National Climatic Data Center

DATA DOCUMENTATION

FOR

DATA SET 9641L (DSI-9641L)

EXTREME SNOWFALL RETURN PERIOD STATISTICS

December 23, 2002

National Climatic Data Center
151 Patton Ave.
Asheville, NC 28801-5001 USA

Table of Contents

1. **Abstract**: The purpose of this project was to prepare snowfall extremes and return period statistics for official weather stations across the contiguous United States and Alaska. A subset of the output data was provided to the Federal Emergency Management Agency (FEMA) for use as an aid in making disaster declarations for record or near-record snowstorms.

COMPUTATIONAL METHODOLOGY

Data from 1948-1996 were analyzed. The output for this project includes observed extreme snowfall values and the extreme snowfall values that correspond to four return periods. These values were computed as follows:

(1) Four values (corresponding to four time units) were determined for each year of the data period: the greatest 1-day, greatest 2-day, and greatest 3-day snowfall amounts, and the August-July total snowfall amount.

(2) Each time unit was analyzed separately. For example, 1-day snowfall might have had 35 values (35 extreme values, one for each of 35 years), 2-day snowfall might have had 30 values, 3-day snowfall might have had 26 values, and August-July total snowfall might have had 21 values.

(3) The highest of these values was selected as the observed maximum snowfall value.

(4) These extreme values were analyzed using the Generalized Extreme-Value statistical distribution estimated using the method of L-moments and L-moment ratios described by Hosking and Wallis (1997). This analysis method can be used to compute the extreme snowfall values corresponding to any desired return period (i.e., probability level). Extreme snowfall values corresponding to the 10-year, 25-year, 50-year, and 100-year return periods were computed for each of the four time units.

STATISTICAL CONSIDERATIONS

It should be noted that a statistical distribution can be determined only from non-zero values. If too many values are zero (which will happen, for example, in warm climate regions such as the Gulf Coast states, southern New Mexico, southern Arizona, and coastal and southern California, where it rarely snows), then a statistical distribution cannot be determined and return period statistics cannot be computed. Hosking and Wallis note that at least 20 non-zero values are needed in order to determine the statistical distribution, but a study by Guttman (1994) indicates that at least 30 non-zero values are needed for stable return period statistics.

In this study, return period statistics were computed if at least 20 non-zero values were available, in order to generate return period statistics for as many stations as possible. The number of years with non-zero data are included in the output to enable the user to decide if they want to use a particular station's values. If the number of years with non-zero data is 20 or more but less than 30, then the user should exercise caution when using the return period statistics for that time unit.

Some time units may have had 20 or more years with non-zero data, while other time units for the same station may have had fewer years. In these cases, only some of the return period statistics were computed; the others have a not

:
:

available code.

DATA COMPLETENESS CONSIDERATIONS

Guidance from the World Meteorological Organization (WMO) was used, where available, for determining the values for each year. The August-July total seasonal snowfall was computed by adding the total monthly snowfall amounts from the corresponding months for each August-July Ayear@. If any month was missing, the August-July value for that Ayear@ was missing. The total monthly snowfall amounts were computed by summing the daily snowfall amounts in the month. If any daily snowfall value was missing, the monthly total was missing. In this way, the August-July total seasonal snowfall has no tolerance for missing data.

For the 1-day, 2-day, and 3-day snowfall extremes, the data for a month was discarded if more than five days were missing. If five or fewer days were missing, then the highest value for that month was used for the month. The extreme value for the August-July snow Ayear@ was the highest value of the twelve available corresponding months. It may be possible, for some locations that experience only a few days of snowfall each year, to have a year with no snowfall when, in fact, snow did fall but the data was discarded because it failed a QC test.

As noted above, the August-July total seasonal snowfall has no tolerance for missing data. The 1-day, 2-day, and 3-day snowfall extremes are more tolerant of missing data. For this reason, the August-July total seasonal snowfall will always have the same or fewer number of years with non-missing data than the 1-day, 2-day, and 3-day extremes. It may be possible for the 1-day, 2-day, or 3-day extreme value(s) to be greater than the August-July extreme seasonal total if the corresponding year(s) for the August-July total value was missing.

The 1-day extreme snowfall value is the greatest single daily snowfall amount. The 2-day extreme snowfall value is the greatest two-day snowfall amount, where data were available for both days. If heavy snow fell on one day but the day before and the day after were both missing, then that snowfall amount could not be included in a two-day total or a three-day total. In this way, it may be possible for the 1-day extreme snowfall value to be larger than the 2-day or 3-day extreme snowfall value. Likewise for 3-day snowfall. The 3-day extreme snowfall value is the greatest three-day snowfall amount, where data were available for each of the three days. If heavy snow fell on one or two days, but the day before and the day after this one-day or two-day period were missing, then that snowfall amount could not be included in a 3-day total. In this way, it may be possible for the 1-day or 2-day extreme snowfall value to be larger than the 3-day extreme snowfall value.

CAUTIONS AND WARNINGS

The users of these snow extremes and return period statistics should use considerable caution if they compare the values from one station to those of another station. The stations may not be directly comparable to one another, due to several reasons:

(1) The properties of snow make it difficult to accurately and consistently measure snowfall. As noted by Doesken and Judson (1997), snow often melts as it lands or as it lies on the ground, snow settles as it lies on the ground, and snow is easily blown and redistributed. These properties can be affected

:

by location, time of day the observations are taken, and how often they are measured.

(2) The synoptic weather patterns that generate snow can result in snowfall amounts that vary greatly over small distances (snow bands).

(3) Local topography can have a major effect. Snowfall amounts can vary greatly depending on elevation and on slope aspect, steepness, and orientation (especially with regard to the prevailing wind patterns and the wind patterns associated with any given storm).

(4) The data period is an important factor. Ideally, the same data period (with no missing data) would be desired for all stations if any inter-station comparisons were to be made. In reality, the stations have varying data periods with differing amounts of missing data.

(5) The results of a statistical analysis partly depend on how much data are analyzed (sample size). A bigger sample size (60 or 70 years of non-zero data) would provide more stable results for this type of analysis (Guttman, 1994). Unfortunately, this amount of data was not available. The preferred minimum sample size is 20 to 30 years of non-zero data, but the user should exercise caution when using return period statistics if the number of years with non-zero data is 20 or more but less than 30.

(6) Even if two stations have the same number of years with non-zero data, the history of snowfall at a location can affect the shape of the statistical distribution, which determines the snowfall amounts corresponding to the selected return periods.

2. Element Names and Definitions: This data set consists of three files on one magnetic tape. The first file contains snowfall observed extremes and return period statistics. The second file contains station metadata. The third file contains the narrative and ASCII tables that were provided to the Federal Emergency Management Agency (FEMA).

There are 8718 stations in this data set, of which 7464 were identified as “current” stations and were used in the FEMA sub-project. All of the stations have observed extreme snowfall data for four time frames, however some of the stations did not have sufficient data to compute return period statistics.

The magnetic tape and the data in each of the three files have the following specifications:

Record Length	: Fixed 160 characters
Blocked	: 8000 characters
Media	: ASCII 18-Track IBM-Type 3480 cartridge
Parity	: Odd
Label	: ANSI Standard Labeled

DATA FILE

Snowfall return period statistics were computed for extreme one-day, two-day, and three-day snowfall amounts, and for August-July snow season total snowfall, using data from 1948-1996. The return period statistics plus observed extremes output data are archived by ascending station number sort. Each data record consists of station identification information, return period statistics, observed extremes data, and number of years information.

:

:

The data consist of decimal values, integer values, and special codes. Snowfall amount units are inches. The special codes are defined as follows:

CODE VALUE	DEFINITION
-99.9	code for missing value due to data not available
-66.6	code indicating value could not be computed due to insufficient data
-8.8	observed trace value

Each data record has the following record format:

POSITION	WIDTH	TYPE	CONTENTS
001-006	6	numeric	Cooperative Network station number
007-012	6	numeric	1-day snowfall amount for 100-year return period
013-018	6	numeric	1-day snowfall amount for 50-year return period
019-024	6	numeric	1-day snowfall amount for 25-year return period
025-030	6	numeric	1-day snowfall amount for 10-year return period
031-036	6	numeric	maximum observed 1-day snowfall amount
037-039	3	numeric	1-day number of years with non-missing data
040-042	3	numeric	1-day number of years with non-zero data
043-048	6	numeric	2-day snowfall amount for 100-year return period
049-054	6	numeric	2-day snowfall amount for 50-year return period
055-060	6	numeric	2-day snowfall amount for 25-year return period
061-066	6	numeric	2-day snowfall amount for 10-year return period
067-072	6	numeric	maximum observed 2-day snowfall amount
073-075	3	numeric	2-day number of years with non-missing data
076-078	3	numeric	2-day number of years with non-zero data
079-084	6	numeric	3-day snowfall amount for 100-year return period
085-090	6	numeric	3-day snowfall amount for 50-year return period
091-096	6	numeric	3-day snowfall amount for 25-year return period
097-102	6	numeric	3-day snowfall amount for 10-year return period
103-108	6	numeric	maximum observed 3-day snowfall amount
109-111	3	numeric	3-day number of years with non-missing data
112-114	3	numeric	3-day number of years with non-zero data
115-120	6	numeric	Aug-Jul total snowfall amount for 100-year return period
121-126	6	numeric	Aug-Jul total snowfall amount for 50-year return period
127-132	6	numeric	Aug-Jul total snowfall amount for 25-year return period
133-138	6	numeric	Aug-Jul total snowfall amount for 10-year

:

			return period
139-144	6	numeric	maximum observed Aug-Jul total snowfall amount
145-147	3	numeric	Aug-Jul number of years with non-missing data
148-150	3	numeric	Aug-Jul number of years with non-zero data
151-155	5	numeric	first year of data period analyzed
156-160	5	numeric	last year of data period analyzed

