SEGY to ASCII Conversion and Plotting Program 2.0

By Mark. R. Goldman[1]

Open-File Report 2005-1311

This Open-File Report supersedes USGS Open-File Repor99-126

This document and any updates to it may be found at:
http://pubs.usgs.gov/of/2005/1311/

U.S. DEPARTMENT OF THE INTERIOR
U.S. GEOLOGICAL SURVEY

[1]Menlo Park, Calif.

CONTENTS AND PROGRAM REQUIREMENTS

Date of Latest Revision: July, 2005

U.S. Department of the Interior
U.S. Geological Survey

Open-File Report 2005-1311

SEGY to ASCII Conversion and Plotting Package 2.0

Mark Goldman
U.S. Geological Survey
MS 977
345 Middlefield Rd.
Menlo Park, CA 94025

This directory contains the source code and shell scripts necessary to convert SEGY files to ASCII format and to create postscript plots, along with a makefile and a README file. This directory also contains a subdirectory with an example SEGY file, ASCII file, shell script file, and resulting postscript file. As a test, one can run the shell script file to recreate the xyz and postscript files.

The shell scripts run using the csh shell command interpreter, which is present on most Unix systems.

A requirement for compiling the source code is a C++ compiler The C++ programs documented here have been successfully compiled using Gnu's g++ version 3.3.2. The g++ compiler is a free C++ compiler and may be downloaded from the web site:
http://gcc.gnu.org

A requirement for plotting the seismic data is the GMT plotting package. The shell scripts documented here have successfully produced postscript files using Gmt version 3.3.4. The GMT plotting package is a free set of plotting routines, and may be downloaded from the web site:
http://gmt.soest.hawaii.edu

INTRODUCTION

SEGY has long been a standard format for storing seismic data and header information. Almost every seismic processing package can read and write seismic data in SEGY format. In the data processing world, however, ASCII format is the "universal" standard format. Very few general purpose plotting or computation programs will accept data in SEGY format.

The software presented in this report, referred to as SEGY to ASCII (SAC), converts seismic data written in SEGY format (Barry et al., 1975) to an ASCII data file, and then creates a postscript file of the seismic data using a general plotting package (GMT, Wessel and Smith, 1995). The resulting postscript file may be plotted by any standard postscript plotting program.

There are two versions of SAC: one version for plotting a SEGY file that contains a single gather, such as a stacked CDP or migrated section, and a second version for plotting multiple gathers from a SEGY file containing more than one gather, such as a collection of shot gathers. Note that if a SEGY file has multiple gathers, then each gather must have the same number of traces per gather, and each trace must have the same sample interval and number of samples per trace.

SAC will read several common standards of SEGY data, including SEGY files with sample values written in either IBM or IEEE floating point format. In addition, utility programs are present to convert non-standard Seismic Unix (.sux) SEGY files and PASSCAL (.rsy) SEGY files to standard SEGY files.

SAC allows complete user control over all plotting parameters including label size and font, tick mark intervals, trace scaling, and the inclusion of a title and descriptive text.

SAC shell scripts create a postscript image of the seismic data in vector rather than bitmap format, using GMT's pswiggle command. Although this can produce a very large postscript file, the image quality is generally superior to that of a bitmap image, and commercial programs such as Adobe Illustrator ® can manipulate the image more efficiently.

LIST OF FILES

PROGRAM NAME	DESCRIPTION
README.txt	this file as a text document
README.doc	this file as a Microsoft Word document
makefile	makefile for use in compilation
splot	template csh shell script for plotting a single gather
splot2	template csh shell script for plotting multiple gathers
splot.body	csh shell script called by splot (do not edit)
splot2.body	csh shell script called by splot2 (do not edit)
segy2xyz	program to convert a SEGY file to ASCII, called by shell script
segy2xyz2	program to convert a particular gather within a multi-gather SEGY file to ASCII, called by shell script
segyheader	utility to list a SEGY file's header values
segyibm2ieee	utility to convert sample values from IBM to IEEE
su2segy	utility to combine several Seismic Unix (.sux) SEGY files into a single SEGY file
rsy2segy	utility to combine several Passcal (.rsy) SEGY files into a single SEGY file
xyz2segy	utility to convert xyz data to SEGY format
testEndian	utility to determine if your computer's is big or little Endian
segyswap	utility to convert a SEGY file between big/little Endian

In Example Subdirectory:

README	readme file for this directory
example	edited copy of the splot template shell script
example.segy	input SEGY file of a stacked depth section
example.xyz	output ASCII file
example.1000wv.ps	output postscript file, 1000 ms depth, combined wiggle trace and variable area display
example.cdpx	cdp VS x coordinate file

INPUT DATA FORMAT

SAC reads standard 4 byte IBM or IEEE floating point SEGY files, reading the number of samples, sample rate, channel, and common depth point values directly from the SEGY file.

SAC assumes that the input SEGY file starts with a 3600-byte reel header, with bytes 3217-3218 containing the sample rate in microseconds and bytes 3221-3222 containing the number of samples.

Each trace should be preceded by a 240-byte trace ID header, with bytes 13-16 containing the channel number and for a cdp gather, bytes 21-24 containing the cdp for the trace.

SEGY files may contain more than one gather per trace, though this is non-standard. For example, a SEGY files will often contains several shot gathers. For SAC to run, each gather must have the same number of traces per gather, and each trace must have the same sample interval and number of samples.

A common problem when dealing with binary formats like SEGY is the byte order, which can be either big-Endian (on Mac's and Unix machines) or little-Endian format (on PC's). A SEGY file created on a little-Endian machine will look like garbage if read on a big-Endian machine, and visa-versa. A utility program, segyswap, is included to toggle the SEGY file between these two types. If you encounter an "unexpected end of SEGY file" error message, or the output looks like garbage, try swapping the byte order of the SEGY file.

OUTPUT DATA FORMAT

SAC converts an input SEGY file to an ASCII file in xyz format, where x is the channel number, CDP, or the x coordinate of the trace, y is the time or depth of the trace sample, and z is the amplitude of the trace sample.

A line containing only the '#' character in column 1 precedes the listing for each trace. After this line the trace samples are listed, one sample per line, with the x, y, and z values of the sample separated by spaces or tabs.

Samples are listed in order of ascending time or depth. All values are right justified within specific column ranges. The x value is in columns 1-8, the y value in columns 10-17, and the z value in columns 19-30.

Below is an example for 2 traces (channels 1 and 2), each with 2 samples per trace (sample times of 0 and 5 ms).

```
#
   1.00       0     -1.6961
   1.00       5     -0.1384
#
   2.00       0      0.1334
   2.00       5      1.9543
```

CDP CONVERSION FILE FORMAT

The CDP to x coordinate conversion file has a free-form format, where the first field is the CDP number, and the second field is the x coordinate. Header and comment lines are not necessary, but may be included if they start with a '#' character in column 1.

Below is an example for a typical CDP to x coordinate conversion file.

```
# optional comment line
1         97.2345
2         100.1256
```

GETTING STARTED

Move to the directory containing the SAC code and this README file.

Compile the C++ source code using the makefile provided by typing "make all".

Create a working subdirectory within the SAC directory, and copy the splot or splot2 template shell script to your working directory, perhaps renaming the shell script to a more descriptive name within your working directory (see the Example subdirectory).

Move to your working directory

Edit your working directory's version of the template shell script, and set the conversion and plotting parameters within that file.

note: do not edit the template splot or splot shell scripts in the main SAC directory, only edit copies of these shell scripts in your own working directory

Move the input SEGY file to your working directory. The SEGY file must have the suffix ".segy"

note: if using FTP, make sure you set the transfer type to "binary"

If needed, create a cdp to x coordinate conversion file in your working directory.

Run the shell script in your working directory by typing the name of the shell script, followed by the name of the input SEGY file. This should create an ASCII xyz file and a postscript file.

The postscript file will have the suffix "xy.ps", where "x" is the maximum time or depth, and "y" is the trace display type (c for grayscale display, w for wiggle display, v for variable area display). For example, a postscript plot of a gather down to 500 ms, with wiggle and variable area trace display, will have the suffix "500wv.ps"

If plotting multiple gathers using the splot2 shell script, then the postscript file will have the suffix "gather.xy.ps", where "gather" is the gather number + offset, with offset defined in the splot2 shell script. The gather number is the sequential gather in the SEGY file, starting at one and incrementing by one with every subsequent gather.

SHELL SCRIPT PARAMETERS

A large number of parameters must be specified within the shell script. Some parameters near the bottom of the shell script, such as “text_title offset”, may not need alteration. Other parameters, such as trace scale and text messages, may be changed quite frequently.

The shell script sets various parameters using the "set" command. Comments may be inserted in the file by starting a line with the "#" character. Parameter values may be toggled by including several lines that set the same parameter, with all but one of the lines commented out. For example:

```
#set text_1 = "agc set to 150"
#set text_1 = "agc set to 300"
set text_1 = "agc set to 500"
```

The shell script in your working directory calls a shell script in the main SAC directory. These shell scripts in the main SAC directory should not be edited.

The following is a list of parameters in the shell scripts:

```
# INPUT FILE PARAMETERS (PRESENT IN SPLOT2 SHELL SCRIPT ONLY)
set shotno_first =     1          # first sequential shot in SEGY file to plot
set shotno_last  =     2          # last  sequential shot in SEGY file to plot
                                  # by definition, the first shot in a SEGY file is 1
set traces_per_shot = 602         # traces per shot in SEGY file (must be constant)
set label_offset  =  1000         # value to add to shot number for text_1 label

# INPUT FILE PARAMETERS
set ieee_flag    =     1          # 0 if input SEGY in IBM floating point format,
                                  # 1 if input SEGY in for IEEE format

# TRACE DISPLAY PARAMETERS
set trace_type   =     3          # 0 for no trace display,
                                  # 1 for wiggle display, 2 for variable area, 3 for both
set w_scale      =   100          # scale for pswiggle, smaller number = larger traces

# COLOR DISPLAY PARAMETERS
set color_type =     0           # 0 for no color image, 1 for grayscale color image
set c_scale      =     0.1        # for color image, smaller number = more color variation
set x_grid       =     2.5        # x axis grid spacing (trace spacing in chan, cdp, or dist units)
set y_grid       =     1          # y axis grid spacing (sample spacing in time or depth units)
set x_final_grid =  2.5          # final x axis grid spacing (set = to x_grid for no smoothing)
set y_final_grid =  1.0          # final y axis grid spacing (set = to y_grid for no smoothing)

# X AXIS ATTRIBUTES AND MODE FOR SEISMIC PLOT
set x_start      =     1          # x axis start, in cdp, chan, or distance units
set x_end        =   602          # x axis end, in cdp, chan, or distance units
set x_tick       =    10          # x axis tick mark interval
set x_anot       =    50          # x axis tick mark annotation interval
```

```
set x_mode    = "chan"          # x-coordinate mode, must be "cdp", "chan", or "dist"
set dist_file = "CDPX"          # for "dist" mode use this file as conversion file
set x_title   = "CHANNEL"       # x axis label

# Y AXIS ATTRIBUTES AND DATUM FOR SEISMIC PLOT
set datum     =  0              # datum added to sample's y value
                                # datum < 0 will shift traces upward
                                # datum > 0 will shift traces downwards
                                # example:  if trace samples are from 0 to 100 ms,
                                #    then datum = -40 will shift samples so that
                                #    they are from -40 to 60 ms

set y_start   =    0            # y axis start (usually set to datum)
set y_end     = 7000            # y axis end, in time or depth units
set y_tick    =  100.0          # y axis tick mark interval
set y_anot    =  500.0          # y axis tick mark annotation interval
set y_title   = "TIME (ms)"     # y axis label

                                # note: one can limit the amount of data displayed
                                #    by setting the start/end parameters to be
                                #    less than the actual amount of data present

# TITLE AND TEXT MESSAGES
set title     = "Title"
set text_1    = "First line of text'      # overwritten in splot2 with line: "ffid x",
                                          # where x is the shot_number + label_offset
                                          # to not overwrite, comment out line 156
                                          # in splot2.body shell script
set text_2    = "Second line of text"

# OVERALL SIZE AND PLACEMENT OF SEISMIC PLOT (IN INCHES)
set plot_length  =  9.0         # length of seismic plot, in inches
set plot_height  =  4.0         # height of seismic plot, in inches
set x_offset     = 1.25         # offset from left edge of paper
set y_offset     = 1.00         # offset from bottom edge of paper
set dpi          = 1200         # dots per inch of image

# TRACE POLARITY PARAMETERS
set polarity     =  0           # 0 to fill positive excursions, 1 to fill negative
set bias         =  0.001       # bias to add to sample values
                                # (needed to ensure that zero traces are plotted)

# FONT SIZES
set title_size       = 16       # title size
set text_1_size      = 16       # text 1 size
set text_2_size      = 10       # text 2 size
set plot_label_size  = 12       # axis label size of seismic plot
set plot_anot_size   = 12       # annotation label size of seismic plot
```

```
# SEISMIC PLOT PEN WIDTH AND COLOR ATTRIBUTES
set fill_color      = 0         # 0-255 grayscale OR red/green/blue format
set plot_color      = 0         # 0-255 grayscale OR red/green/blue format
set plot_pen        = 0.1       # pen width/style for seismic traces
set ft_pen          = 5         # pen width/style for frame and tick marks
set tick_ln         = 0.12      # tick length
set anot_off        = 0.06      # offset of annotation from frame

# OFFSETS SEPARATING IMAGE OBJECTS (IN INCHES)
set title_text_off    = 0.25    # offset between title and text 1
set text1_text2_off   = 0.20    # offset between text 1 and text 2
set text_plot_off     = 0.70    # offset between text and seismic plot
```

PASSCAL SEGY DEVIATIONS FROM STANDARD SEGY:

1) missing initial 3600 byte reel header

2) number of samples are at bytes 229-232 (allows for higher values than 32767)

3) sample values are in 2 or 4 byte integer format
(bytes 205-206 = 0 then 2 byte integer format, = 1 then 4 byte integer format)

4) bytes 221-224 contain a floating point scaling factor,
where true value = sample value * scaling factor

5) bytes 225-226 have integer DAS instrument number + 10,000

SEISMIC UNIX SEGY DEVIATIONS FROM STANDARD SEGY:

1) missing initial 3600 byte reel header

2) number of samples is still at bytes 115-166 (limited to 32767 samples/trace)

3) sample values are in 4 byte IEEE floating point format

4) bytes 175-176 have integer DAS instrument number + 10,000

REFERENCES

Wessel, P., and W.H.F. Smith, 1995, The Generic Mapping Tools (GMT) version 3.0 Technical Reference & Cookbook, SOEST/NOAA.

Barry, K.M., Cavers, D.A., and Kneal, C.W., 1975, Recommended Standards for Digital Tape Format, Geophysics, vol 40, no 2, p. 344-352.