

National Climatic Data Center

DATA DOCUMENTATION

FOR

DATA SET 3200 (DSI-3200)

SURFACE LAND DAILY COOPERATIVE SUMMARY OF THE DAY

March 18, 2003

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

Table of Contents

Topic	Page Number
1. Abstract.....	3
2. Element Names and Definitions:	4
3. Start Date.....	13
4. Stop Date.....	13
5. Coverage.....	13
6. How to order data.....	14
7. Archiving Data Center.	14
8. Technical Contact.....	14
9. Known Uncorrected Problems.....	14
10. Quality Statement.....	14
11. Essential Companion Data Sets.....	14
12. References.....	14
Appendix A.....	15
Appendix B.....	16

1. **Abstract:** This data set has been used in countless climatological studies, legal litigations, insurance claims, and various other research applications. It contains various parameters consisting of the previous days maximum and minimum temperatures, snowfall, and 24-hour precipitation totals that are initially obtained from state universities, state cooperatives, and the National Weather Service. Currently, there are approximately 8,000 active stations, known as cooperative observing stations with cooperative observers, but data are in these files for approximately 23,000 stations for various years. Selected Summary of the Day data from related file DSI-3210 for National Weather Service "first order" or principal climatological stations and "second order" stations have also been included in this file.

The period of record and number of stations varies among the states. Most states began collecting data during 1948, but some began in 1946. Prior to 1948, most of these data were collected through cooperative agreements with state universities and other state organizations. Many years of records were digitized with some going as far back as the 1850s. It must be noted that NCDC has the observations from the time the station opened, but the [NWS](#) has the current data. Official surface weather observation standards can be found in the [Federal Meteorological Handbook](#).

Stations

The DSI-3200 database is comprised primarily of stations in the National Weather Service (NWS) cooperative station network. The vast majority of the observers are volunteers (non-paid, private individuals). However, the cooperative network also includes the NWS principal climatological stations, which are operated by highly trained observers. The NWS cooperative station network also includes stations that are supported by other federal agencies (e.g., the Department of Interior and Department of Transportation). Commonly the observers at these stations are employees of the Federal Aviation Administration (FAA), National Park Service (NPS), Bureau of Land Management (BLM), U.S. Forest Service (USFC), U.S. Geological Survey (USGS), and Tennessee Valley Authority (TVA). A handful of stations in the DSI-3200 database are stations of the U.S. Department of Defense (DoD) (e.g., Air Force and Navy bases).

The observing equipment used at all of the stations, whether at volunteer sites or federal installations, are calibrated and maintained by NWS field representatives, Cooperative Program Managers (CPMs) and Hydro-Meteorological Technicians (HMTs).

Years ago, stations were selected according to a grid plan. Today stations are selected if another station closes or a different requirement exists for a specific area. When a previously established station closes, a new station compatible to it is established within 5 miles when possible. There are no specifics for distribution of stations, other than an attempt to give a reasonable sampling of the entire region. Currently there are approximately 8,000 active stations in the data set. Historically, approximately 23,000 stations are included for various years.

Elevations for fixed surface locations for the data set are mostly below 1,000 meters above sea level. The minimum elevation is -30 meters and the maximum is 3,000 meters.

:
:

The accuracy of the maximum-minimum temperature system (MMTS) is +/- 0.5 degrees C, and the temperature is displayed to the nearest 0.1 degree F. The observer records the values to the nearest whole degree F. A Cooperative Program Manager calibrates the MMTS sensor annually against a specially-maintained reference instrument.

Observations

The primary intent of the cooperative observing program today is the recording of 24-hour precipitation amounts, but about 55% of the stations also record maximum and minimum temperatures. These observations are for the 24-hour period ending at the time of observation and might be called Summary of the Day observations.

The vast majority of observation times are near either 7:00 a.m. or 7:00 p.m. local time. Principal climatological stations are usually fully instrumented, and therefore record a more complete range of meteorological parameters than other stations. These observations are usually for the 24-hour period midnight to midnight. Observation times do vary however, because of special program needs, and because many of the observers are unpaid volunteers with their own scheduling needs. Ideally, all observations should be taken at the same time, but it not possible to establish and enforce such requirements. The next best thing is to have adequate and accurate information about the observing practices at each station. Even this is an elusive target. Volunteer observers occasionally change their observing schedule without notifying either the NWS or NCDC. In some instances precipitation observations are taken in the morning and temperature observations in the evening, or vice versa. It must be noted that all observations for precipitation amount and temperature max and min are from the previous 24 hour period. Users should also be aware of a potential for a time lapse between a change of observation times at the observing site and the first appearance of the time change in NCDC records (that is, several months may be archived under an "old" observation time before NCDC received notification).

The DSI-3200 data itself contains certain observation-time aberrations: Observations through June 1967 originally carried only the designation "AM" or "PM"; these were later set to 06 or 18 respectively. (The actual hour is shown in the Climatological Data bulletin.) From July 1967 through December 1981, all observation times are shown as 18 (and in fact the majority are p.m. observations. Users must take these facts into consideration. From January 1982 onward, the official time of observation is carried with the data.

2. Element Names and Definitions: DSI-3200 data are archived in a variable length element file structure. The element file structure is designed to allow maximum flexibility in requesting data. Only those elements or groups of elements of particular interest need be ordered.

A DSI-3200 record contains one month of daily data of one data type for one weather station. The record consists of two distinct portions:

The first portion, consisting of the first 30 characters of the record, is the ID section. The second portion, consisting of the rest of the record, is the data. These portions are discussed in more detail below.

:
:

In general, there is only one record with a particular station, data type, year and month. Exceptions are the soil temperature data types and the "days with weather" data type, in which there may be multiple records.

The record format lends itself to variable-length records, and that is one way the user may receive it. Data may also be received in a fixed length record structure. Data may also be received in a fixed length record structure described below.

- a. COBOL Data Description (1 example)
- b. FORTRAN Data Descriptions (2 examples)
- c. Control Language Notes
- d. List of Variables ("Elements") (Described below)
- e. Schematic Variable Length Record Format Layout (Described below)

List of Variables

<u>ELEMENT</u>	<u>WIDTH</u>	<u>POSITION</u>
001 RECORD TYPE (= DLY)	3	001-003 --
002 STATION ID	8	004-011
003 METEOROLOGICAL ELEMENT TYPE	4	012-015
004 MET. ELEMENT MEASUREMENT UNITS CODE	2	016-017 -- ID
005 YEAR	4	018-021 <u>PORTION</u>
006 MONTH	2	022-023
007 FILLER (= 9999)	4	024-027
008 NUMBER OF DATA PORTIONS THAT FOLLOW	3	028-030 --
009 DAY OF MONTH	2	031-032 --
010 HOUR OF OBSERVATION	2	033-034
011 SIGN OF METEOROLOGICAL ELEMENT VALUE	1	035 -- DATA
VALUE OF METEOROLOGICAL ELEMENT	5	036-040 PORTION
QUALITY CONTROL FLAG 1	1	041
QUALITY CONTROL FLAG 2	1	042 --
DATA GROUPS IN THE SAME FORM AS ELEMENT	12	043-054 DATA PORT.
POSITIONS 31-42 REPEATED AS MANY	12	055-066 DATA PORT.
TIMES AS NEEDED TO CONTAIN ONE MONTH	12	067-078 DATA PORT.
OF RECORDS.
608	12	1219-1230 DATA PORT.

Format (Variable Length Record Layout)

1. The first eight elements (positions 001-030) constitute the ID PORTION of the record and describe the characteristics of the entire record. The next six elements, the DATA PORTION of the record contains information about each meteorological element value reported. This portion is repeated for as many values as occur in the monthly record.

2. Each logical record is of variable length with a maximum of 774 characters. (In the FORTRAN and COBOL examples, a larger number of characters may be allowed for.) Each logical record contains a station's data for a specific meteorological element over a one month interval. The form of a record is:

:
:

ID PORTION (30 Characters) Fixed length

```

*****
* REC | STATION | ELEM |   |   |   |   | NUM >
* TYP |   ID   | TYPE | UNT | YEAR | MO | FILL | VAL >
*****|*****|*****|*****|*****|*****|*****|*****
* XXX | XXXXXXXX | XXXX | XX  | XXXX | XX  | XXXX | XXX >
*****
ELEMENTS 001      002      003      004      005      006      007      008

DATA PORTION (12 Characters, repeated "NUM-VAL" times--up to 100)

*****
< DY | HR | MET. ELEM | FL | FL | DY | HR | MET. ELEM | FL | FL >
<   |   |   |   | 1 | 2 |   |   |   | 1 | 2 >
<   |   | *****|   |   |   |   | *****|   |   >
<   |   | DATA |   |   |   |   | DATA |   |   >
<   |   | S | VALUE |   |   |   |   | S | VALUE|   |   >

<*****|*****|*****|*****|*****|*****|*****|*****|*****| *****|*****|*** >
< XX | XX | X | XXXXX | X | X | XX | XX | X | XXXXX | X | X >
*****
ELEMENTS 009 010 011 012 013 014 015 016 017 018 019 020

*****
< DY | HR | MET. ELEM | FL | FL *
<   |   |   |   | 1 | 2 *
<   |   | *****|   |   *
<   |   | DATA |   |   *
<   |   | S | VALUE |   |   *
<*****|*****|***|*****|*****|*****|*****|*****
< XX | XX | X | XXXXX | X | X *
*****
ELEMENTS 603 604 605 606 607 608

```

3. The Number of Data Portions (position 008) for the logical record of type "DLY" ranges from 1 to 62.

In addition to a variable length record structure, users may also receive data in a fixed length record structure. However, the user must specify whether to extract either the original or edited data values. Supplied data are in the same sort as archived data.

Provided within this section are information and examples of how to access the fixed length data records, specifically:

- a. COBOL Data Description
- b. FORTRAN Data Description
- c. List of Variables ("Elements")
- d. Schematic Fixed Length Record Format Layout

The type of meteorological elements stored in this record and range of values are listed below.

DYSW - The different types of weather occurring that day (reference Table "C"; See topic 46 "Data Quality: Confidence Factors").

:
:
6:

EVAP - Daily evaporation (not reported when temperature below freezing). Unit Measurement, Inches & Hundredths of Inches.

MNPN - Daily minimum temperature of water in an evaporation pan (effective September 1963). Unit Measurement, Whole Degrees Fahrenheit.

MXPN - Daily maximum temperature of water in an evaporation pan (effective September 1963). Unit Measurement, Whole Degrees Fahrenheit.

PRCP - Daily precipitation. (Precipitation reading for 24 hours ending at time of observation. Trace is less than 0.005 inch. Unit Measurement, Inches to Hundredths.

SNOW - Daily Snowfall (Snowfall includes sleet). Amount is for 24-hour period ending at observation time. Hail was included with snowfall from July 1948 through December 1955. Hail occurring alone was not included with either snowfall or snow depth before and after that period. Trace is less than 0.05 inch. Unit Measurement, Inches to Tenths.

SNWD - Snow depth at observation time. (Snow depth is depth of snow on the ground at time of observation. Trace is depth less than 0.5 inch.) Unit Measurement, Whole Inches.

TMAX - Daily maximum temperature. (Maximum temperature reading for 24 hours ending at time of observation.) Unit Measurement, Whole Degrees Fahrenheit.

TMIN - Daily minimum temperature. (Minimum temperature reading for 24 hours ending at time of observation.) Unit Measurement, Whole Degrees Fahrenheit.

TOBS - Temperature at observation time. Unit Measurement, Whole Degrees Fahrenheit.

WDMV - 24-hour wind movement. Unit Measurement, Whole Miles.

WTEQ - Water equivalent of snow depth. (For principal stations only. Effective October 1963 for snow depth equal or greater than 2 inches). Unit Measurement, Inches to Tenths.

SNyz - Daily minimum soil temperature (see note below). Unit Measurement, whole degrees Fahrenheit.

SOyz - Soil temperature at observation time (see note below). Unit Measurement, whole degrees Fahrenheit.

SXyz - Daily maximum soil temperature (see note below). Unit Measurement, whole degrees Fahrenheit.

Soils Note 1: Even though TD-3200 is daily data, many cases of two or more soil temperature measurements per day, at different times, can be found.

Soils Note 2: Positions "y" and "z" of the soil temperatures are encoded using reference Table "D", e.g., SN12 indicates that the daily minimum soil temperatures that follow are measured in an area covered with grass and at a depth of four inches or 10 centimeters.

:
:

METEOROLOGICAL ELEMENT MEASUREMENT UNITS CODE - The units and decimal position (precision) of the data value for this record (reference Table "E"). See topic "Known Uncorrected Problems" for additional details.

YEAR - This is the year of the record. Range of values is 1850-current year processed.

MONTH - This is the month of the record. Range of values is 01-12 LST.

FILLER - Filler value is 9999.

NUMBER OF DATA PORTIONS THAT FOLLOW

The number of data values in this record. Range of values is 01-62.

NOTE: A record may contain fewer or more data values than you might expect.

A maximum of two DATA PORTIONS are used for each day of the month so as to allow one original data value and one edited data value. At most, 62 data values may be contained in any given record (e.g., $30 + (62 \times 12) = 774$ characters). Thus, while a maximum of 1,230 characters has been assigned in some of the program examples, no more than 774 characters will be used for the daily data record types.

If a particular data value was not taken or is unavailable, there is no entry for it. For meteorological elements observed once a day, if all the daily observations of a given month are received and pass QC checks, there will be one DATA PORTION for each day. And, if every value failed the quality control, there may be two DATA PORTIONS for every day of that month. When two DATA PORTIONS for a daily record are encountered (with the exception of DYSW), one finds that the original data values are flagged and the second DATA PORTION is the best available replacement. (See code definitions for the Flag 2 element).

DAY OF MONTH - Contains the day of the month on which an observation was made. Range of values is 01-31.

HOURLY OF OBSERVATION - Contains the hour of the daily observation. Hour of observation is reported using the 24-hour clock with values ranging from 00-23 LST, except in the cases of soil temperatures element-type (where the hour is 99 to indicate missing) and "days with weather" (where the hour is 24). Through June 1967 observations were designated as "AM" or "PM"; these values were set to 06 or 18 respectively during a conversion of earlier data to DSI-3200. From July 1967 through 1981, all observations were set to hour 18 (because the majority are p.m. observations). Beginning January 1982, the actual hour of the observation is available and is indicated.

SIGN OF METEOROLOGICAL VALUE - The algebraic sign of the meteorological data value is given as either a blank or a minus sign (-). Blank indicates a positive value and a minus sign represents a negative value.

VALUE OF METEOROLOGICAL ELEMENT - The actual data value is given as a five-digit integer. One major exception does exist however, for the DYSW (days with weather code) element-type values as explained in Table "C".

NOTE: For fixed length records only - when a data value is missing, the sign of the data value is set to "-", the data value is set to "99999", flag

:
:

position 1 is set to "M" and flag position 2 is blank. For variable-length records, the minus sign is omitted for any such values.

A special case of missing data occurs in the precipitation data types when, for one or more days, the observer made no measurements. On a later day, the observer resumed measurements. The first new measurement showed the accumulated total of precipitation during all of the missing days. This case is represented in TD-3200 data by observations for two days: The first day is the first on which no measurement was made. It has a data value of "99999" (missing) and has "S" in flag 1. The second day is the day of the first new measurement. The data value of the second day is the accumulated total and has "A" or "B" in flag 1.

FLAG1 - The Data Measurement FLAG (reference Table "F").

FLAG2 - The Data Quality FLAG (reference Table "G").

:
:

TABLE "A"

State-Code Table

01 Alabama	28 New Jersey
02 Arizona	29 New Mexico
03 Arkansas	30 New York
04 California	31 North Carolina
05 Colorado	32 North Dakota
06 Connecticut	33 Ohio
07 Delaware	34 Oklahoma
08 Florida	35 Oregon
09 Georgia	36 Pennsylvania
10 Idaho	37 Rhode Island
11 Illinois	38 South Carolina
12 Indiana	39 South Dakota
13 Iowa	40 Tennessee
14 Kansas	41 Texas
15 Kentucky	42 Utah
16 Louisiana	43 Vermont
17 Maine	44 Virginia
18 Maryland	45 Washington
19 Massachusetts	46 West Virginia
20 Michigan	47 Wisconsin
21 Minnesota	48 Wyoming
22 Mississippi	49 Not Used
23 Missouri	50 Alaska
24 Montana	51 Hawaii
25 Nebraska	66 Puerto Rico
26 Nevada	67 Virgin Islands
27 New Hampshire	91 Pacific Islands

TABLE "B"

Cooperative Network Division Table

NOTE: The division number for a station may change over time.

HAWAII (STATE 51)

ISLAND NAME	DIVISION
Kauai	01
Oahu	02
Molokai	03
Lanai	04
Maui	05
Hawaii	06

PACIFIC ISLANDS (STATE 91)

Division:

02 - East of 180th Meridian - Phoenix Islands, Line Islands, American Samoa
03 - Western Pacific Islands, North of 12N
04 - Caroline and Marshall Islands

:

:

10:

TABLE "C"

DYSW - Daily Occurrence of Weather Tables

POR = Period of Record.

- 00 - Day of no occurrence
- 01 - Day with smoke or haze (POR through 1963 and 1982 to Present)
- 02 - Day with fog (POR through 1963 and 1982 to Present)
- 04 - Day with drizzle (POR through 1963 and 1982 to Present)
- 05 - Day with ice pellets (sleet)
- 06 - Day with glaze
- 07 - Day with thunder
- 08 - Day with hail
- 09 - Day with dust or sand storm (POR through 1963 and 1982 to Present)
- 10 - Day with blowing snow
- 11 - Day with high wind (POR through 1963 and 1982 to Present)
- 12 - Day with tornado (POR through 1963 and 1982 to Present)
- 13 - Day with rain (1982 to Present)
- 14 - Day with snow (1982 to Present)

Note: These two-character DYSW element-type codes are stored into the rightmost four characters of the data value portion of the meteorological element. The leftmost character is always zero (0). Within the four characters used, the weather codes are entered left justified. Thus, if one type of weather occurs during a day, the data values would appear as OXX00, where XX is the appropriate weather code. If two types of weather occur, the data value will contain OXXYY, where XX is the first code and YY is the second. If more than two types of weather occur on the same day, they will be stored into additional records of the element-type code "DYSW" as needed.

TABLE "D"

Soil Temperature Table

(y = Code for soil cover) (z = Code for soil depth)

```

*****
|Code| Cover      | |Code  |  Depth  | Depth  |
|  |  |  |  |  (inches) | (cm)  |
|*****|*****|*****|*****|*****|
|y=1 | Grass      | |      |          |        |
| 2 | Fallow     | |z = 1 | 2        | 5      |
| 3 | Bare       | |      |          |        |
|  | ground    | | 2    | 4        | 10     |
| 4 | Brome     | |      |          |        |
|  | grass     | | 3    | 8        | 20     |
| 5 | Sod        | |      |          |        |
| 6 | Straw     | |      |          |        |
|  | mulch     | | 4    | 20       | 50     |
| 7 | Grass     | |      |          |        |
|  | muck      | | 5    | 40       | 100    |
| 8 | Bare       | |      |          |        |
|  | muck      | | 0    | Unknown  | Unknown|
| 0 | Unknown   | |      |          |        |
*****

```

:
:

NOTE: Soil records are kept since 1982. Some stations may report soil temperatures at observation time twice a day. Separate records will occur for both observation times.

TABLE "E"

Units of Measurement Table

Range of values where b = Blank:

bF	Whole degrees Fahrenheit (right justified)
HI	Hundredths of inches
bI	Whole inches (right justified)
bM	Whole miles (right justified)
NA	No units applicable (nondimensional)
TI	Tenths of inches

TABLE "F"

Data Measurement Flag 1

- A - Accumulated amount since last measurement.
- B - Accumulated amount includes estimated values (since last measurement).
- E - Estimated (see Table "G" for estimating method).
- J - Value has been manually validated.
- M - For fixed length records only. Flag1 is "M" if the data value is missing. In this case, the sign of the meteorological value is assigned "-" and the value of the meteorological element is assigned "99999".
- S - Included in a subsequent value. (data value = "00000" OR "99999").
- T - Trace (data value = 00000 for a trace).
- (- Expert system edited value, not validated.
-) - Expert system approved edited value.
- Blank - Flag not needed.

Notes:

An original observation that is followed by an edited observation is indicated by "2" in Flag 2.

Flag 1 values of "S" and "A" occur in pairs (i.e. a daily value will have "S" in Flag 1 and the next daily value will have "A" or "B" in Flag 1). These flags have traditionally been associated with PRCP and SNOW data, but can occur in any of the data types in which the data may be cumulative from day to day.

:
:

TABLE "G"

Data Quality Flag 2

- 0 - Valid data element.
- 1 - Valid data element (from "unknown" source, pre-1982).
- 2 - Invalid data element (subsequent value replaces original value).
- 3 - Invalid data element (no replacement value follows).
- 4 - Validity unknown (not checked).
- 5 - Original non-numeric data value has been replaced by its deciphered numeric value.
- A - Substituted TOBS for TMAX or TMIN.
- B - Time shifted value.
- C - Precipitation estimated from snowfall.
- D - Transposed digits.
- E - Changed units.
- F - Adjusted TMAX or TMIN by a multiple of + or -10 degrees.
- G - Changed algebraic sign.
- H - Moved decimal point.
- I - Rescaling other than F, G, or H.
- J - Subjectively derived value.
- K - Extracted from an accumulated value.
- L - Switched TMAX and/or TMIN.
- M - Switched TOBS with TMAX or TMIN.
- N - Substitution of "3 nearest station mean".
- O - Switched snow and precipitation data value.
- P - Added snowfall to snow depth.
- Q - Switched snowfall and snow depth.
- R - Precipitation not reported; estimated as "0".
- S*- Manually edited value.
- T - Failed internal consistency check.
- U - Failed areal consistency check (beginning Oct. 1992).

* Manually edited value could be derived by any procedure - Procedure unspecified.

3. Start Date: Data from 1948 onward comprise the bulk of the data set. This is when punch cards were used to help summarize climatological data. Data from earlier periods were placed upon punch cards as a result of various agreements with state universities.

Meteorological Element-Type

Daily Temperature	1850
Daily Precipitation	1862
Daily Evaporation	1951
Soil Data	1982

4. Stop Date: Ongoing.

5. Coverage: the USA, Alaska, Hawaii, US Pacific Islands, Puerto Rico and the US Virgin Islands

a. Southernmost Latitude	15S
b. Northernmost Latitude	75N

:
:

c. Westernmost Longitude 130E
d. Easternmost Longitude 60W

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: Users should also be aware of a potential for a "lag" in the change of observation times used in the logical record and what is actually in practice at the site (that is, several months may be archived digitally under an "old" observation time before NCDC received notification).

