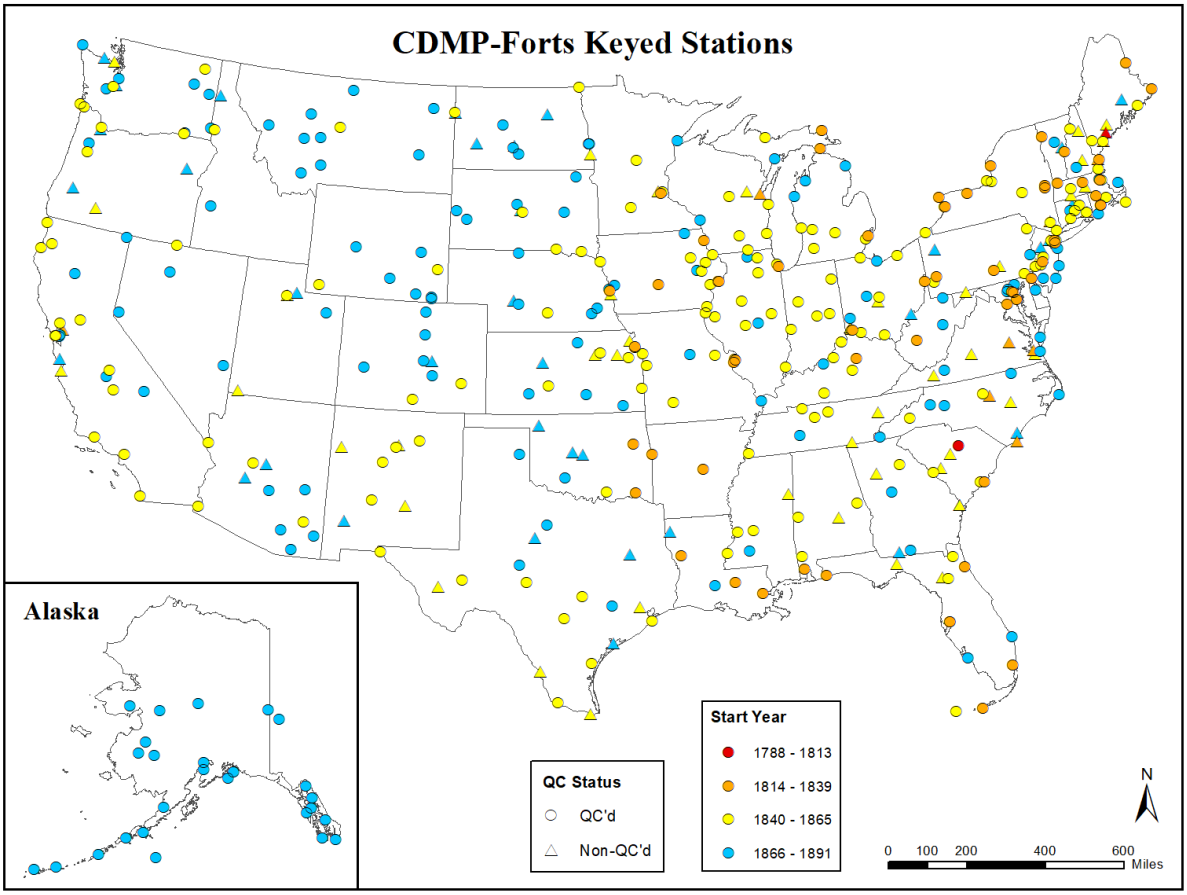


Figure 1: Map of keyed CDMP-Forts stations with shape denoting completion (circle) or not (triangle) of quality control and colors indicating station start year.

of original observer forms for 450 stations in the continental U.S. and Alaska (Fig. 1), and data for 355 of these stations were quality controlled (displayed as circles in Fig. 1). These stations were selected for their duration (10 years or more) and to ensure adequate geographic representation of all regions of the U.S. By extending station records backward in time and adding data in new locations, the dataset, while not complete, has a great potential to advance understanding of historical climate behavior at local, regional, national, and global scales and further document climate change and variability.

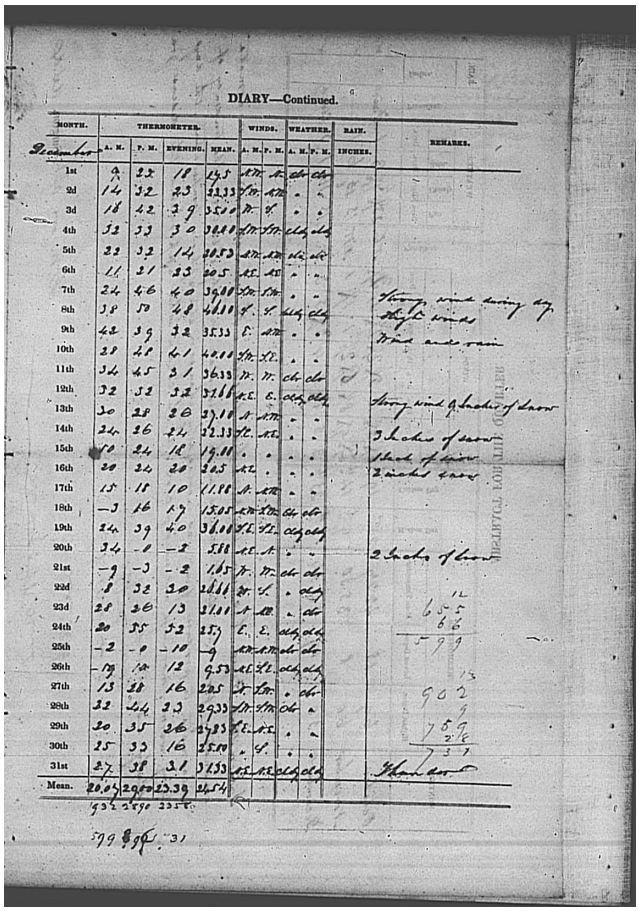
This article first provides an overview of the data and quality control (QC) procedures detailed in Westcott et al. (2011) and its strengths and some caveats regarding its use. Next, it highlights example applications of these data. Finally, it discusses the relationship of these data to the Global Historical Climate Network (GHCN) and other international datasets and data rescue efforts.

2. Data and Methodology

The processes involved in the keying of data and the QC tests briefly presented here are generally well known and straightforward, as are the caveats considered in the analyses. They are included to illustrate the lengthy procedures involved in moving from the keying of data to producing the analyses so that one is better prepared when embarking on either a new data rescue effort or before employing the data in a historical analysis.

a. Keyed Data

Modern-day NWS Cooperative Observer data consist of daily maximum and minimum temperature, precipitation, and snowfall from the preceding 24-hour period, and snow depth at the time of observation. However, like the Surface Airways Network and some automated stations of today, many of the CDMP-Forts stations recorded additional variables or at more frequent times. In all, 39 variables were keyed (see Table 1 in Westcott et al., 2011). Although data entry is a time-consuming task, and many applications involve only temperature and/or precipitation, the likelihood of rekeying the data is low and the value of including the additional variables is high. Beyond temperature, precipitation, and snow, keyed observed variables could include various measures of humidity, pressure, cloud and sky information, winds, and remarks (e.g., Fig. 2). These later


DIARY—Continued.

MONTH.	THERMOMETER. A. M.	P. M.	EVENING.	MEAN.	WINDS. A. M.	P. M.	WEATHER. A. M.	P. M.	RAIN. INCHES.	REMARKS.
December										
1st	9	22	18	14.5						
2d	14	32	25	23.33						
3d	18	42	39	33.00						
4th	32	33	30	31.66						
5th	22	32	14	22.66						
6th	11	21	23	18.5						
7th	24	46	40	36.66						
8th	38	50	48	44.00						
9th	42	39	32	37.33						
10th	29	48	41	39.33						
11th	34	45	31	36.66						
12th	32	32	32	32.00						
13th	30	28	26	28.00						
14th	24	36	24	28.00						
15th	18	24	18	20.00						
16th	20	32	20	24.00						
17th	15	18	10	14.33						
18th	−3	16	17	10.00						
19th	24	39	40	34.33						
20th	34	−0	−2	10.66						
21st	−9	−3	−2	−4.66						
22d	8	32	30	23.33						
23d	28	36	13	25.66						
24th	20	35	32	29.00						
25th	−2	−0	−10	−4.00						
26th	−19	12	12	1.66						
27th	13	28	16	19.00						
28th	32	44	33	36.33						
29th	30	35	26	30.33						
30th	25	33	16	24.66						
31st	27	38	31	32.00						
Mean.	20.03	29.90	23.39	24.44						

Figure 2: December 1836 U.S. Army Surgeon observer form from Fort Des Moines, IA, which is near modern-day Des Moines, IA.

variables can aid in validating daily temperature and precipitation observations and allow spatial analyses similar to those found on modern weather maps. Variables were keyed as they appeared on the original forms, were processed and digitally preserved, and are presented here in units used at the time (e.g., temperatures in degrees Fahrenheit, precipitation in inches, cloud cover in quarters).

Not all variables were present on each form, and reporting times often varied by station and observer. Temperatures were most commonly read three times daily and are referred to as at-hour observations. The times were in the morning (denoted as AM or 0700); early afternoon (denoted as PM, 1200, or 1400); and evening (denoted as EVENING, 1700, or 2100) Local Time. Observations were sometimes read at sunrise and sunset and are denoted as such. Rainfall and other variables also could be observed at one or more of these times. In the digital files, observation times are included in variable names (e.g., TEMP07 for temperature at 0700 LT). Maximum and minimum temperatures were not consistently recorded until about 1865, first at stations located at forts and then at volunteer stations. Based on a cursory review of the available variables by date and station, the transition to daily summary observations from at-hour temperatures to maximum and minimum temperatures was about 90 percent complete by 1895 (Westcott et al., 2011). Detailed information on variables and data formatting are outlined in Andsager et al. (2010).

As with modern-day observation networks, observation procedures changed over the period of record and even from observer to observer, and factors such as siting requirements and details and instrumentation type were often inadequately (or not) documented. Latitude, longitude, and elevation were not always provided on the forms, and when they were provided, they were not to the precision of today's GPS measurements.

The keyed station file can include nearby locations (within about 5 miles), as stations moved or were decommissioned and recommissioned over time. The station data and metadata files preserve all available station identification names, numbers, and locations. Extensively researched histories for 72 stations are available on the MRCC website (https://mrcc.purdue.edu/FORTS/histories1.jsp) and in the NCEI archives. These histories address station locations, moves, observers, instrumentation, instrument siting, and observing practices.

Keying values handwritten in 19th century script (e.g., Fig. 2) is challenging, especially as observer forms varied throughout the century, penmanship varied by observer, and microfiche/microfilm image quality varied (Truesdell et al., 2008). To reduce keying errors, data were double-keyed (i.e., keyed independently by two transcribers) into templates specifically designed for each of the various forms. In an extensive study, Barchard and Pace (2011) found some 30 times more errors were missed by visual inspection of single keying than by double keying data, although double keying could still result in some errors (<1%). Further, staff at NCEI (then NCDC) and the MRCC determined that use of a template with metadata specifications (e.g., station name, location, variables present by column, and date) reduced the number of metadata errors. Once double-keyed (and compared for agreement), the digital data were sent to the MRCC where both the keyed metadata and the keyed observed data were quality controlled.

b. Quality Control

The MRCC developed software to QC the keyed data (Westcott et al., 2011) and performed QC tests on 355 stations (see station list and status, http://mrcc.purdue.edu/FORTS/qc3.jsp). The purpose of the QC is to ensure that the station and variable data specified by the metadata keyers were in fact the data keyed, and that the keyed observations represented the intent of the observers. The QC centered on reviewing potential keying, instrument, observer, and metadata errors and not on the more refined analyses of data-series homogenization.

The QC tests flag suspicious values, some of which can be corrected automatically, and some that must be evaluated manually. The tests are run in sequence, moving from metadata checks to data checks, and from more general issues to

more specific readings. In a step-wise manner, the output of each QC test serves as input for the next test (Andsager et al., 2005). The metadata, temperature and precipitation data are rigorously evaluated. The other variables are subject only to gross error (range) tests.

i. Metadata Tests

As metadata specific to the station location are examined, in the context of keying, metadata tests also include issues related to the recording and transfer of data. Metadata errors could indicate formatting issues or systematic data quality issues. Initial tests include checking for 1) duplicate element types within each month for each station; 2) "gross errors" from mislabeled elements exceeding an extreme range of values or inappropriate units; 3) inconsistencies between the observation date of the keyed data and that identified in the metadata; and 4) trace precipitation amounts and precipitation amounts accumulated over several days for flagging.

A second group of metadata tests checks for known data issues that can be automatically corrected, or if not correctable, then flagged as an error. One such issue that was automatically resolved (set to missing) relates to wet bulb temperatures when the bulb is doused with too much water and the air temperature falls below freezing. As the water freezes, a crust is formed, insulating the bulb, and the latent heat of fusion released artificially heats the bulb. A last test examines the form type and determines if the observation time is in the appropriate time zone. This was necessary in part because Signal Service took observations at the standardized Washington Meridian (or Mean) Time (WMT) of 0735, 1435, and 2300 WMT from April 1873 to December 1892. This allowed weather maps to be produced for approximately the same moment in time across the continental U.S. (Conner, 2008). The old Naval Observatory in Washington, DC was selected as the standard for WMT. This is similar in concept to Coordinated Universal Time (UTC). The volunteer observers continued to report at 0700, 1400, and 2100 LT (Westcott, 2011; Conner, 2008).

ii. Data Tests

Suspicious values are flagged by various tests including 1) checks on daily values by comparing newly computed monthly total precipitation and mean temperature with those written on the original observation forms; 2) internal consistency checks comparing daily measurements of the same variable at different observation times or different measures of a related variable; and 3) extreme value checks of individual values compared with the standard deviation of calendar day values. Internal consistency tests for temperature would identify, for example, when at-hour temperatures were smaller than minimum temperatures, when minimum temperatures exceeded maximum temperatures, or when wet bulb temperatures were higher than the corresponding dry bulb temperatures.

Extreme values are often examined graphically (e.g., Kunkel et al., 2005). To identify and evaluate daily extremes, all years of daily data from a station are plotted on a single graph, with each Julian Day along the x-axis (e.g., Fig. 3 for maximum temperature), providing a visual climatology. The standard deviation can be computed (with three or more years) and plotted for each day. The graphs are inspected manually by climatologists to determine a standard

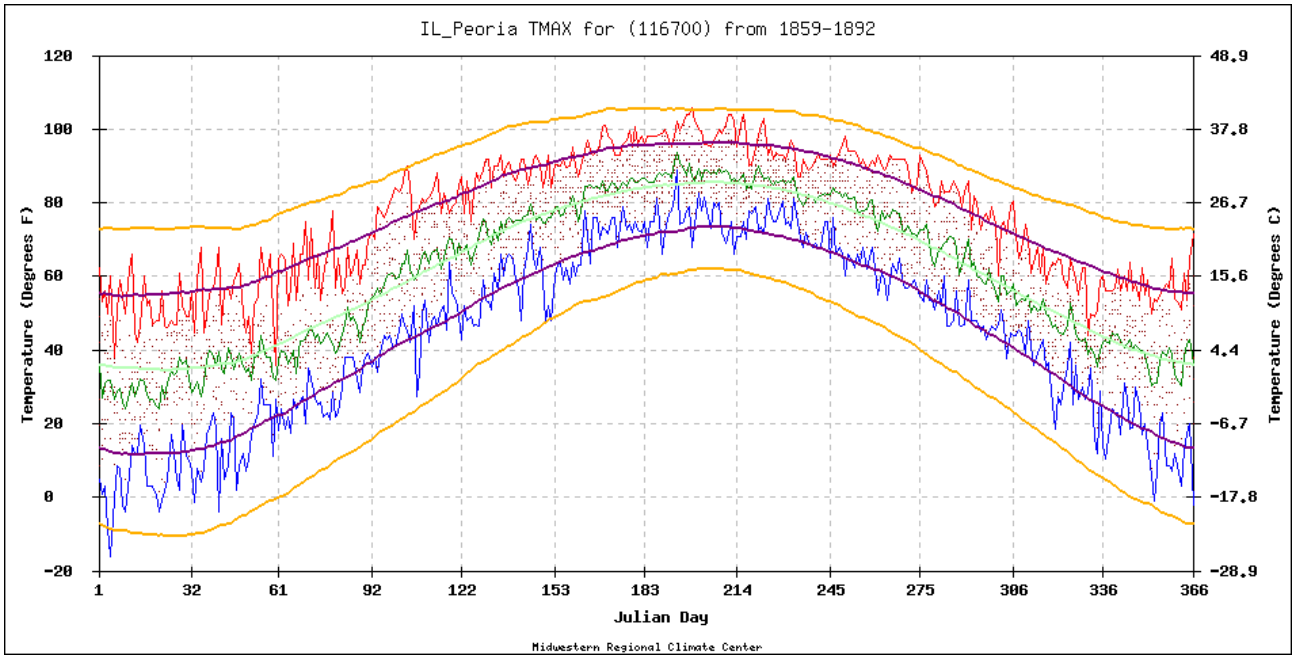


Figure 3: Graph used in quality control of extreme maximum temperature values in Peoria, Illinois. The red line indicates the highest values for each Julian Day, the blue line the lowest, the green line the mean, and the yellow lines are the standard deviation cutoffs. From Westcott et al., 2011.

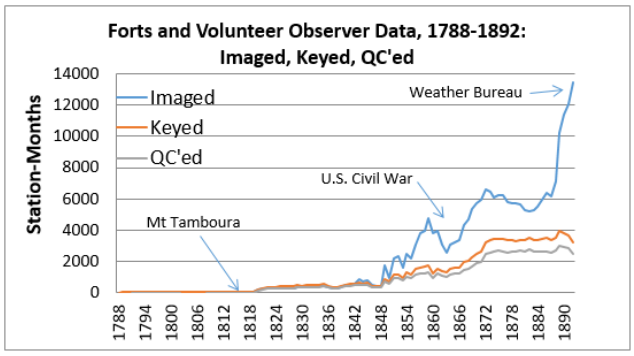


Figure 4: Annual number of station-months imaged, keyed, and those with metadata, temperature, and precipitation quality-controlled and available at the MRCC.

deviation limit that captures possible errant values. When daily values exceed the specified standard deviation limit, that value is flagged, and the scientist would check the value against the form image. Other graphical tests include large day-to-day differences using a spike test, systematic jumps (often monthly) in data values using a step test, and within-day differences using a diurnal range test. When standard deviation thresholds are used to flag possible errors, the allowable range of daily values is station and time-of-year dependent. Larger standard deviation thresholds are set for continental stations due to their wider variations in temperature than at coastal stations. Also, the range of values flagged is smaller in summer than in the transitional seasons. Generally, the threshold is around ±5 standard deviations for the spike test and ±4 standard deviations for the step and diurnal temperature range tests. Extreme individual daily precipitation values are also flagged as outliers and are usually 8 to 9 standard deviations from the mean daily value. As precipitation values typically follow a skewed statistical distribution characterized by a long tail, large standard deviation limits are reasonable.

iii. Flagged Value Verification

Common observer errors found in the keyed data include transposed numbers, missing or misplaced decimal points, missing digits (such as a zero), and missing values. Other errors include other observer errors (e.g., non-conventional units), instrument malfunctions, metadata problems (station ID or naming convention inconsistency), metadata errors (incorrect year or month), and unclear forms (faded ink, poor copy).

A flag was set to identify the reason for each error. Once the verification of the flagged error value is completed, the assessor selects the option that best describes the appropriate verification type. The value may verify as 1) correct, matching the form; 2) suspect, with no obvious reason for error, thus no change to data; 3) in error, with a replacement value specified; 4) in error, corrected to match the original form; 5) in error, set to missing; 6) non-data, from stray pen/pencil or imaging copy marks, deleted; or finally 7) data were missed by keyers, and scheduled to be keyed.

A preliminary evaluation of errors was provided in Westcott et al. (2011). The type and frequency of errors varied from station to station and observer to observer. The correctable errors typically resulted from unclear forms, miss-keyed data, or errors in the metadata for the image. Also, it was found that when the CDMP project introduced form-by-form instructions, the number of errors decreased for subsequent keyed stations.

c. Status of Keyed and Quality-Controlled Datasets

Since keying of the original 450 stations, no additional stations have been keyed, but the QC of Alpena, MI and Grand Rapids, MI has been completed. The number of completed station-months of keyed and quality-controlled data by year compared to the number of imaged station-months are presented in Fig. 4. From Figures 1 and 4, one can imagine the westward expansion of the U.S. forts and population in the 1800s and the effects of the Civil War on weather observations. Note in Figure 4, the initial increase in stations followed the Mount Tambora eruption on April 1815, which resulted in the "Year Without a Summer" in New England, Southeast Canada, and Europe (Ludlum, 1948; Hamilton, 1986). Although only about 10 percent of imaged stations have been keyed, those stations make up about 50 percent of the total number of available images. Thus, many of the stations remaining to be keyed are of a shorter duration. These short-duration stations can be quite valuable for verifying the validity of nearby stations or for in-filling of spatial or temporal gaps in the data record.

Figure 5 shows the number of station-months keyed by select variable categories: temperature, precipitation, pressure, winds, clouds and sky observations, and humidity. The quantity of data for most variables follows that of the total number of keyed station-months, with the exception of snow and relative humidity. The shift to daily maximum/minimum temperature observations from at-hour

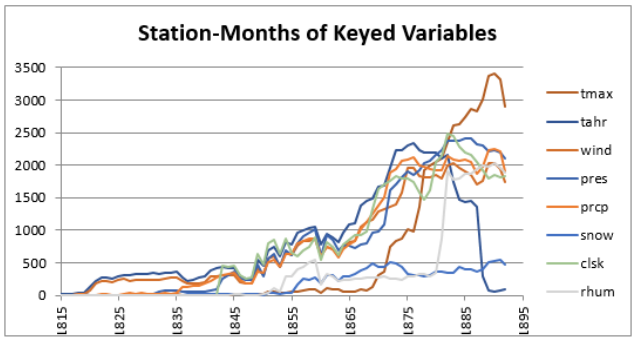


Figure 5: Annual number of station-months keyed for maximum and at-hour temperatures (tmax and tahr), wind (wind), pressure (pres), precipitation (prcp), snow (snow), clouds and sky cover (clsk), and relative humidity (rhum).

temperature observations is also apparent. The variety of variables could be quite useful in reconstructing weather patterns beyond what is available in the Daily Historical Weather Map series (https://library.noaa.gov/Collections/Digital-Collections/US-Daily-Weather-Maps), which began on 1 January 1871 and continues through today, and in validating the analyses provided by the Twentieth Century Reanalysis dataset (20CR, Compo et al., 2011).

3. Data Applications

Over the past 10–15 years, the MRCC and NCEI have distributed data to numerous users who have requested it, for example, to ensure accuracy in historical accounts and historical fiction, to commemorate local historical events, and to extend the period of record for long-running stations. In the following section, three specific applications of these CDMP-Forts data are presented to highlight both possible uses and data limitations to consider when performing such analyses.

Several datasets were used to augment the historical data provided by the CDMP-Forts database. In the first application of trends in heavy precipitation, daily ThreadEx data (http://threadex.rcc-acis.org/ ; e.g., Lai and Dzombak, 2019) for the Minneapolis, MN; Milwaukee, WI; and Chicago, IL areas were appended to the CDMP-Forts data after 1892. ThreadEx data often are used in climate service applications to represent large cities. During both the 19th and 20th centuries and even today, multiple station locations are possible. Generally, the longest reliable station is used in the CDMP-Forts and ThreadEx time series. Neither dataset is adjusted beyond the basic quality-control steps applied.

For the second application examining the 1877–18R8 El Niño event, 1981-2010 gridded normals from the PRISM (Parameter elevation Regression on Independent Slopes Model); e.g., Slinskey et al., 2019) dataset were used for comparison with the CDMP-Forts and GHCN (Menne et al., 2012) station data to compute monthly maximum temperature and total precipitation anomalies. The GHCN data will be further discussed in section 3b.

In the third application -- the case study of a flash freeze event in 1836 -- quality-controlled CDMP stations were used as well as some non-quality-controlled keyed stations and some stations with data read directly from handwritten images. Twenty-nine of the 38 stations used by Obrecht et al. (2012) were observations from the U.S. Army Surgeon Generals Network, and the remainder were from civilian volunteers.

a. Heavy Precipitation Event Climatology

One of the most common uses of historical data is the development of time-series for examining trends in temperature and/or precipitation for public interest applications (e.g., daily records extending back to the 1870s disseminated by broadcast meteorologists) and for research (e.g., Kunkel, et al., 1999; Burnette et al., 2010; Slonosky, 2015; Burnette, 2012; McKitrick and Christy, 2019; Lai and Dzombak, 2019). Here, example precipitation trends for three major cities in the upper Midwestern United States -- Minneapolis, MN; Milwaukee, WI; and Chicago, IL -- were constructed from 1838, 1851, and 1867, respectively, to 2019. Rainfall totals from the top-ten rainfall events per year are presented in Figure 6. The top ten events must be treated with some caution. While manual QC includes checking for and flagging accumulated (2–3 day) events, some observers did not note such events, and not every accumulated event may have been caught. Thus, the most extreme top-10 annual values were compared with total annual precipitation. Total annual precipitation, on the other

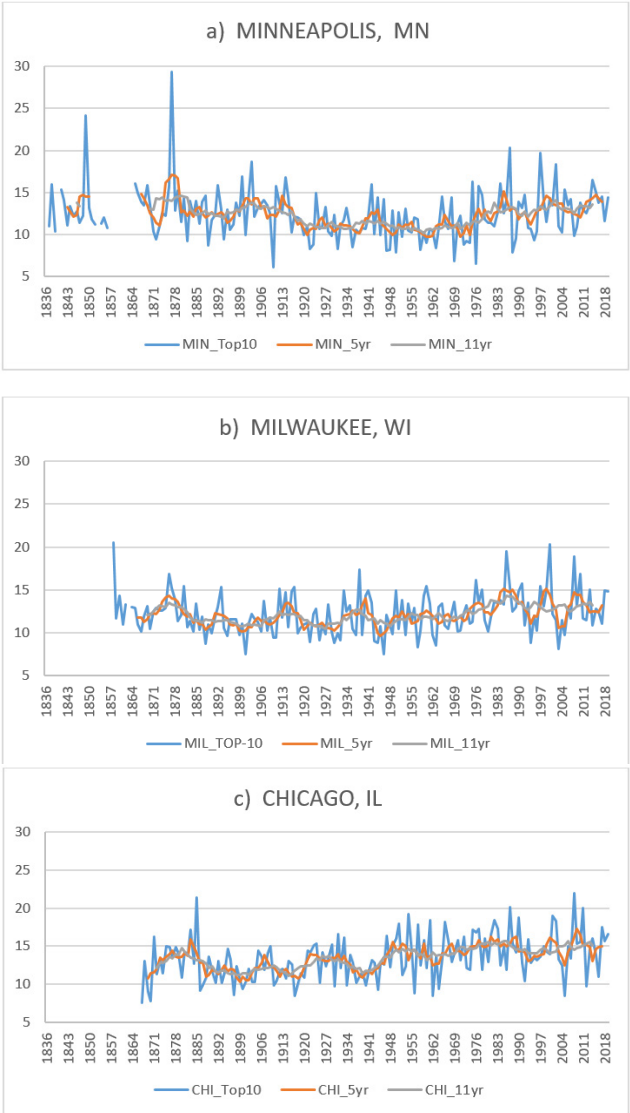


Figure 6: Precipitation (inches) from the CDMP-Forts and ThreadEX datasets totaled annually from the 10 largest events, along with the 5- and 11-year running means, for a) Minneapolis MN, b) Milwaukee, WI and c) Chicago, IL.

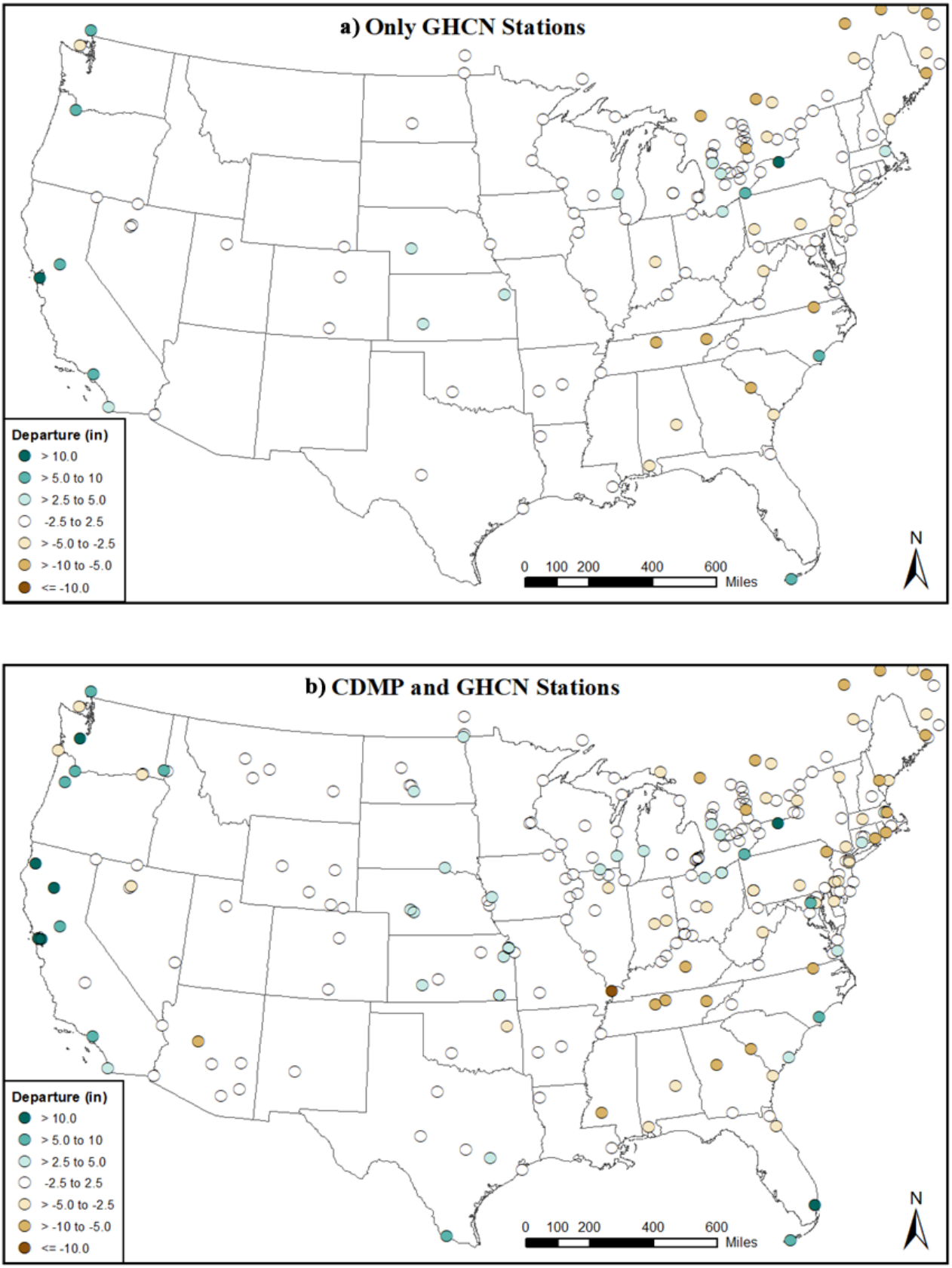


Figure 7: For the Dec 1877–March 1878 period, maps of the precipitation anomaly (inches) using a) GHCN stations alone and b) CDMP-Forts and GHCN stations.

hand, may be slightly underrepresented as some observers did not record precipitation amounts under 0.10 or 0.25 inches (Burnette and Stahle, 2012). Also, not every year includes all months, particularly during winter months with prolonged extreme weather. The annual precipitation totals for years with missing months are excluded from this analysis. Generally, only one or two months are missing from a year and make up about 10 to 15 percent of the station record. Years with more missing months often occur when stations are commissioned or decommissioned. Here it was found that the years with the largest top-ten events are also the years with high annual precipitation totals, providing some credence to their validity.

Plotting such series facilitates the identification of climate patterns of potential interest. For example, in Figure 6, a relatively dry period can be seen from about the 1920s through the 1960s in Minneapolis, from about 1885 to 1970 in Milwaukee, and from about 1885 to 1945 in Chicago. The increased rainfall in the past 50 years is apparent in all three cities. Additionally, very high rainfall amounts occurred in all three cities in the mid- to late 1800s, although not in the same decades. The Minneapolis station is a combination of several stations in the area, including Fort Snelling (Grice and Boulay, 2005). The large rainfall amounts occurring in 1849 and 1877 have been evaluated by the Minnesota Climate Office. In 1849, heavy rains were common from April through August at Fort Snelling (https://www.climatestations.com/minnesota-weather-for-1849/). In 1877, nearly half of the annual total precipitation fell in May and June. Annual precipitation and precipitation from the top-10 events were both high for Minneapolis in 1849 and 1877, for Milwaukee in 1858, and in 1885 for Chicago.

Unfortunately, time series are plagued by missing data during the early years of the network. Missing data result from station closures or moves and missing or unreadable forms. Investigating these further may be possible if other copies of the forms or other corroborating evidence is found. As in the case studies below and as evidenced in the Shein et al. (2013) evaluation of precipitation extremes, early observations are given more credence where accompanied by written accounts from other sources such as newspapers, diaries, or from nearby stations.

b. El Niño 1877-1878

The El Niño of 1877–1878 was a well-defined event comparable in strength to recent events (1982/1983, 1997/1998, and 2015/2016; Huang, 2020; Kiladis and Diaz, 1986). The 1877–1878 El Niño season was dubbed the "Year Without a Winter" in Minnesota (https://www.dnr.state.mn.us/climate/journal/1877_1878_winter.html). The largest monthly precipitation and maximum temperature anomalies occurred during the months of December 1877 to March 1878. The composite (Dec–Mar) four-month deviations of total precipitation and maximum temperature are presented in Figures 7 and 8. Considerable data exist across the U.S. to give a general view of this large-scale atmospheric phenomena (Fig. 7). Note the very high rainfall anomalies in locations on the West Coast and very high temperature anomalies in the northern and northeastern U.S. and southeastern Canada compared with the 1981–2010 normals (Fig. 8).

Figure 7a includes 133 GHCN precipitation stations and Figure 7b includes those plus an additional 119 CDMP-Forts stations. Figure 8a includes 110 daily GHCN stations with maximum temperatures and Figure 8b includes the GHCN stations and 58 additional CDMP-Forts stations. The extra CDMP-Forts stations approximately double the number of precipitation stations available and increase the number of available maximum temperature stations by about 30 percent. The fewer temperature stations compared to precipitation stations was due in part to many stations still practicing at-hour temperature observations during this 1877–1878 period.

For this case, 45 Canadian stations with precipitation values and 35 with maximum temperatures were obtained from the GHCN dataset. Also, some 70 stations are present in both the GHCN and the CDMP-Forts datasets for this case, although not all 70 stations have both precipitation and max/min temperatures (56 temperature and 59 precipitation stations overlap). When duplicate variables were present from both sources, the GHCN data were used for simplicity.

During this El Niño investigation, several actions were identified that would improve the quality and completeness of the 19th century database. First, it was recognized that the CDMP-Forts and the GHCN overlapping data have not been systematically compared. A preliminary examination of monthly values suggests that the overlapping values agreed to within 0.5 degrees F and 0.5 inches approximately 95 percent of the time. Some of the smaller differences may be due to averaging, GHCN's conversion to metric units, and/or CDMP's use of tenths of degrees Fahrenheit when present. An additional step in the QC process should be to review the monthly summary images where GHCN and CDMP-Forts daily values differ. This would be made easier if both datasets were in a common database, including all station names, identifiers, latitudes, longitudes, elevations, and all variables.

Secondly, additional data would greatly improve station density, resulting in a more nuanced view at finer scales, as well as providing a means to better evaluate the confidence of temperature and precipitation values at nearby stations. The QC of the remaining 95 CDMP-Forts keyed stations should be completed to increase the spatial density of stations. Further, for the 1877–1878 period, approximately 350 stations imaged but un-keyed should be both keyed and quality controlled. The full database for 1820 to 1892 would benefit from the keying and QC of the remaining 4350 stations.

