Station File Formats for *mloc*

Eric Bergman
Global Seismological Services
September 4, 2017

This document describes the formats for use for station data files in the multiple event relocation program *mloc*. There is a "standard" format designed specifically for *mloc* which is used for the master station file, but for convenience and backward compatibility, several other formats are supported for use as supplemental station data files (command "sstn"). The format used for the master file has changed several times during development of *mloc*. The current master station file format was introduced with *mloc* v10.0.0, released on April 28, 2014.

Format Index

The first line of all station data files must be a header line that carries a 1-digit integer (*isstn*) in the first column that defines the format of the data file. *mloc* reads this digit and branches to different sections of code to read and process the appropriate format. The rest of the header line (up to 96 characters total for the line) is considered an optional comment, which can be used to describe the dataset.

The current list of defined values for *isstn* is given in the following table:

ISSTN	Description
0	Master station data file format
1	ISC FFB format (geographic coordinates in deg-min-sec*10, elevation in m)
2	SEISAN station list, with station code starting in column 3. Degree-minute-decimal seconds.
3	Simplified format, decimal degrees, geographic coordinates. A basic format for putting in supplemental station data, supports 6-character station codes.
4	China Seismic Bureau format
5	NEIC format
6	MSU format (Kevin Mackey)
7	undefined
8	undefined
9	Previous master station file format

Details of each format can be found in subroutine "redsta" in mloclib_stations.f90.

Master Station File Format (*isstn* = 0)

The master station file used by mloc is keyed to the International Registry of Seismograph Stations (IR) maintained at the ISC. Specifically, it should contain only station codes that have been registered in the IR, even if the details (e.g., coordinates) have been altered (corrected). To use a station with a code that has not been registered, a supplemental station data file should be declared (command sstn) in *mloc*. A supplemental station file can use the master station file format.

The format used for the master station file is unique to *mloc*. It departs from standard practice in permitting multiple entries for the same station code. There are two main reasons for this complexity:

1. In some cases stations have been moved without changing the station code. The format includes the concept of epochs that define when a station occupied the given location. Therefore it is necessary to check the date for which coordinates of a given station code are needed to ensure that the correct coordinates are used.
2. The coordinates carried in the IR are sometimes found to be in error, for example, by investigation with Google Earth. Rather than delete the IR coordinates, it is preferable to put the improved information in as a new entry for the same epoch, thus documenting the discrepancy. The concept of authorship of an entry also helps in this regard.

The order in which "duplicate" entries in the master station file are listed controls which coordinates are used: **the first entry that matches the station code and the operational epoch will be used**.

Station Codes

The master station file format carries a maximum of 5 characters for station code, following the standard for the IR. *mloc* can accommodate 6-character station codes, but information about such stations must be provided in a supplemental station data file in a format that allows 6 character codes, for example, the Simplified format (*isstn* = 3) defined below.

Operational Epoch

The period of time during which a set of coordinates is valid for a given station code is called the "operational epoch". It is defined by two variables, *date_on* and *date_off*, which are given as a 7-digit integer composed of the year (left-most four digits) and Julian date (right-most three digits). Either one or both may be left blank, with the obvious meaning.

Seeding of the operational epochs for many codes in the master station file was based on an analysis of data holdings at Lawrence Livermore National Laboratory by Steve Myers,

supplemented by information acquired by Bob Engdahl. I have found that there are many errors in these epoch data, especially in the *date_off* field, such that *mloc* often reports a "date range" failure in processing actual arrival time data against the master station file. In the case of erroneous *date_off* entries the problem is usually that Steve filled in the field with the date of the most recent data holding at that time, but obviously most stations have continued to operate past that point; therefore the correct course of action is to delete the *date_off* entry completely. In the case of violations of the *date_on* entry, the preferred course of action is usually to update the *date_on* value to the earliest date of the available data.

The definition of the SEISAN station file format (*isstn* = 2) has been extended to include optional *date_on* and *date_off* fields, because this format is often used for temporary networks that had a limited period of operation, and the station codes for such deployments often conflict with registered stations.

Authorship

I have introduced the notion of authorship to the master station data file to help document discrepancies in reported station coordinates that are often encountered. The author of a set of coordinates is carried in a character variable with 8 characters, the same as authorship for hypocenters and phase readings in the ISF format and the MLOC Native Format for arrival time data. There are no standard definitions for these authorship entries, but I have used "IR" as the code for entries downloaded from the International Registry, the base set of coordinate information for the master station file. Definitions of authorship codes can be carried in the beginning of the master station file in lines that begin with '#' in column 1 (comment lines).

A standard variation on authorship codes is the use of the '+' character to indicate a different author for the information on station elevation and depth of burial (if given) than the author appropriate for the latitude and longitude. This is useful because the IR in many cases does not carry any information on station elevation. A very easy way to rectify this lapse is to enter the coordinates in Google Earth and take the elevation from that, in which case the authorship code would be amended from "IR" to "IR+GE". In some cases, where the seismic vault can be clearly identified in Google earth, I have used the authorship code "EAB+GE" to indicate that I specified all coordinates of the station using Google Earth.

Station Elevation and Depth of Burial

The master station file format follows the IR practice of carrying two values related to the vertical coordinate of a station. Elevation is the elevation of the ground surface at the latitude and longitude provided, and depth of burial defines how far below the surface the sensor is emplaced. Both values are given in meters. Elevation can be a negative value, but depth of burial is always positive. Both values follow the IR in assuming the reference plane to be the reference geoid WGS84.

In reality, depth of burial information is often missing in the IR (and therefore, in the master station file) and in some cases the elevation field has been filled with the elevation at the sensor. It is very difficult and time-consuming to correct these errors, but in most cases the discrepancy in elevation will be on the order of 100 m or less, which is usually not very important in the context of earthquake location. If corrections to depth of burial and/or elevation are made, the authorship code should be modified accordingly.

Format

The first line must contain the digit "0" in column 1, to indicate the format. Comment lines, indicated by '#' in column 1, can be included anywhere in the file. All other lines in the file have the following format:

Column	Description
1:5	Station code (a5)
7:15	Latitude (f9.5)
17:26	Longitude (10.5)
28:32	Elevation, m (i5)
34:37	Depth of burial, m (i4)
39:46	Author (a8)
48:52	Agency (a5)
54:61	Deployment (a8)
63:64	Location (a2)
66:72	Date_on (i7)
74:80	Date_off (i7)
82:	Station name (no limit)

Supplemental Station File Formats

Master Station File Format (*isstn* = 0)

See above.

ISC Fixed Format (*isstn* = 1)

This format is based on the record type 91 used in the ISC's old Fixed Format ("96 Byte") data format, extended to support 6-character station codes. It features geographic coordinates in deg-min-sec*10, elevation in m.

Column	Description
15:20	station code (a6)
62:63	Latitude degrees (i2)
64:65	Latitude minutes (i2)
66:68	Latitude seconds*10 (i3)
69:69	"N" or "S" (a1)
70:72	Longitude degrees (i3)
73:74	Longitude minutes (i2)
75:77	Longitude seconds*10 (i3)
78:78	"E" or "W" (a1)
79:82	Elevation, m (i4)

SEISAN Station Format (*isstn* = 2)

The standard form of station data produced by the SEISAN software. Limited to 4-characters for station code. Geographic coordinates in degrees-decimal minutes. Elevation in m. The fields *date-on* and *date_off* are optional.

Column	Description
3:6	Station code (a4)
7:8	Latitude degrees (i2)
9:13	Latitude minutes (f5.2)
14:14	"N" or "S" (a1)
15:17	Longitude degrees (i3)
18:22	Longitude minutes (f5.2)
23:23	"E" or "W" (a1)
24:27	Elevation, m (i4)
34:40	Date_on (i7)
42:48	Date_off (i7)

Simplified Format (*isstn* = 3)

This is a basic format designed for quick entry of a few supplemental stations. Geographic coordinates in decimal degrees. Supports 6-character station codes, operational epoch and both station elevation and depth of burial (positive number). Elevation, depth of burial and operational epoch are all optional.

Column	Description
1:6	Station code (a6)
8:15	Latitude (f8.4)
17:25	Longitude (f9.4)
27:31	Elevation, m (i5)
33:37	Depth of burial, m (i5)
39:45	Date_on (i7)
47:53	Date_off (i7)

China Seismic Bureau Format (*isstn* = 4)

A format used by the China Seismic Bureau for their station lists. Uses a 3-character, lower-case station code. Only defined for positive latitude and longitude, but could be easily extended.

Column	Description
1:3	Station code (a3)
5:8	Elevation, m (i4)
10:11	Latitude degrees (i2)
14:15	Latitude minutes (i2)
18:21	Latitude seconds (f4.1)
25:27	Longitude degrees (i3)
30:31	Longitude minutes (i2)
34:37	Longitude seconds (f4.1)

NEIC Format (*isstn* = 5)

This format is for use with station coordinate data returned from the NEIC metadata server. If you use arrival time data from the NEIC ComCat server, especially for events within the U.S., you will probably need a supplemental station file to handle the many stations in U.S. regional networks that have not been registered with the International Registry at the ISC and also to solve the many station code conflicts that exist with registered stations.

The easiest method of building type 5 supplemental station files is to use the *mdget* utility program written and distributed by Dave Ketchum at NEIC <ketchum@usgs.gov>. The format used here is returned by the basic command *mdget -coord -s* (followed by a network/station regular expression and a *-b* and *-e* options to define a time interval, if desired).

Two forms of the command are especially useful for this purpose. It is necessary first to determine which network's version of a given station code is desired. There are cases of several regional networks using the same code and many cases where a network has used a code that is registered for a different station at the ISC. For example the station TCU is searched by

mdget -a *.*.TCU -b all

returning

```
FDSN.IR.TCU.--:24.1475 120.676 84.0
* <EOE>
FDSN.UU.TCU.--:41.1173 -111.4078 2269.0
* <EOE>
FDSN.UU.TCU.01:41.1173 -111.4078 2269.0
* <EOE>
HOLD.TAP.TCU.--:24.1475 120.676 84.0
* <EOE>
ISC.IR.TCU.--:24.1475 120.676 84.0
* <EOE>
NEIC.TAP.TCU.--:24.1475 120.676 84.0
* <EOE>
UUSS.UU.TCU.--:41.1173 -111.4078 2269.0
* <EOE>
UUSS.UU.TCU.01:41.1173 -111.4078 2269.0
* <EOE>
* <EOR>
```

In this case there is a conflict between TCU in the Utah network (network code UU) and TCU in the International Registry (FDSN.IR.TCU). To obtain the properly formatted version of FDSN.UU.TCU for the supplemental station file, and because if you need one UU station you will probably need others, it is convenient to use a different form of mdget to dump the entire UU network:

mdget -coord -s UU.....Z.. -b all > uu_stn.dat

where we redirect the output to a file. The regular expression has 12 characters:

1-2	Network code
3-7	Station code
8-10	Channel code
11-12	Location code

In the example the character "Z" in the 10th position selects only vertical components, which is usually adequate to obtain the station coordinates. The first few lines returned by this command are:

```
UU AIUT  -- EHZ:     0.0 -90.0    0.0:  40.8558 -112.1755 1334.0 1989-09-28 00:00 to 1990-05-10 23:59
* <EOE>
UU AIUT  -- EHZ:     0.0 -90.0    0.0:  40.8558 -112.1755 1334.0 1987-06-25 00:00 to 1989-09-27 23:59
* <EOE>
UU AIUT  -- EHZ:     0.0 -90.0    0.0:  40.8558 -112.1755 1334.0 1985-06-04 00:00 to 1987-06-24 23:59
* <EOE>
```

If you're going to save a full list for this network it is worthwhile to edit the returned list to remove "end of entry" lines, or the desired line can simply be copied and pasted into the type-5 supplemental station file. Although the output of the *mdget* utility includes operational epoch information, *mloc* does not use it.

Column	Description
4:8	Station code (a5)
40:47	Latitude (f8.4)
49:57	Longitude (f9.4)
58:62	Elevation, m (i5)

MSU Format (*isstn* = 6)

A format used by Kevin Mackey at Michigan State University, especially for data from the former Soviet Union.

Column	Description
1:5	Station code (a5)
6:7	Latitude degrees (i2)
9:10	Latitude minutes (i2)
12:15	Latitude seconds (f4.1)
16:16	"N" or "S" (a1)
17:19	Longitude degrees (i3)
21:22	Longitude minutes (i2)
24:27	Longitude seconds (f4.1)
28:28	"E" or "W" (a1)
30:33	Elevation, m (i4)