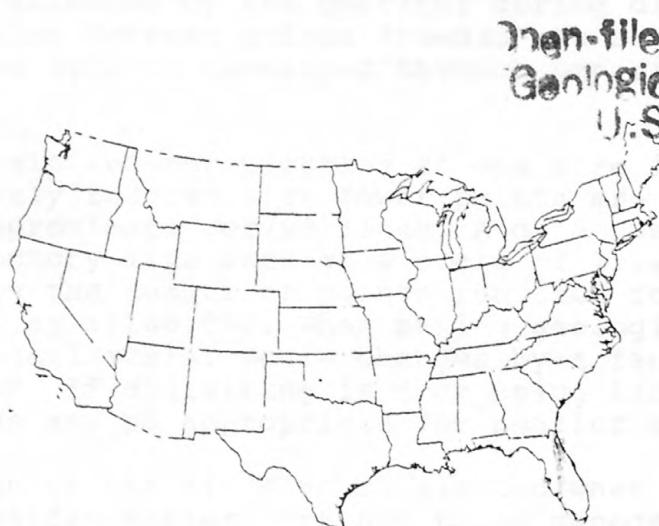UNITED STATES DEPARTMENT OF INTERIOR
GEOLOGICAL SURVEY

Filter programs GSDFIL and GSMFIL to generalize lines in GSDRAW
and GSMAP data bases, using an IBM PC (or compatible)
microcomputer.

by

Gary I. Selner and Richard B. Taylor

Open File Report
88-665A
Program disk
88-665B

Two anal

Disclaimer

This report is preliminary and has not been reviewed for conformity
with U.S. Geological Survey editorial standards. Any use of trade
names is for descriptive purposes only and does not imply
endorsement by the U.S. Geological Survey.

Although program tests have been made, no guarantee
(expressed or implied) is made by the authors regarding program
correctness, accuracy, or proper execution on all computer
systems.

Denver, Colorado

November 1988

GSDFIL and GSMFIL have been written to generalize or "filter" data points defining lines in GSDRAW and GSMAP data bases (Versions 3.0 and higher). Filtering reduces the number of points defining lines, the size of the data base and the time required for plotting. GSDFIL is used to filter lines from GSDRAW data bases; GSMFIL is used to filter lines from GSMAP data bases. Databases created using Versions 3.0 through 5.0 can be filtered by these programs.

In GSDRAW and GSMAP lines are digitized as a series of points that define straight-line segments. These points are stored as X and Y coordinates in GSDRAW and as latitude and longitude coordinates in GSMAP. The plotter draws straight lines between the digitized points. The points needed to approximate curved lines are selected by the operator during digitizing. A feel for the spacing between points "needed" to satisfactorily represent a curved line is developed through experience and practice.

When data digitized for plotting at one size is to be plotted at a greatly reduced size fewer points are "needed" to satisfactorily approximate curved lines. A data base with points producing satisfactory line work at a scale of 1:24,000 may contain many times the number of points required for a satisfactory plot at 1:100,000. When making geologic map compilations at small scale, scale changes by a factor of more than 10 are common. If digitizing is done using large scale maps, filtering of lines may be appropriate for smaller scale plots.

The selection of the lines to be plotted when changing map scales is a scientific matter, one not to be managed by a computer. The filtering of digitized points to reduce the number of points in a line, yet meet requirements for plots at smaller sizes can be done well by a computer. The original (input) data base is not modified by these programs. A new (output) data base is created.

When operating these programs the user specifies a **TOLERANCE** distance. The program selects a series of points from each original line. The first and last points of the line are always selected. Intermediate points are selected to form a generalized line in such a way that all points in the original line are closer to the generalized line than the specified <u>tolerance</u> distance. The filtering algorithm used here was developed by Douglas, D. H. and Peucker T. K. (1973). Its method of operation is described in comments in the source code quoted on page 6.

System requirements

IBM PC (or compatible) microcomputer with math coprocessor chip operating under DOS 2.0 or higher. The operation of the filter programs is computationally intensive. The filtering of large data bases may take considerable time.