10. Quality Statement: These data have received a high measure of quality control through computer and manual edits. These data are subjected to internal consistency checks, compared against climatological limits, checked serially, and evaluated against surrounding stations. Quality control "flags" are appended to each element value to show how they fared during the edit procedures and to indicate what, if any, action was taken. The historical data prior to 1982 were converted from existing files then placed in the element file structure format after being processed only through a gross value check. In November 1993 the entire historical period of record was processed through a stringent quality control. Another round of quality control in November 2000 increased the data set's quality still more. The DYSW "days with weather" element data, which is fair to poor, depending on the station, cannot be used with any measure of confidence. Principal climatological stations operating 24 hours a day are expected to be the most reliable source of "days with weather". Reporting of this element by cooperative observers is not a requirement and criteria for reporting is not definitive. Most stations do not record this information.

11. Essential Companion Datasets: DSI-3200 requires use of NCDC's in-house Station History file.

12. References:

National Weather Service, 1987: Cooperative Program Management, Weather Service Operations Manual B-17 (revised), NOAA-NWS, Silver Springs, MD.

:
:

Reek T., S. R. Doty, and T. Owen, 1991: ValHiDD Documentation and Users Guide, Internal NCDC document, 20 pp.

APPENDIX A:

Program 3200fix_1.c
/*

WHAT THIS PROGRAM DOES:

This program is the first of two that make corrections to NCDC DSI-3200 data. The program itself is generic "C".

COMPILE AND RUN INSTRUCTIONS:

The source code file is "3200fix_1.c".

A suggested compile command for the SUN Solaris UNIX environment is

```
cc -O -o 3200fix_1 3200fix_1 -lm
```

(-O and -o are upper- and lower-case letters.) This command creates an optimized program named "3200fix_1".

To run the program: Place all input files in one directory. Make that directory the current directory. Then type in

```
<path>3200fix_1
```

where <path> is the location of the program. <Path> can be omitted if the program is located in the current (input files) directory.

DETAILS:

Five types of files are involved in running this program. All but (1) are ASCII text.

- (1) The program itself, "3200fix_1".
- (2) DSI-3200 input files (in NCDC's standard element format for DSI-3200).
- (3) One output data file per input file.
- (4) One output file that is a summary log of changes.
- (5) One output file that is a detailed log of changes.

Referring to (2): "3200fix_1" will attempt to open, as input files, all ASCII files in the current directory whose names do NOT contain any periods (.). Files that are directories or programs will be ignored, regardless of name. Records may be fixed or variable length; it makes no difference.

Referring to (3): One output file is created for each input file. Each output file name is the same as the input file name, except that the extension ".nd" is appended. Output files are variable length ASCII.

Referring to (4): One summary log file named "3200log.summary1" is created. The file shows the number of records read, deleted, modified, swapped and output for each input file and station.

:
:

Referring to (5): One detailed log file named "3200log.details1" is created. It shows each modified record before and after the changes. For deleted records, it shows the record before deletion. Because it contains record images, it is variable length ASCII like (3).

OTHER PROGRAM INFORMATION:

Functions appear in alphanumeric order in the "Function Prototypes" section.

Variable names try to suggest what the variable represents. But names longer than a dozen characters are rare and local work variable names are usually quite short, often 1 or 2 characters. (The author's prejudice is that overly long variable names are more clumsy than helpful.)

Variable names are composed of some combination of letters, underscores, and numbers. Underscores and numbers may occur in any name. Letters occur in all names. Letter case is used to indicate the type of variable.

- a. All letters are upper case in global definition names and macro names.
- b. An upper case 'G' is always the first character of a global variable name. Other letters are always lower case.
- c. All letters in function names and in names of local variables are lower case.

A few words about pointers and copy-overwrites, methods heavily used in this program, can be found at the end of this file.

APPENDIX B:

Program 3200fix_2.c
/*

WHAT THIS PROGRAM DOES:

This program is the second of two that make corrections to NCDC TD-3200 data. The program itself is generic "C".

COMPILE AND RUN INSTRUCTIONS:

The source code file is "3200fix_2.c".

A suggested compile command for the SUN Solaris UNIX environment is

```
cc -O -o 3200fix_2 3200fix_2 -lm
```

(-O and -o are upper- and lower-case letters.) This command creates an optimized program named "3200fix_2".

To run the program: Place all input files in one directory. Make that

:
:

directory the current directory. Then type in

<path>3200fix_2

where <path> is the location of the program. <Path> can be omitted if the program is located in the current (input files) directory.

DETAILS:

Five types of files are involved in running this program. All but (1) are ASCII text.

- (1) The program itself, "3200fix_2".
- (2) DSI-3200 input files (in NCDC's standard element format for DSI-3200).
- (3) One output data file per input file.
- (4) One output file that is a summary log of changes.
- (5) One output file that is a detailed log of changes.

Referring to (2), input file names must be of the form "*.nd" where "*" is any group of characters that make a legal file name in your computer's operating system. The extension ".nd" is required.

Referring to (3), an output file is created for each input file. Each output file name is the same as the input file name, except the extension is ".nd2". Output files contain variable-length records.

Referring to (4), one summary log file named "3200log.summary2" is created. It shows the number of records read, deleted, modified, and output for each input file and station.

Referring to (5), one detailed log file named "3200log.details2" is created. It shows each modified record before and after the changes. For deleted records, the original record is shown. Records with no changes do not appear in this log file.

OTHER USEFUL INFORMATION:

Function prototypes appear in alphanumeric order in the "Function Prototypes" section.

Functions appear in the following order in the code itself:

- (1) "Main()".
- (2) Specialized functions in alphanumeric order.
- (3) General-purpose functions in alphanumeric order.

The phrase "significant obs" or "significant observation" is often seen in the comments. The basic meaning of this is "an edited obs, or an original obs without an edited obs following". Conversely, a "non-significant obs" is "an original obs with an edited obs following". Sometimes, a "significant obs" must also not be error-flagged.

:
:

Variable names usually suggest what the variable represents. Names longer than a dozen characters are rare. Work/scratch variable names are usually quite short, often 1 letter.

Names are composed of letters, underscores, and numbers. Letter case is significant:

- a. Global variable names always begin with an upper case letter, often a 'G'. Any other letters in the names are lower case.
- b. Global definition names and macro names never contain lower case letters.
- c. Ordinary variable names and function names never have upper case letters.

More about pointer arithmetic and character string copying:

One form of character string copying that is much used in this program is the copy-overwrite, in which pointer arithmetic is rampant. Typical are lines of code like

```
(1) mystrcpy(a, a+12) ;  
(2) mystrcpy(a+12, a) ;
```

where "mystrcpy()" is a general-purpose function in this program. "a" and "a+12" are character addresses inside the same long character string, whose starting address is typically stored in some other variable that is not mentioned; "a+12" is 12 bytes to the right (toward higher addresses) of "a". "a-12" is 12 bytes to the left (toward lower addresses) of "a".

Code line (1) copies, to "a", the character substring that begins at "a+12" and ends after the long string's NULL character. In effect, the 12 bytes between "a" and "a+11" are lifted out of the long string and discarded; and the right end of the long string is slid to the left to fill the gap. The long string is 12 bytes shorter. This trick removes a 12-byte observation in a DSI-3200 record.

Code line (2) copies, to "a+12", the character substring that begins at "a" and ends after the long string's NULL character. In effect, the right end of the long string, starting at "a+12", is slid to the right, leaving a 12-byte gap; and the 12 bytes between "a" and "a+11" are duplicated in the gap. The long string is 12 bytes longer. A new 12-byte observation in a DSI-3200 record can be copied into the gap.

:
: