



## Short note

# A Visual Basic Program to Generate Sediment Grain-Size Statistics and to Extrapolate Particle Distributions<sup>☆</sup>

L.J. Poppe<sup>a,\*</sup>, A.H. Eliason<sup>b</sup>, M.E. Hastings<sup>a</sup><sup>a</sup> *US Geological Survey, 384 Woods Hole Road, Woods Hole, MA 02543, USA*<sup>b</sup> *Eliason Data Services, 230 Meetinghouse Road, Mashpee, MA 02649, USA*

Received 10 September 2003; received in revised form 20 April 2004; accepted 1 May 2004

## 1. Introduction

Measures that describe and summarize sediment grain-size distributions are important to geologists because of the large amount of information contained in textural data sets. Statistical methods are usually employed to simplify the necessary comparisons among samples and quantify the observed differences. The two statistical methods most commonly used by sedimentologists to describe particle distributions are mathematical moments (Krumbein and Pettijohn, 1938) and inclusive graphics (Folk, 1974). The choice of which of these statistical measures to use is typically governed by the amount of data available (Royse, 1970). If the entire distribution is known, the method of moments may be used; if the next to last accumulated percent is greater than 95, inclusive graphics statistics can be generated. Unfortunately, earlier programs designed to describe sediment grain-size distributions statistically do not run in a Windows environment, do not allow extrapolation of the distribution's tails, or do not generate both moment and graphic statistics (Kane and Hubert, 1963; Collias et al., 1963; Schlee and Webster, 1967; Poppe et al., 2000)<sup>1</sup>.

Owing to analytical limitations, electro-resistance multichannel particle-size analyzers, such as Coulter Counters, commonly truncate the tails of the fine-fraction part of grain-size distributions. These devices do

not detect fine clay in the 0.6–0.1  $\mu\text{m}$  range (part of the 11-phi and all of the 12-phi and 13-phi fractions). Although size analyses performed down to 0.6  $\mu\text{m}$  microns are adequate for most freshwater and near shore marine sediments, samples from many deeper water marine environments (e.g. rise and abyssal plain) may contain significant material in the fine clay fraction, and these analyses benefit from extrapolation.

The program (GSSTAT) described herein generates statistics to characterize sediment grain-size distributions and can extrapolate the fine-grained end of the particle distribution. It is written in Microsoft Visual Basic 6.0 and provides a window to facilitate program execution. The input for the sediment fractions is weight percentages in whole-phi notation (Krumbein, 1934; Inman, 1952), and the program permits the user to select output in either method of moments or inclusive graphics statistics (Fig. 1). Users select options primarily with mouse-click events, or through interactive dialogue boxes.

## 2. System requirements

Windows 95/98/ME/NT/2000/XP computers with minimal specifications and the necessary support and associated files will be able to run the compiled Visual Basic version of the program without difficulty. If recompiling the program from its raw code, users with Windows 95 or early versions of Windows 98 may need to install updated versions of three support files into their C:\Windows\System directory. These files include the Common Dialog control file COMDLG32.OCX, the Visual Basic “Run-Time” driver MSVBVM60.DLL,

<sup>☆</sup> Code available from server at <http://www.iamg.org/CGEditor/index.htm>.

\*Corresponding author. Tel.: +1-508-457-2314.

E-mail address: [lpoppe@usgs.gov](mailto:lpoppe@usgs.gov) (L.J. Poppe).

<sup>1</sup> Rockware Inc, 2003. <http://www.rockware.com/>.

