Long-Term Daily Climate Records from Stations Across the Contiguous United States

 USHCN Home  Daily Data  Monthly Data  Monthly Data Documentation
 Web Interface

Investigators

M.J. Menne, C.N. Williams, Jr., and R.S. Vose
National Climatic Data Center, National Oceanic and Atmospheric Administration

Table of Contents

- Introduction
- Data Sources and Compilation
- Quality Control
- Bias Adjustments
- Station Information
- Data Files
- Data Access
- References
- Contacts

Introduction

The United States Historical Climatology Network (USHCN) is essentially a subset of the U.S. Cooperative Observer Network operated by NOAA's National Weather Service (NWS). The approximately 1200 USHCN stations were originally selected according to factors such as record longevity, percentage of missing values, spatial coverage as well as the number of station moves and/or other station changes that may affect data homogeneity. Most USHCN stations are situated in rural areas or small towns; however, a smaller number of stations are also part of the NOAA NWS synoptic network, whose stations are generally located at airports in more urbanized environments. USHCN datasets have been developed at NOAA's National Climatic Data Center (NCDC) in collaboration with the Department of Energy's Carbon Dioxide Information Analysis Center (CDIAC).

The USHCN project dates to the mid-1980s (Quinlan et al. 1987). At that time, in response to the need for an accurate, unbiased, modern historical climate record for the United States, personnel at the Global Change Research Program of the U.S. Department of Energy and at NCDC defined a network of 1219 stations in the contiguous United States whose observation would comprise a key baseline dataset for monitoring U.S. climate. Since then, the USHCN *monthly* dataset has been updated several times (e.g., Karl et al., 1990; Easterling et al., 1996). The USHCN Version 2 Serial Monthly data release is the most recent update to the USHCN datasets.

The first database released by NCDC to contain *daily* data from USHCN stations, the HCN/Daily (HCN/D; Hughes et al. 1992; hereafter H92), contained daily maximum and minimum temperatures and precipitation totals from 138 select U.S. stations. The temperature and precipitation records from these stations were considered to be the most reliable, internally consistent, and unbiased records from the USHCN. These records were compiled from digital and nondigital data sets archived at NCDC that come from a variety of sources, including climatological publications, universities, federal agencies, individuals, and data archives. Records were subjected to extensive manual and automated quality assurance (QA) checks. The selection of stations for inclusion in H92 was performed with the following data quality issues in mind.

- The degree to which each station maintained a constant observation time for maximum and minimum temperatures, excursions from a station's predominant observing time of no more than four years being desired.
- At least 95% of a station's pre-1951 data should be contained in NCDC digital daily archives.
- A station's potential for heat island bias over time should be low.
- Quality assessments based upon the decile ranking assigned by Karl et al. (1990) to the stations' monthly maximum/minimum temperature data for certain quality characteristics.

Since the release of H92, much more work has been conducted at NCDC involving compilation and digitizing of daily data. However, to enable the compilation of a database providing better spatial coverage of the contiguous United States, the four station selection criteria listed above were not strictly adhered to in later versions of the USHCN daily database (e.g., Williams et al. 2004).

The USHCN daily data contain observations of maximum and minimum temperature, precipitation amount (liquid equivalent), snowfall amount, and snow depth. Records from about 700 stations begin before 1900 and most station records are essentially complete for at least 60 years extending through a recent year. It's common for temperature and precipitation records to have started earlier than snowfall and snow depth measurements, especially at stations with longer periods or record.

Data Sources and Compilation

The current version of the USHCN Daily database contains the same variables, and was subjected to the same quality control checks (details below), as NCDC's Global Historical Climatology Network-Daily (GHCN-Daily) database. USHCN-Daily is in fact a subset of GHCN-Daily, which serves as the official archive for daily data from the Global Climate Observing System (GCOS) Surface Network (GSN). Following is a list of the sources of data currently contained in USHCN-Daily along with a brief description of each source.

- U.S. Cooperative Summary of the Day (NCDC DSI-3200) - Includes daily observations at 19,000 stations in the United States and its territories. While most measurements are taken by volunteer observers as part of the U.S. NOAA National Weather Service's Cooperative Observer (COOP) Program, manual and automated measurements from some "First Order" sites are also included. Some records extend back to the late 1800s, but most do not begin until 1948 or later.

- CDMP U.S. Cooperative Summary of the Day (NCDC DSI-3206) - Contains mostly pre-1948 observations from more than 11,000 COOP stations that were keyed as part of NCDC's Climate Data Modernization Program (CDMP).
- U.S. First Order Summary of the Day (NCDC DSI-3210) - Contains historical and present-day manual and automated observations from approximately 1600 sites, including U.S. First Order stations, a selection of Canadian sites, and U.S.-operated stations in other countries.

Quality Control

As part of the GHCN-Daily quality control procedures, USHCN-Daily data have been subjected to a variety of internal consistency, frequent-value, outlier, and spatial consistency checks. The quality control efforts are always improving and expanding, so users are urged to keep up with them by visiting the Quality Control section of the NCDC/GHCN website.

Bias Adjustments

At present, the USHCN daily data contain no adjustments for biases resulting from historical changes in instrumentation and observing practices. This is true of the GHCN-Daily database as a whole, which includes the USHCN daily data. However, there is ongoing work at NCDC to develop adjustments that can be applied to daily maximum and minimum temperatures, and a GHCN daily derived product containing adjusted daily temperatures may become available in the future.

Station Information

The format of each record in the USHCN Daily station inventory file (ushcn-stations.txt) is as follows.

Variable	Columns	Type
COOP ID	1-6	Character
LATITUDE	8-15	Real
LONGITUDE	17-25	Real
ELEVATION	27-32	Real
STATE	34-35	Character
NAME	37-66	Character
COMPONENT 1	68-73	Character
COMPONENT 2	75-80	Character
COMPONENT 3	82-87	Character
UTC OFFSET	89-90	Integer

These variables have the following definitions:

COOP ID is the U.S. Cooperative Observer Network station identification code. Note that the first two digits in the Coop ID correspond to the assigned state number

(see Table 1 below).

LATITUDE	is latitude of the station (in decimal degrees).
LONGITUDE	is the longitude of the station (in decimal degrees).
ELEVATION	is the elevation of the station (in meters, missing = -999.9).
STATE	is the U.S. postal code for the state.
NAME	is the name of the station location.
COMPONENT 1	is the Coop Id for the first station (in chronologic order) whose records were joined with those of the USHCN site to form a longer time series. "------" indicates "not applicable".
COMPONENT 2	is the Coop Id for the second station (if applicable) whose records were joined with those of the USHCN site to form a longer time series.
COMPONENT 3	is the Coop Id for the third station (if applicable) whose records were joined with those of the USHCN site to form a longer time series.
UTC OFFSET	is the time difference between Coordinated Universal Time (UTC) and local standard time at the station (i.e., the number of hours that must be added to local standard time to match UTC).

TABLE 1. State numbers and abbreviations for the contiguous United States.

State number	State abbreviation	State
01	AL	Alabama
02	AZ	Arizona
03	AR	Arkansas
04	CA	California
05	CO	Colorado
06	CT	Connecticut
07	DE	Delaware
08	FL	Florida
09	GA	Georgia
10	ID	Idaho
11	IL	Idaho
12	IN	Indiana
13	IA	Iowa
14	KS	Kansas
15	KY	Kentucky
16	LA	Louisiana
17	ME	Maine
18	MD	Maryland
19	MA	Massachusetts
20	MI	Michigan
21	MN	Minnesota

22	MS	Mississippi
23	MO	Missouri
24	MT	Montana
25	NE	Nebraska
26	NV	Nevada
27	NH	New Hampshire
28	NJ	New Jersey
29	NM	New Mexico
30	NY	New York
31	NC	North Carolina
32	ND	North Dakota
33	OH	Ohio
34	OK	Oklahoma
35	OR	Oregon
36	PA	Pennsylvania
37	RI	Rhode Island
38	SC	South Carolina
39	SD	South Dakota
40	TN	Tennessee
41	TX	Texas
42	UT	Utah
43	VT	Vermont
44	VA	Virginia
45	WA	Washington
46	WV	West Virginia
47	WI	Wisconsin
48	WY	Wyoming

Data Files

USHCN data files may be downloaded from CDIAC's anonymous FTP area (see the USHCN Data Access page).

There are 48 state files (ASCII) using the following naming convention: stateno_stateabbrev.txt.gz.

Example: state01_AL.txt.gz.

station_file_format.txt contains ASCII file format information and variable definitions.

There are also five NetCDF files containing data organized by variable type (e.g., ushcn_prcp.nc.gz).

ushcn-stations.txt contains a listing of all of the stations along with their latitude and longitude.

us.txt.gz contains data for the 48 states represented.

The format of each record in an ASCII data file, be it a state-level file (e.g., state01_AL.txt) or the file for the entire U.S. (us.txt) is as follows. (Each record in a file contains one month of daily data.)

Variable	Columns	Type
COOP ID	1-6	Character
YEAR	7-10	Integer
MONTH	11-12	Integer
ELEMENT	13-16	Character
VALUE1	17-21	Integer
MFLAG1	22	Character
QFLAG1	23	Character
SFLAG1	24	Character
VALUE2	25-29	Integer
MFLAG2	30	Character
QFLAG2	31	Character
SFLAG2	32	Character
VALUE31	257-261	Integer
MFLAG31	262	Character
QFLAG31	263	Character
SFLAG31	264	Character

These variables have the following definitions:

COOP ID	is the U.S. Cooperative Observer Network station identification code. Note that the first two digits in the Coop Id correspond to the state.
YEAR	is the year of the record.
MONTH	is the month of the record.
ELEMENT	is the element type. There are five possible values: PRCP = precipitation (hundredths of inches) SNOW = snowfall (tenths of inches) SNWD = snow depth (inches) TMAX = maximum temperature (degrees F) TMIN = minimum temperature (degrees F)
VALUE1	is the value on the first day of the month (missing = -9999).
MFLAG1	is the measurement flag for the first day of the month. There are five possible values: Blank = no measurement information applicable

B = precipitation total formed from two 12-hour totals

D = precipitation total formed from four six-hour totals

H = represents highest or lowest hourly temperature

L = temperature appears to be lagged with respect to reported hour of observation

P = identified as "missing presumed zero" in DSI 3200 and 3206

T = trace of precipitation, snowfall, or snow depth

QFLAG1 is the quality flag for the first day of the month. There are fourteen possible values:

Blank = did not fail any quality assurance check

D = failed duplicate check

G = failed gap check

I = failed internal consistency check

K = failed streak/frequent-value check

L = failed check on length of multiday period

M = failed megaconsistency check

N = failed naught check

O = failed climatological outlier check

R = failed lagged range check

S = failed spatial consistency check

T = failed temporal consistency check

W = temperature too warm for snow

X = failed bounds check

Z = flagged as a result of an official Datzilla investigation

SFLAG1 is the source flag for the first day of the month. There are fifteen possible values:

Blank = No source (e.g., data value missing)

0 = U.S. Cooperative Summary of the Day (NCDC DSI-3200)

6 = CDMP Cooperative Summary of the Day (NCDC DSI-3206)

7 = U.S. Cooperative Summary of the Day -- Transmitted via WxCoder3 (NCDC DSI-3207)

A = U.S. Automated Surface Observing System (ASOS) real-time data (since January 1, 2006)

B = U.S. ASOS data for October 2000-December 2005 (NCDC DSI-3211)

F = U.S. Fort Data

G = Official Global Climate Observing System (GCOS) or other government-supplied data

H = High Plains Regional Climate Center real-time data

M = Monthly METAR Extract (additional ASOS data)

N = Community Collaborative Rain, Hail, and Snow (CoCoRaHS)

R = NCDC Reference Network Database (Climate Reference Network and Historical Climatology Network-Modernized)

S = Global Summary of the Day (NCDC DSI-9618)

NOTE: "S" values are derived from hourly synoptic reports exchanged on the Global Telecommunications System (GTS). Daily values derived in this fashion may differ significantly from "true" daily data, particularly for precipitation (i.e., use with caution).

T = SNOwpack TELemtry (SNOTEL) data obtained from the Western Regional Climate Center

U = Remote Automatic Weather Station (RAWS) data obtained from the Western Regional Climate Center

W = WBAN/ASOS Summary of the Day from NCDC's Integrated Surface Data (ISD)

X = U.S. First-Order Summary of the Day (NCDC DSI-3210)

Z = Datzilla official additions or replacements

VALUE2	is the value on the second day of the month.
MFLAG2	is the measurement flag for the second day of the month.
QFLAG2	is the quality flag for the second day of the month.
SFLAG2	is the source flag for the second day of the month.

...and so on through the 31st day of the month. Note: if the month has less than 31 days, then the remaining variables are set to missing (e.g., for April, VALUE31 = -9999, MFLAG31 = blank, QFLAG31 = blank, SFLAG31 = blank).

Data Access

The USHCN Daily data are available via FTP or a Web interface that allows users to query, plot, and download data for individual states, stations, and variables. Please see the USHCN Data Access page.

References

- Easterling, D. R., T. R. Karl, E. H. Mason, P. Y. Hughes, and D. P. Bowman. 1996. United States Historical Climatology Network (U.S. HCN) Monthly Temperature and Precipitation Data. ORNL/CDIAC-87, NDP-019/R3. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee. 280 pp.
- Hughes, P. Y., E. H. Mason, T. R. Karl, and W. A. Brower. 1992. United States Historical Climatology Network Daily Temperature and Precipitation Data. ORNL/CDIAC-50, NDP-042. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee. 140 pp.

- Karl, 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, Tennessee.
- Quinlan, F. T., T. R. Karl, and C. N. Williams, Jr. 1987. United States Historical Climatology Network (HCN) serial temperature and precipitation data. NDP-019. Carbon Dioxide Information Analysis Center. Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- Williams, C. N., R. S. Vose, D. R. Easterling, and M. J. Menne, 2004. United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data. ORNL/CDIAC-118, NDP-070. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Contacts

Questions regarding the USHCN Daily web site or data may be directed to Dale Kaiser at CDIAC.

_____

Please cite data as: M. J. Menne, C. N. Williams, Jr., and R. S. Vose, 2015. United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.