University of Wisconsin-Extension

GEOLOGICAL AND NATURAL HISTORY SURVEY
3817 Mineral Point Road
Madison, Wisconsin 53705

M.E. Ostrom, State Geologist and Director

# STATION HISTORY DATA FILE FORMAT FOR PERSONAL COMPUTER APPLICATIONS USING DAILY CLIMATOLOGICAL DATA

by

D.R. Clark

Open-File Report 86-4
6 p.



1986

# A Station History Data File Format for Personal Computer Applications Using Daily Climatological Data

Douglas R. Clark
State Climatologist
Geological and Natural History Survey
University of Wisconsin-Extension

NC-94 Annual Meeting
October 8-9, 1986
East Lansing, Michigan

This paper presents a draft data file format for information about climatological observing stations. The purpose of such a file is to enable applications programs that use daily climatological data to have convenient access to information about the station, such as name, location, period of record, and time of observation. The format is simple to create, maintain and use, and requires no special software. The information selected for this file format has been limited to items that may have use in quantitative applications, such as physically- or statistically-based models. It is not intended to be a complete station history database. Easy availability of station history information will encourage development of applications that take station characteristics and changes into account. A standard format will allow the states to exchange station history information readily, and jointly develop software that employs this information.

Wisconsin Geological and Natural History Survey
Open File Report WOFR 86-4.

Station History File for Users

Purpose: Provide PC users with a convenient data file on a station's history to go with the long-term daily data. It also provides auxiliary data about the station in a file separate from the data.

Information in file:

1) station name
2) station identifying numbers: state, station no., division
3) current location: latitude, longitude, elevation
4) current observation times
5) list of station changes, from beginning of record
   a) station moves
   b) changes in observation time

Requirements:

1) Ordinary ASCII datafile, in form readable by most languages, spreadsheets, and analysis packages and text editors on PC's (i.e., comma-separated data).
2) Simple file structure to make it easy for users to access station history data for simple applications programs.
3) Easily updated and documented.
4) Set up so that users can read and use the basic identifying and location information from the file without having to process the more detailed historical data in the file. (Do

this by making first few lines of the file contain the most-used information, so users can just read those lines and then quit.)

5) Each station would have a separate information file so that users only need the files for the stations they want to use, and they do not have to access a large dataset to get that information.

6) Applications that only need current station information can read just the first few lines of the file.

Layout and Specifications:

Line 1: Station name, 24 characters long. Could also be longer to allow for more descriptive name.

Line 2: State code, station number, division

Three values; separated by commas. These are the NCDC station identifiers.

Line 3: Year1, month1, day1

Three integer values separated by commas giving the date of the first day of data in the daily data file. (first data line) Year is given in full form, e.g. 1948.

Line 4: Year2, month2, day2

Three integer values separated by commas, giving the date of the last (most recent) day of data in the data file. This is also the date for the "current" station information on the next lines.

Line 5: Latitude, longitude, elevation

Three decimal values separated by commas. Latitude and longitude in decimal degrees, positive north and west. Elevation in meters above mean sea level. These values are the "current" values for the station as of the date on line 4 above.

Line 6: Time of temperature, time of precipitation, time zone, daylight flag

Three decimal values giving the hour (and fraction of hour) of observation times in local standard time and the time zone of local standard time, in hours after Greenwich Mean Time (GMT or UT), and an integer value, 0 or 1, indicating whether the observation time shifts to daylight time (1) or remains on standard time all year (0). These values are the "current" observation times as of the most recent date in the file (line 4). Hours are 1 to 24, with 24 being midnight. 99 if not observed.

Line 7 to end of file: station history data lines

Each line has one of three formats, but begins with the date and a code indicating the "type" of the line. The date always indicates the day on which the new status begins.

Type 1: Station opening, move, or reopening. Includes new location information.

Format: YYYY,MM,DD,1,location,latitude,longitude,elevation.

Location is an integer index for each site at which observations have been taken, in chronological order beginning with 1. Latitude, longitude and elevation are the values effective on the given date for the new station location.

Type 2: Change of observation time.

Format: YYYY,MM,DD,2,TempTime,PrecipTime,TimeZone,Daylight-Code.

The times are the same format as described for line 6, as is the daylight time code, 0 or 1. If the element is not observed, the hour is given as 99.

Type 3: Station closing. Used to indicate the first day on which observations were no longer taken. This type is included to allow breaks in station observations to be noted, even when the missing data has been filled in with estimates in the data file.

Format: YYYY,MM,DD,3.

Note that the month and day numbers should have leading zeros (01, 02, ...) to maintain a fixed length for the date fields.

File naming convention:

Files will have a six-character file name and a three-character extension following MS-DOS conventions. The filename will consist of the 2-digit NCDC State Code followed by the 4-digit NCDC Station Number. The extension will be ".SHD" for "station history data." This name convention will permit files to be accessed given only the station number. It also permits files to be arranged conveniently in disk directories.

Other information not included in this file format:

1. Distance from previous location
2. Types of instruments and instrument changes
3. Types of shelters and wind-shields
4. Exposure (characteristics around instruments, such as lawn vs. woods, slope, etc.)
5. County, township and range
6. Observer information, such as name and address
7. Nearby stations