STATION METADATA FILE

The station metadata are archived by ascending station number sort, with each record having the following record format:

POSITION	WIDTH	TYPE	CONTENTS
001-006	6	numeric	COOP station identification number
007-007	1	alpha	not used
008-012	5	alpha-num	latitude (ddmmh, where dd=degrees, mm=minutes, h=hemisphere [N for North, S for South])
013-013	1	alpha	not used
014-019	6	alpha	longitude (dddmmh, where ddd=degrees, mm=minutes, h=hemisphere [W for West, E for East])
020-024	5	numeric	elevation, to whole feet (-999 = missing)
025-025	1	alpha	not used
026-027	2	alpha	two-letter state abbreviation
028-028	1	alpha	not used
029-058	30	alpha	station name
059-059	1	alpha	not used
060-089	30	alpha	county name (blanks = county name not available)
090-160	71	alpha	not used

FEMA NARRATIVE AND TABLES FILE

Stations were identified as “current” based on the last year of data in the DSI-3200 database, or on the ending year information in the Station History file. The extreme snowfall and return period statistics for this subset of stations were provided to FEMA in a set of ASCII narrative and table files. These ASCII files were merged into one file and added to this data set as the “FEMA Narrative and Tables File”. This file contains the following information: 1. a narrative introduction describing the project, the data, and what was done; 2. observed snowfall extremes and return period statistics for stations in the contiguous U.S., grouped by county within each state, in a simple ASCII table format; 3. observed snowfall extremes and return period statistics for stations in Alaska, in a simple ASCII table format; 4. observed snowfall extremes for stations in the contiguous U.S., grouped by county within each state, in a simple ASCII table format; 5. observed snowfall extremes for stations in Alaska, in a simple ASCII table format; 5. metadata for stations in the contiguous U.S., grouped by county within each state, in a simple ASCII table format; 7. metadata for stations in the contiguous U.S., in ascending station sort within each state, in a simple ASCII table format; and 8. metadata for stations in Alaska, in a simple ASCII table format.

:

:

Each record in this file consists of narrative character information and is 160 characters long.

3. **Start Date**: 19480101

4. **Stop Date**: 19961231

5. **Coverage**: North America

a. Southernmost Latitude: 20N
b. Northernmost Latitude: 75N
c. Westernmost Longitude: 170W
d. Easternmost Longitude: 65W

6. **How to Order Data**:

Ask NCDC's Climate Services about the cost of obtaining this data set.
Phone: 828-271-4800
FAX: 828-271-4876
E-mail: **NCDC.Orders@noaa.gov**

7. **Archiving Data Center**:

National Climatic Data Center
Federal Building
151 Patton Avenue
Asheville, NC 28801-5001
Phone: (828) 271-4800.

8. **Technical Contact**:

National Climatic Data Center
Federal Building
151 Patton Avenue
Asheville, NC 28801-5001
Phone: (828) 271-4800.

9. **Known Uncorrected Problems**: The properties of snow make it difficult to accurately and consistently measure snowfall. Snow often melts as it lands or as it lies on the ground, snow settles as it lies on the ground, and snow is easily blown and redistributed. These properties can be affected by location, time of day the observations are taken, and how often they are measured (Doesken and Judson, 1997). For these reasons, it is important for observers to adhere to a standard methodology (National Weather Service, 1972) for observing and reporting snow. Unfortunately, stations change location, observers, and sometimes observation time. Such changes introduce inhomogeneities into the snow record. No acceptable adjustment algorithms exist to statistically adjust daily snow data for inhomogeneities.

10. **Quality Statement**: Three levels of quality control were employed in order to obtain the best snow data possible. The first level involved using the ValHiDD edited DSI-3200 values (Reek, *et al.*, 1992). The second level employed a number of additional internal consistency checks wherein the daily snowfall was compared to the corresponding maximum and minimum temperature, precipitation, and snow depth. The third level was an extremes check.

:
:

Snowfall values which failed the checks were corrected (where possible) or set to missing.

11. Essential Companion Datasets: None.

12. References:

Doesken, N.J. and A. Judson, 1997: ***A Guide to the Science, Climatology, and Measurement of Snow in the United States***, Second Edition, Colorado State University Department of Atmospheric Science: Fort Collins.

Guttman, N.B., 1994: On the sensitivity of sample L-Moments to sample size@. ***Journal of Climate***, vol. 7, pp. 1026-1029.

Hosking, J.R.M. and J.R. Wallis, 1997: ***Regional Frequency Analysis: An Approach Based on L-Moments***, Cambridge University Press.

National Weather Service, 1972: ***National Weather Service Observing Handbook No. 2: Substation Observations***, First Edition, Revised December 1972 (Supersedes Circular B), U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, Silver Spring, MD.

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." ***Bulletin of the American Meteorological Society***, vol. 73, pp. 753-762.

: