

# The US Temperature Record 4: Data Flags

/ SEP 22, 2021

The data provided has flags you should understand. Definitions below are from the USHCN Status file.

## Data Measurement Flag

```
blank = no measurement information applicable
a-i = number of days missing in calculation of monthly mean
      temperature
[] E = The value is estimated using values from surrounding
[]   stations because a monthly value could not be computed
[]   from daily data; or,
[]       the pairwise homogenization algorithm removed the value
[]   because of too many apparent inhomogeneities occurring
[]   close together in time.
```

## Quality Control Flag

```
BLANK = no failure of quality control check or could not be
        evaluated.

D = monthly value is part of an annual series of values that
    are exactly the same (e.g. duplicated) within another
    year in the station's record.

I = checks for internal consistency between TMAX and TMIN.
    Flag is set when TMIN > TMAX for a given month.

L = monthly value is isolated in time within the station
    record, and this is defined by having no immediate non-
    missing values 18 months on either side of the value.

M = Manually flagged as erroneous.

O = monthly value that is >= 5 bi-weight standard deviations
    from the bi-weight mean. Bi-weight statistics are
    calculated from a series of all non-missing values in
    the station's record for that particular month.

S = monthly value has failed spatial consistency check.
    Any value found to be between 2.5 and 5.0 bi-weight
```

standard deviations from the bi-weight mean, is more closely scrutinized by examining the 5 closest neighbors (not to exceed 500.0 km) and determine their associated distribution of respective z-scores. At least one of the neighbor stations must have a z score with the same sign as the target and its z-score must be greater than or equal to the z-score listed in column B (below), where column B is expressed as a function of the target z-score ranges (column A).

A	B
4.0 - 5.0	1.9
3.0 - 4.0	1.8
2.75 - 3.0	1.7
2.50 - 2.75	1.6

W = monthly value is duplicated from the previous month, based upon regional and spatial criteria and is only applied from the year 2000 to the present.

Quality Controlled Adjusted (QCA) QC Flags:

A = alternative method of adjustment used.

M = values with a non-blank quality control flag in the "qcu" dataset are set to missing the adjusted dataset and given an "M" quality control flag.

Data Source Flag

- Blank = Value was computed from daily data available in GHCN-Daily
- Not Blank = Daily data are not available so the monthly value was obtained from the USHCN version 1 dataset. The possible Version 1 DSFLAGS are as follows:
- 1 =

2 =

3 =

4 =

5 =

6 =
- Books, U.S. Department of Commerce, Weather Bureau, USGPO (1960)
- section), F.H. Bigelow, U.S. Weather Bureau (1912); and, Bull

Summary of the Climatological Data for the United States, 2nd  
7 = Local Climatological Data (LCD), monthly NCDC publication  
8 = State Climatologists, various sources  
B = Professor Raymond Bradley - Refer to Climatic Fluctuations of  
United States During the Period of Instrumental Records, Brad  
Contribution No. 42, Dept. of Geography and Geology, Universi  
Massachusetts (1982)  
D = Dr. Henry Diaz, a compilation of data from Bulletin W, LCD, a  
Deck 3220 (1983)  
G = Professor John Griffiths - primarily from Climatological Data

Most of these flags I ignore because they either won't change the annual averages or they are the consequence of an opinion. The only flag I care about is the **I** flag, meaning the reported Tmin is larger than the Tmax for that month. I'll report later how many flags there are in the example dataset I'll use.

---

Revision #1

Created 7 April 2024 19:08:20 by bruce

Updated 7 April 2024 19:09:09 by bruce