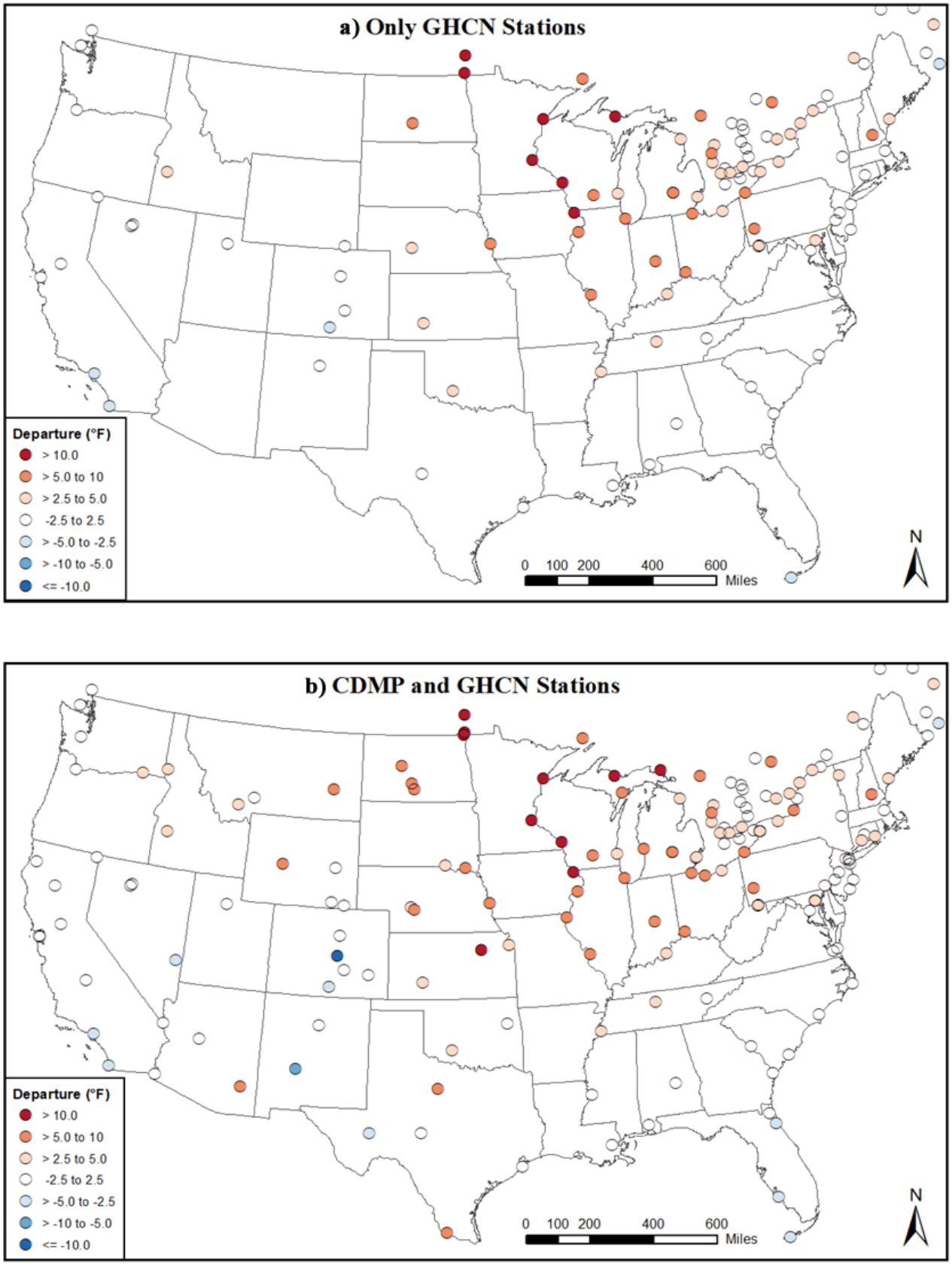


Figure 8: For the Dec 1877–March 1878 period, maps of the maximum temperature anomaly (F) using a) GHCN stations alone and b) CDMP-Forts and GHCN stations.

c. Flash Freeze Case, December 1836

The CDMP-Forts data are ideal for examination of case studies of meteorological phenomena, as many variables beyond temperature and precipitation were keyed. Ludlum (1968) provided written accounts of numerous weather events in the early United States using observational data, government reports, and personal accounts from newspapers and diaries. In one such case, the movement of a ferocious cold wave produced a severe flash freeze event (December 1836). Obrecht et al. (2012) reconstructed this event in detail, tracking the progress of the cold front from the upper Midwest to the East Coast over 3 days, with 17 additional stations not available to Ludlum (1968). The impact of this flash freeze was most severe where the temperature fell below freezing during daylight hours. The unexpected nature of this event led to many being unprepared, resulting in deaths of persons and livestock and a few "frozen to the saddle" (Ludlum, 1968).

Wendland (1987) developed a climatology of flash freeze events, defining them as having a > 39.6°F (22°C) difference between the maximum temperatures on one day and minimum temperatures the following day and when the minimum temperature fell to freezing or below. In this case, the 24-hour difference in at-hour temperatures was used.

The Obrecht et al. (2012) case study benefitted from additional data, employing 38 stations, as opposed to the 21 stations used by Ludlum (1968), with 11 of the stations being common to both studies. Ludlum (1968) accessed meteorological data from network stations and from written accounts such as diaries and newspapers. A visual summary of this flash freeze event is presented in Figure 9 with select images from Obrecht et al. (2012). Even with a limited number of stations and with uncertainty in the exact accuracy of the variables reported, these historical data present a relatively clear picture of the magnitude and progression of this cold front. With careful use of three-times-daily observations and the inclusion of pressure and wind, case studies of early weather phenomena can be reconstructed to a higher degree of resolution and timing than previous studies could achieve. This would be especially important for events occurring prior to the first publication of daily

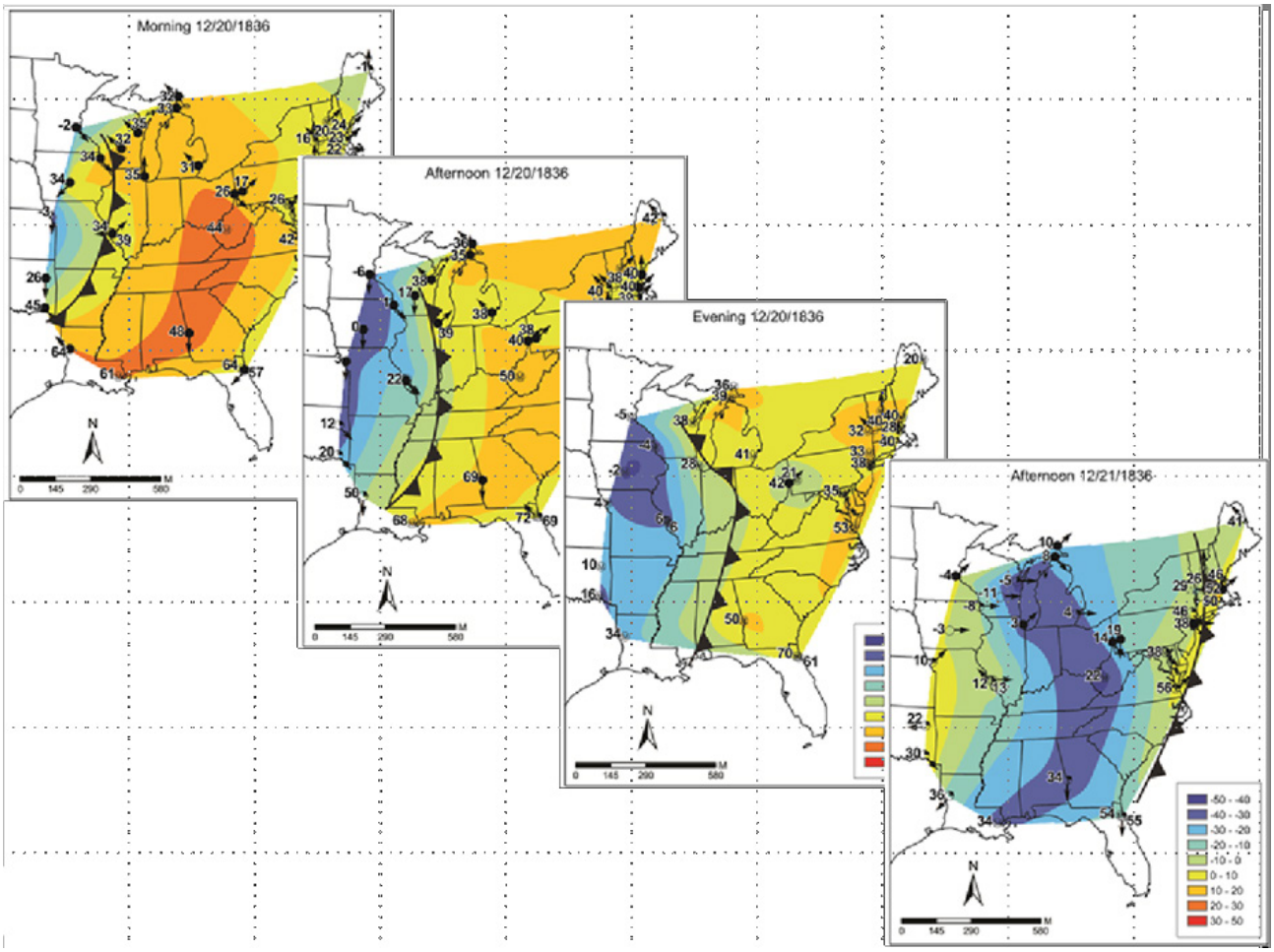


Figure 9: Weather maps from the morning of 20 Dec 1836 to the afternoon of 21 Dec 1836 with the most likely position of the cold front plotted. The station model at each site shows current air temperature (°F), cloud cover (fourths), wind direction, and when available, pressure in italics. Shading indicates the previous 24-hour temperature change (°F).

weather maps in 1871.

d. Additional Applications

Requests for 19th century data have been made for studies to evaluate other data, such as tree rings, 1876 weather maps, Native American climate records, and information available from diaries, historical accounts, and historical fiction. Other requests for early period data have been made to provide insight into weather and climate impacts on mortality and morbidity of Union Army veterans, forest management practices, frost bite injuries, frost damage in woody forests, and the effect of severe cold on the wine industry.

Studies of note using 19th century data include the study of the Laura Ingalls Wilder's The Long Winter (Boustead et al., 2020), research on drought periods in the Great Plains (Gallo and Wood, 2015; Burnette and Stahle, 2012), an investigation of winter season counts by Native Americans (Therrell and Trotter, 2011), and a collection of case studies of historical climate events in Dupigny-Giroux and Mock (2010). These studies relied in varying degrees on newly keyed and quality-controlled CDMP-Forts data, GHCN data, and on data the user keyed and quality controlled. Although the ability to conduct research on historical events was aided by the CDMP program, the coordinated effort to preserve, digitize, and quality control early weather data ceased in 2012 because of budgetary cuts and changing federal priorities. Revitalization of these efforts would facilitate and enhance future historical studies, as well as improve climate change model verification and climate trend and variability studies.

4. Relationship to Global Datasets and Data Rescue Efforts

Despite resource limitations that have halted further keying and QC of these important historical data for the present, the MRCC and NCEI continue to maintain the CDMP-Forts database. Both institutions provide the quality-controlled 19th century CDMP-Forts data to users upon request. These data have been incorporated into the International Pressure Database (Cram et al., 2015) and thus are a major component of the Twentieth Century Reanalysis (20CR; Compo, et al., 2011; Slivinski et al., 2019). This reanalysis dataset provides an estimate of three-hourly and daily weather at a grid spacing of 75 km at the equator, extending back to 1836 (https://www.psl.noaa.gov/data/gridded/data.20thC_ReanV3.html). The additional QC of temperature and precipitation data provides a means of verifying the 20CR, and keying additional pressure data would be invaluable in filling in spatial and temporal gaps for use in the reanalysis.

Data also have been provided to the Global Surface Temperature Initiative (Rennie et al., 2014) for eventual use in climate modeling and in the extension of global historical Climate Extreme Indices prior to 1910 (Brown et al., 2010). CDMP-Forts metadata are included in a pre-1850 metadata database (Bronnimann et al., 2019). This metadata project highlights the worldwide desire for the global recovery of historical weather data. In many regions such as the U.S. and Europe, data rescue involves keying missing data prior to 1900. In other regions, data to be rescued often include lost colonial-era data (prior to about 1960) or more recent data lost due to poor storage practices, war, and natural disasters, or simply data for which digitization was impractical or not a high priority.

A focus of the World Meteorological Organization (WMO) for many years has been to aid individual National Metrological and Hydrological Services in collecting and preserving their climate data. The Atmospheric Circulation Reconstruction over the Earth (ACRE) initiative (Allan et al., 2011) likewise spearheads and helps support data rescue around the world. More recent efforts by the WMO Commission on Climate and the European Union through the Copernicus C3S Data Rescue Service have been to develop an infrastructure to consolidate data rescue project information (https://idare-portal.org/), collect to-be-rescued metadata and data (https://datarescue.climate.copernicus.eu/), and facilitate the development of technical resources for data rescue efforts. The recent role of NOAA has been to store and catalog monthly summary forms from around the world.

With minimal governmental support over the past decade, much of the focus of data rescue has shifted to citizen science projects run largely by volunteer organizations. Most notably, Old Weather has been quite successful in digitizing ship logs (https://www.oldweather.org/), providing invaluable weather information across the oceans during the first half of the 20th century. In North America, two groups are actively involved in citizen science-based climate data rescue: ACRE-Canada (https://citsci.geog.mcgill.ca/en/) and the International Environmental Data Rescue Organization (IEDRO). The Data Rescue: Archives and Weather (DRAW) program undertaken by ACRE-Canada is focused on digitizing weather observations from the Montreal McGill Weather Observatory since 1863 (Slonosky, 2019, Slonosky and Sieber, 2020). IEDRO has undertaken many international data rescue missions in Africa and Central and South America and is developing both an alpha-numeric digitization software and chart-tracing software for use in citizen science and National Meteorological and Hydrological Services (NMHS)-staffed data rescue efforts. These citizen science programs, while being able to attract keyers and tracers for data entry, still require development of infrastructure and staff to manage and quality control the digitized data (Slonosky, 2020). With a large number of un-keyed images from approximately 4350 stations and demonstrated user applications, the CDMP-Forts dataset is a good candidate for either a citizen science program or a dedicated governmental effort.

5. Conclusions

Digitization and analysis of historical data is challenging and requires painstaking effort. Changing observational practices, observer errors, illegible penmanship, and poorly documented siting and instrumentation all contribute to uncertainty. Hence, considerable time, effort, and expense are required to not only key and QC the data, but also to integrate them with existing digital data and reliably analyze them. These observations are an important part of our written history and are exceptionally valuable to the climate community for documentation and validation of paleoclimate techniques such as tree ring and ice core analyses, reconstructed datasets such as the 20CR, and verification of non-standard observations by indigenous peoples or others without access to meteorological instrumentation. Economists, historians, biologists, and social scientists can all benefit from these data when studying the effects of weather and a changing climate on the environment and on societal responses to these changes.

Completing the keying and quality control of additional stations would provide more information to evaluate both the variability and long-term trends in weather phenomena. Our knowledge of large-scale atmospheric events would be better defined in time and space, particularly with respect to maximum and minimum temperatures for the 1870–1892 period and for daily precipitation from 1788 onward. The 1836 flash-freeze event demonstrates that individual case studies can be recreated using multiple variables even if those observations did not conform to current observational practices.

The digital data for the 355 quality-controlled sites are available upon request and at no cost from the MRCC or NCEI. An inventory by year and month for the U.S. stations and station histories are available on the MRCC website. Actual images of the observation forms from which the digital data are keyed are available from the MRCC or from the NCEI archives.

6. Acknowledgments

The authors wish to acknowledge the foundational inspiration and dedication of Stephen Doty of NCEI to CDMP, as well as the many who contributed to the program. Thanks also go to Zoe Zaloudek and Eileen Deremiah for their technical assistance of this manuscript and to the diligence of three anonymous reviewers.

REFERENCES

Allan, R., P. Brohan, G.P. Compo, R. Stone, J. Luterbacher, and S. Bronnimann, 2011: The International Atmospheric Circulation Reconstructions over the Earth (ACRE) Initiative. Bull. Amer. Meteor. Soc., 92, 1321-1425. https://doi.org/10.1175/2011BAMS3218.1

Andsager, K., M. C. Kruk, and M. L. Spinar, 2005: A comprehensive single-station quality control process for historical weather data. Preprint, 15th Applied Climatology Conference, Amer. Meteor. Soc., Savanna, GA. http://ams.confex.com/ams/15AppClimate/techprogram/paper_91763.htm

Andsager, K., M.C. Kruk, M.L. Spinar, and L. Stoecker, 2010: DSI-3297 CDMP Coop Summary of the Day Forts. Version 1.1. Prepared for the Midwestern Regional Climate Center, 41p.

Barchard, K.A., and L.A. Pace, 2011: Preventing human error: The impact of data entry methods on data accuracy and statistical results. Comput. in Hum. Behav., 27, 1834-1839.

Boustead, B.M., M. Shulski, and S.D. Hilberg, 2020: The long winter of 1880-81. Bull. Amer. Meteor. Soc., 101, E797-E814. https://doi.org/10.1175/BAMS-D-19-0014.1

Bronnimann, S., and co-authors, 2019: Unlocking pre-1850 instrumental meteorological records. A global inventory. Bull. Amer. Meteor. Soc., 100, ES389-ES413. https://journals.ametsoc.org/view/journals/bams/100/12/bams-d-19-0040.1.xml

Brown, P.J., R.S. Bradley, and F.T. Keimig, 2010: Changes in extreme climate indices for the northeastern United States, 1870-2005. J. Climate, 23, 6555-6572. https://doi.org/10.1175/2010JCLI3363.1

Burnette, D.J., and D.W. Stahle, 2012: Historical perspective on the Dust Bowl drought in the central United States. Clim. Change, 116, 479-494.