General operating instructions

We suggest that you copy the executable version of the needed program GSDFIL.EXE or GSMFIL.EXE from the release disk to a working disk or subdirectory for use. Operation of GSMFIL requires an appropriate projection parameter file as used by GSMAP Version 5.0; this ASCII projection file must be created before starting GSMFIL. Projection files used by older versions of GSMAP will not work with GSMFIL.

Codes, code groups, projection parameter files and map projections are described in documentation for GSDRAW and GSMAP Version 5.0 (Selner G. I., and Taylor R. B., 1988a). Familiarity with these programs is assumed here.

GSMFIL and GSMFIL operate from prompts similar to those of GSMAP and GSDRAW. Responses to prompts are entered by typing the response, then hitting the Enter key.

A data base is specified by entering its name. The various extensions for the several files that make up the data base are not entered. Drives, or directories other than the default can be specified by following DOS conventions.

A file is specified by entering its full name with extension. Drives or directories other than the default may be specified by following DOS conventions.

Lines codes 1-99 and 400-499 can be filtered by these programs. Entries belonging to other code groups are copied to an output data base without modification. Deleted entries are not copied into the output data base.

Nodes are not created by this program. Node points at intermediate positions on lines may not survive filtering. If preservation of all node points is important, we recommend use of the fragment option of GSDUTIL or GSMUTIL before filtering the data base. Nodes at the beginning and ending points of lines in a filtered data base can be created using option 1 of GSDUTIL or GSMUTIL (Selner and Taylor 1988b.

Enter **GSDFIL** to start operation of the program. A sequence of four prompts will be displayed. Answer each in turn:

ENTER GSDRAW INPUT DATA BASE:

Enter the name of the input data base.

ENTER GSDRAW OUTPUT DATA BASE:

Enter the name desired for the output data base.

ENTER TOLERANCE:

Enter the tolerance distance in units of the data base. For most data bases the tolerance will be expressed in inches. Example, **.001**

ENTER CODE TO BE SELECTED(-1=ALL, 0=EXIT):

Enter the code required. After entry of the code the program will operate. The "ENTER CODE TO BE SELECTED" prompt will return to enable selection of an additional code until a **0** is entered to end the session. Enter **-1** to filter all lines codes 1-99 and 400-499 and also copy entries of all other codes to the output data base. After entry of **-1** the program will operate and then will return the system to the DOS prompt .

Notes:

- Selection of a tolerance distance of .001" is equivalent to plotter resolution; selection of 0.003" is about the width of a fine line.

- If GSDFIL is used to reduce the number of points in a data base but plots are to be made at the scale at which digitizing was done, we recommend use of a tolerance distance of 0.001 inches.

Enter **GSMFIL** to start operation of the program. A sequence of five prompts will be displayed. Answer each in turn:

ENTER GSMAP INPUT DATA BASE:

Enter the name of the input data base.

ENTER GSMAP OUTPUT DATA BASE:

Enter the name of the output data base.

ENTER FILENAME OF PROJECTION PARAMETERS:

Enter the name of the projection file to be used.

ENTER TOLERANCE:

Enter the tolerance distance in <u>meters</u> <u>at</u> <u>map</u> <u>scale</u>. Example, **1000**

ENTER CODE TO BE SELECTED(-1=ALL, 0=EXIT):

Enter the code required. After entry of the code the program will operate. The "ENTER CODE TO BE SELECTED" prompt will return to enable selection of an additional code until a **0** is entered to end the session. Enter **-1** to filter all lines codes 1-99 and 400-499 and also copy entries of all other codes to the output data base. After entry of **-1** the program will operate and then will return the system to the DOS prompt .

Notes:

- In GSMFIL <u>don't</u> enter a tolerance in inches as you would in GSDFIL. <u>Calculate</u> <u>the</u> <u>equivalent</u> <u>distance</u> <u>in</u> meters <u>at</u> <u>map</u> <u>scale</u>; ie. a tolerance of .003" for a 1:100,000 scale map would be: 0.003/39.37X100000=7.6 m. 7.6 would be entered at the tolerance prompt. Remember that 1 m = 39.37 inches; slide rule accuracy is appropriate.

- Selection of a tolerance distance of .001" <u>at</u> <u>map</u> <u>scale</u> is equivalent to plotter resolution; selection of 0.003" <u>at</u> <u>map</u> <u>scale</u> is about the width of a fine line.

- Use the projection parameter file corresponding to the map projection that you expect to use for final plotting. The filter algorithm works in terms of distances and directions and the tolerance is specified in meters, so use of an incorrect map projection may not produce satisfactory results.

- If GSMFIL is being used to reduce the number of points in a data base and plots are to be made at the scale at which digitizing was done, we recommend selection of a tolerance distance equivalent to 0.001 inches at map scale.

Douglas Peucker filter

The section below is quoted from the source code GSMFIL.BAS, and describes the operation of the Douglas Peucker filter.

```
' *       THIS PROCEDURE IS AN IMPLEMENTATION OF THE DOUGLAS-PEUCKER   *
' * ALGORITHM FOR LINE GENERALIZATION (SEE 'ALGORITHMS FOR THE REDUC-  *
' * TION OF THE NUMBER OF POINTS REQUIRED TO REPRESENT A DIGITIZED      *
' * LINE OR ITS CARICATURE', BY D.H. DOUGLAS AND T.K. PEUCKER,         *
' * CANADIAN CARTOGRAPHER, VOL. 10, NO. 2, DEC. 1973, PP. 112-122.     *
' *                                                                     *
' *       THIS ALGORITHM SELECTS THE POINT IN THE LINE WITH THE GREATEST*
' * PERPENDICULAR DISTANCE FROM A MATHEMATICAL LINE BETWEEN THE ANCHOR *
' * (1ST POINT) AND THE FLOATER (LAST POINT). TWO LINES ARE FORMED,    *
' * ONE CONSISTING OF THE POINTS BEFORE THE POINT WITH THE GREATEST    *
' * PERPENDICULAR DISTANCE AND ONE CONSISTING OF THE POINT AFTER. EACH *
' * OF THESE LINES IS EXACTLY AS THE ORIGINAL LINE WAS, EACH POSSIBLY  *
' * FORMING 2 MORE LINES. THE PROCEDURE CONTINUES UNTIL ALL POINTS     *
' * BETWEEN THE ANCHOR AND FLOATER ARE WITHIN A PRESET TOLERANCE FOR   *
' * EACH LINE PART. THE GENERALIZED LINE THEN CONSISTS OF ALL ANCHOR   *
' * AND FLOATER POINTS.                                                 *
' *                                                                     *
```

Source Code

The source code supplied here is written in Microsoft QuickBasic and compiled using Version 4.0 of this program.

Files on release disk

CLDFIL	BAT	Batch command file for compilation of GSDFIL
GSDFIL	BAS	Source code for GSDFIL
GSDFIL	EXE	Executable program GSDFIL
CLMFIL	BAT	Batch command file for compilation of GSMFIL
GSMFIL	BAS	Source code for GSMFIL
PROJLIB	BAS	Source code projection routines
GSMFIL	EXE	Executable program GSMFIL

Acknowledgments

The algorithm used here was published by D. H. Douglas and T. K. Peucker (1973). The version used here has been rewritten in Microsoft QuickBasic from Fortran code supplied by personnel of the National Mapping Division of the U.S. Geological Survey. We very much appreciate their assistance.

References

Douglas, David H., and Peucker, Thomas K., 1973, Algorithms for the reduction of the number of points required to represent a digitized line or its caricature: The Canadian Cartographer, v. 10, no. 2, p. 112-122.

Selner, G., I., and Taylor, R. B., 1988a, GSDRAW and GSMAP Version 5.0: prototype programs, level 5, for the IBM PC and compatible microcomputers, to assist compilation and publication of geologic maps and illustrations: U.S. Geological Survey Open File Report 88-295A, documentation and tutorial, 130 p., and 88-295B, two executable program disks.

Selner, G., I., and Taylor, R. B., 1988b, Utility programs for GSDRAW and GSMAP Version 5.0 using an IBM PC (or compatible) microcomputer, digitizer, and plotter: U.S. Geological Survey Open File Report 88-537A, documentation 130 p., and 88-537B, two program disks.