Burnette, D.J., D.W. Stahle, and C.J. Mock, 2010: Daily-mean temperature reconstruction for Kansas from early instrumental and modern observations. J. Climate, 23,1308-1333.

Compo, G.P., J.S. Whitaker, P. Sardeshmukh, N. Matsui, R.J. Allan, X. Yin, … S.J. Worley, 2011: The Twentieth Century Reanalysis Project. Quart. J of the Roy. Meteor. Soc., 137. https://doi.org/10.1002/qj.776

Conner, G., 2008: History of the prescribed times for temperature observations, United States, 1817-1949. Monograph, National Climatic Data Center Library, Asheville, NC. 60p.

Cram, T.A. ,G.P. Compo, X. Yin, R.J. Allan, C. McColl, S. Vose, and S.J. Worley, 2015: The International Surface Pressure Databank version 2. https://doi.org/10.1002/gdj3.25

Dupigny-Giroux, L. and C.J. Mock (eds.), 2010: Historical climate variability and impacts in North America. Springer Publishers, 278p.

Dupigny-Giroux, L., T.F. Ross, J.D. Elms, R. Truesdell, and S.R. Doty, 2007: NOAA's Climate Database Modernizations Program: Rescuing, archiving and digitizing history. Bull. Amer. Meteor. Soc., 88, 1015-1017.

Gallo, K. and E. Wood, 2015: Historical drought events of the Great Plains recorded by Native Americans. Great Plains Research, Univ. Nev. Press., 25, 151-158.

Grice F.R. and P. Boulay, 2005: History pf Weather Observations Fort Snelling MN, 1819-1892. Midwestern Regional Center, 28 p.

Hamilton, K., 1986: Early Canadian weather observers and the "Year Without a Summer", Bull. Amer. Meteor. Soc., 67, 524-532.

Hopkins, E.J., and J.M. Morin, 2010: Monitoring the climate of the old northwest: 1820-1895 in Historical climate variability and impacts.

Huang, B., M. L., Heureux, Z-Z Hu, X. Yin, and H-M. Zhang, 2020: How significant was the 1877/78 El Niño? J. Climate, 33, 4853-4869. https://doi.org/10.1175/JCLI-D-19-0650.1

Kiladis, G.N., and H.R. Diaz, 1986: An analysis of the 1877–78 ENSO Episode and comparison with 1982–83. Mon. Wea. Rev., 114, 1035-1047. https://doi.org/10.1175/1520-0493(1986)114<1035:AAOTEE>2.0.CO;2

Kunkel, K. E., D. R. Easterling, K. Hubbard, K. Redmond, K. Andsager, M. C. Kruk, and M. L. Spinar, 2005: Quality control of pre-1948 Cooperative Observer Network data. J. Atmos. Oceanic Technol., 22, 1691–1705. https://doi.org/10.1175/JTECH1816.1.

Kunkel, K.E., K. Andsager, and D.R. Easterling, 1999: Long-term trends in extreme precipitation events over the Conterminous United States and Canada. J. Climate, 12, 2515-2527. https://doi.org/10.1175/1520-0442(1999)012<2515:LTTIEP>2.0.CO;2

Lai, Y., and D.A. Dzombak, 2019: Use of historical data to assess regional climate change. J. Climate, 32, 4299-4320. https://doi.org/10.1175/JCLI-D-18-0630.1.

Ludlum, D.M., 1948: That summer of 1816. Weatherwise, 1, 54-63.

Ludlum, D.M., 1968: Early American Winters, II 1821-1870. Amer. Meteor. Soc., Boston, 150-152.

McKitrick R. and J. Christy, 2019: Assessing Changes in US Regional Precipitation on multiple time scales. J. Hydrol., 578. https://doi.org/10.1016/j.jhydrol.2019.124074.

Menne, M.J., I. Durre, R.S. Vose, B.E. Gleason, and T.G. Houston, 2012: An overview of the Global Historical Climatology Network Daily database. J. Atmos. Oceanic Technol., 29, 897-910. https://doi.org/10.1175/JTECH-D-11-00103.1.

Miller, E.R., 1931: The evolution of meteorological observations in the United States, Mon. Wea. Rev., 59, 1-6.

Obrecht, R.L., M.L. Spinar, L.A. Stoecker, and N.E. Westcott, 2012: Frozen to the saddle: A 19th Century Flash Freeze Investigation, Preprint. 19th AMS Conference on Applied Climatology, Ashville, N.C., 4pp. https://ams.confex.com/ams/19Applied/webprogram/Paper190356.html

Rennie, J., J.H. Lawrimore, B.E. Gleason, P.W. Thorne, C.P. Morice, M.J. Menne, C.N. Williams, G,W. de Almeida, J.R. Christy, M. Flannery, M. Ishihara, A.M.C. Kamig, A. Mhanda, D.H. Lister, V. Razuvaev, M. Renom, S.J. Rusticucci, V. Venema, W. Angel, M. Brunet, B. Dattore, H.J. Diamond, H.F. La, J. Luterbacher, H. Machel, J. Revadekar, R.S. Vose, X. Yin, 2014: The international surface temperature initiative global land surface databank: Monthly temperature data release description and methods. Geosci. Data J. Open Access. https://doi.org/10.1002/gdj3.8

Shein, K.A., D.P. Todey, F.A. Akyuz, J.R. Angel, T.M. Kearns, and J.L. Zdrojewski, 2013: Revisiting the statewide climate extremes for the United States. Bull. American Meteorological Society, 94, 393-402. https://doi.org/10.1175/BAMS-D-11-00013.1.

Slinskey, E.A., P.C. Loikith, D.E. Waliser, and A. Goodman, 2019: An extreme precipitation categorization scheme and its observational uncertainty over the Continental United States. J. Hydrometeor., 20, 1029-1052. https://doi.org/10.1175/JHM-D-18-0148.1.

Slivinski, L.C., G.P. Compo, J.S. Whitaker, P.D. Sardeshmukh, B.S. Giese, C. McColl, R. Allan, and X. Ying, 2019: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. Q. J. Roy Meteor. Soc.,145(724). https://doi.org/10.1002/qj.3598.

Slonosky, V, 2015: Daily Minimum and Maximum Temperature in the St. Lawrence Valley, Quebec: two centuries of climatic observations from Canada. Int. J. Climatol., 35, 1662-1681.

Slonosky, V. and R. Sieber, 2020: Building a Traceable and Sustainable Historical Climate Database: Interdisciplinary and DRAW. Patterns, 1, 100012.

Slonosky, V., 2019: Historical weather data rescue with McGill's DRAW (Data Rescue: Archives and Weather) in Canada and around the world. Bull. Can. Meteorol. Ocean. Soc. https://bulletin.cmos.ca/draw-historical-weather-data-rescue/

Therrell, M. D. and M. J. Trotter, 2011: Waniyetu Wówapi: Native American records of weather and climate. Bull. Amer. Meteor. Soc., 92, 583-592. https://doi.

org/10.1175/2011BAMS3146.1

Truesdell, R. T., J. Cooper, E. Freeman, and D. L. O'Connell, 2008: 1.3 The Forms Tell a Tale – Unique Weather Observing Practices of the 19th Century, Preprint. 17th AMS Conf. on Appl. Climatol., 8p.

Wendland, W.M., 1987: Prominent November cold waves in the North Central United States since 1901. Bull. Amer. Soc., 68, 616-619.

Westcott, N.E., K. Andsager, L. Stoecker, M.L. Spinar, R. Smith, R. Obrecht, M.C. Kruk, R. Truesdell, and D. O'Connell, 2011: Quality Control of 19th Century Weather Data. Illinois State Water Survey Contract Report 2011-02, 48p. https://www.isws.illinois.edu/pubdoc/CR/iswscr2011-02.pdf