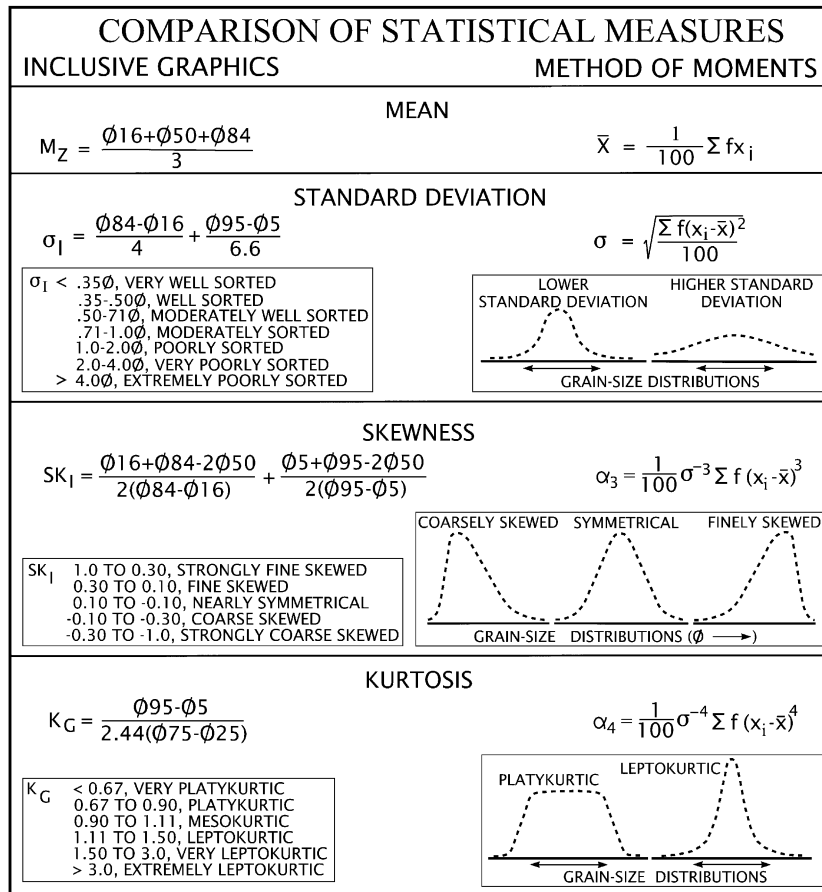


Fig. 1. Equations used to calculate method of moments (Collias et al., 1963) and inclusive graphics with verbal descriptions (Folk, 1974) for grain-size distributions. The figure also shows graphical representations of statistics.

and the standard data formatting object library driver MSSTDFMT.DLL. These support files are included in the compiled software setup package available from the corresponding author or, alternatively, may be downloaded freely from several sites on the internet<sup>2,3</sup>.

### 3. Program details

The GSSTAT code, setup procedures, support files, and documentation are available at <http://www.iamg.org/CGEditor/index.htm> by anonymous ftp at the International Association for Mathematical Geology website and on a CD-ROM from the corresponding author. The program presents the user with several options (Fig. 2). The user must: (1) select either method

of moments or inclusive graphics statistics; (2) select the method of extrapolation from a pull-down menu (either none, linear, exponential, or the average of these two); (3) if extrapolating, enter the smallest particle size actually measured (in microns); (4) select the format of the phi-class percentages in the input file (i.e. frequency or cumulative frequency percent); and (5) enter the format of the phi-class percentages in the output file.

If none is selected as the method of extrapolation, then the program can simply be used to generate statistics for the known particle-size distribution, and entering the smallest particle size measured is not necessary. Fig. 3 presents a stylized plot showing examples of linear, exponential, and averaged extrapolations of the clay fraction to the clay – colloidal clay boundary. The linear extrapolation tends to overestimate the amount of clay present in a typical distribution, and may be used by operators to account for material within the colloidal fraction. The exponential extrapolation tends to underestimate the amount of clay present in a typical distribution, and may be used by operators

<sup>2</sup>Microsoft Product Support Services, 2003. <http://support.microsoft.com/default.aspx?>

<sup>3</sup>WebAttack.com, 2003. <http://www.webattack.com/help/missingfiles.html>.

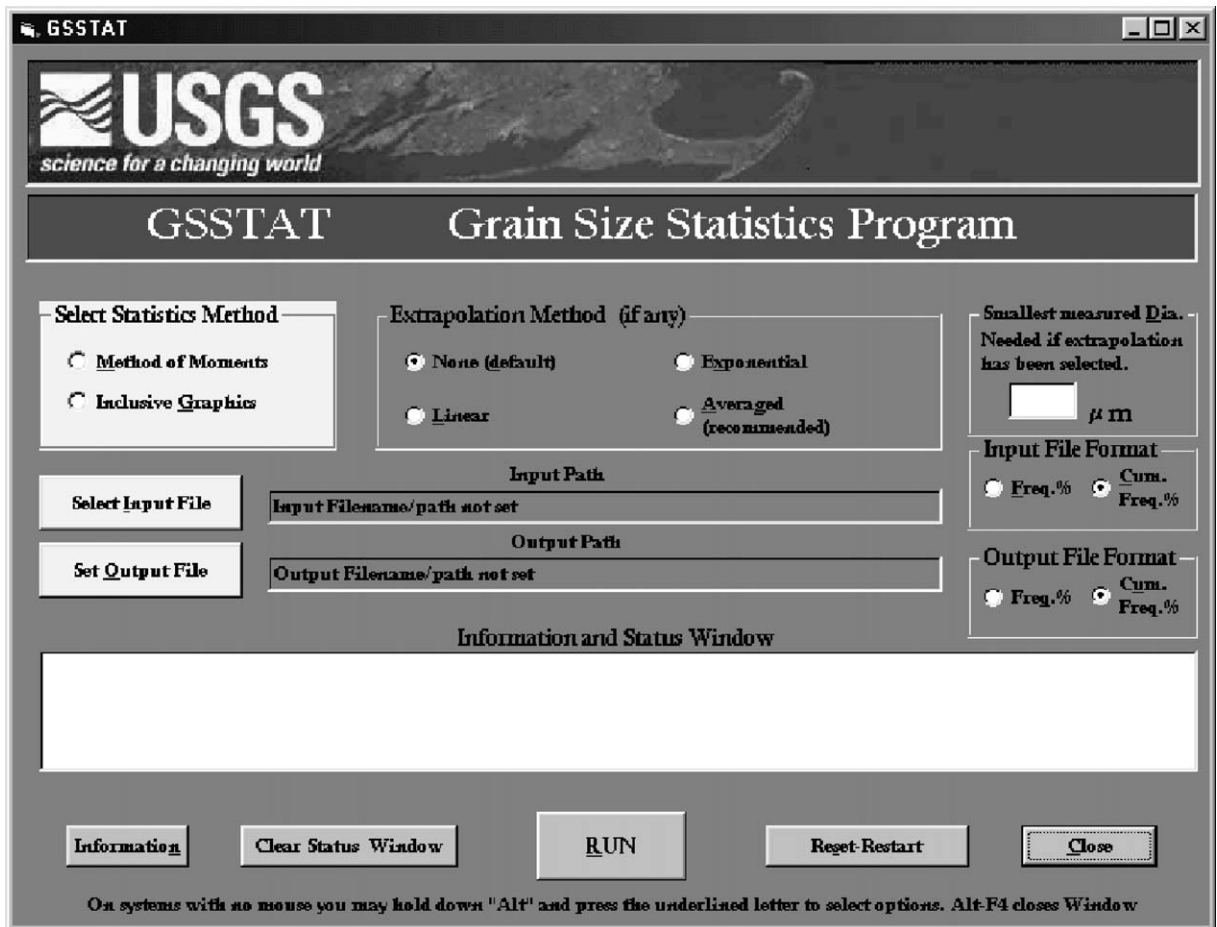


Fig. 2. Window for GSSTAT program. User can select statistical and extrapolation methods, designate input and output files, and start or cancel application. Clicking input and output file options opens text boxes that allow user to enter file names and browse for destination directories.

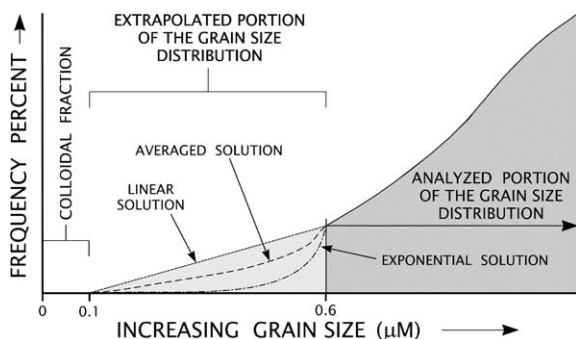


Fig. 3. Stylized plot showing that portion of clay fraction typically truncated by electro-resistance particle size analyzers. Also shown are examples of linear, exponential, and average of linear and exponential extrapolations to 0.1  $\mu\text{m}$ , the clay – colloidal clay boundary by program GSSTAT. Errors inherent in linear and exponential solutions have been exaggerated to show their effect.

who want data that represent the minimal amount of clay present. Although an operator may select linear extrapolation, exponential extrapolation, or the mean of both extrapolations, the mean usually provides the most accurate estimate and is, therefore, the recommended solution for most applications. To prevent “overshoots” found in classic spline methods, curves for the inclusive graphics statistics were generated by constrained cubic spline interpolation (Kruger, 2003).

Selecting the “input file” button opens a text window that allows the user to identify the file to be processed and to specify its location. The input file must be in comma delimited ASCII text (.txt or .csv) and have 20 fields that include: Sample identifier, latitude, longitude, and the percentages (cumulative or frequency) of the 11-phi to  $-5$ -phi fractions (Table 1A). The latitude and longitude fields may be left blank (i.e. unpopulated, but not omitted), but the sum of the phi fraction percentages must total 100% ( $\pm 0.2\%$ ). The program expects the

Table 1

Examples of input and output GSTAT files. (A) Typical input file with twenty fields in comma-delimited ASCII

---

A

SAMPLEID,LATITUDE,LONGITUDE,11\_PHI,10\_PHI,9\_PHI,8\_PHI,7\_PHI,6\_PHI,5\_PHI,4\_PHI,3\_PHI,2\_PHI,1\_PHI,0\_PHI,  
M1PHI,M2PHI,M3PHI,M4PHI,M5PHI  
AQ364,40.83333,-70.38333,2.76,4.67,7.04,10.8,9.12,2.66,0.7,18.63,41.72,0.51,0,1.38,0,0,0,0  
AM306,41.16755,-72.2307,0.16,0.38,0.43,0.46,0.54,0.5,0.38,4.08,28.2,22.37,12.06,13.19,9.96,7.29,,  
AN366,,4.28,8.53,11.96,16.96,14.21,6.67,3.89,32.14,0.53,0.45,0.27,0.11,0,0,0,0

B

SAMPLEID,LATITUDE,LONGITUDE,GRAVEL\_PCT,SAND\_PCT,SILT\_PCT,CLAY\_PCT,SEDCLASS,MEDIAN,MEAN,  
STDDEV,SKEWNESS,KURTOSIS,VERBAL\_DES,PHI\_13,PHI\_12,PHI\_11,PHI\_10,PHI\_9,PHI\_8,PHI\_7,PHI\_6,PHI\_5,  
PHI\_4,PHI\_3,PHI\_2,PHI\_1,PHI\_0,PHIM1,PHIM2,PHIM3,PHIM4,PHIM5,ERRORS  
AQ364,40.83333,-70.38333,0.00,60.22,22.53,17.25,SILTY SAND,3.43,4.83,2.83,0.37,-0.67,,100.00,99.34,98.02,94.08,89.56,82.75,  
72.30,63.47,60.90,60.22,42.20,1.83,1.34,1.34,0.00,0.00,0.00,0.00,0.00,  
AM306,41.16755,-72.2307,17.22,79.74,1.88,1.16,GRAVELLY  
SEDIMENT,1.34,1.07,2.02,0.37,3.02,,100.00,99.96,99.88,99.65,99.27,98.84,98.38,97.84,97.34,96.96,92.89,64.74,42.42,30.38,17.22,7.28,  
0.00,0.00,0.00,  
AN366,,0.00,31.84,39.67,28.49,SAND SILT CLAY,6.60,6.41,2.52,0.10,-0.97,,100.00,99.00,96.99,90.98,82.88,71.51,55.39,41.88,35.54,  
31.84,1.29,0.79,0.36,0.10,0.00,0.00,0.00,0.00,0.00,

C

SAMPLEID,LATITUDE,LONGITUDE,GRAVEL\_PCT,SAND\_PCT,SILT\_PCT,CLAY\_PCT,SEDCLASS,  
MEDIAN,MEAN,STDDEV,SKEWNESS,KURTOSIS,VERBAL\_DES,PHI\_13,PHI\_12,PHI\_11,PHI\_10,PHI\_9,PHI\_8,PHI\_7,PHI\_6,  
PHI\_5,PHI\_4,PHI\_3,PHI\_2,PHI\_1,PHI\_0,PHIM1,PHIM2,PHIM3,PHIM4,PHIM5,ERRORS  
AQ364,40.83333,-70.38333,0.00,62.24,23.28,14.47,SILTY SAND,3.21,4.50,2.46,0.71,0.68,VERY POORLY SORTED; STRONGLY  
FINE SKEWED; PLATYKURTIC,,2.76,4.67,7.04,10.8,9.12,2.66,0.70,18.63,41.72,0.51,0.00,1.38,0.00,0.00,0.00,0.00,  
AM306,41.16755,-72.2307,17.25,79.90,1.88,0.97,GRAVELLY SEDIMENT,1.39,0.95,1.76,-0.34,0.84,POORLY SORTED;  
STRONGLY COARSE SKEWED;  
PLATYKURTIC,,0.16,0.38,0.43,0.46,0.54,0.50,0.38,4.08,28.20,22.37,12.06,13.19,9.96,7.29,0.00,0.00,0.00,  
AN366,,0.00,33.50,41.73,24.77,SAND SILT CLAY,6.48,6.23,2.32,-0.06,0.64,VERY POORLY SORTED; NEAR  
SYMMETRICAL; VERY PLATYKURTIC,,4.28,8.53,11.96,16.96,14.21,6.67,3.89,32.14,0.53,0.45,0.27,0.11,0.00,0.00,0.00,0.00,0.00,

---

First record is a header; next three records contain size distributions in frequency percent. Fields must be in this order, but population of fields is optional. Note last record has no navigation. (B) Output file generated from input file above with method of moments (Collias et al., 1963; Schlee and Webster, 1967), cumulative frequency percent, and mean of linear and exponential extrapolation options selected. Note 13-and 12-phi fields are populated. (C) Output file with 33 fields generated from input file above with inclusive graphics (Folk, 1974), frequency percent, and no extrapolation options selected. Note that verbal description field is populated.

first line of the input file to be a header showing the attribute names; no embedded commas are allowed in any of the data fields. Selecting the “output file” button opens a text window that allows the user to name the file to be generated and to specify the destination directory.

The program generates an output file in the requested destination directory and allows the user to view results in a display window for verification purposes. The output file has a header for its first line, and contains 33 fields including: sample identifier, latitude, longitude, percentages of gravel, sand, silt, and clay, sediment classification, median, mean, standard deviation, skewness, kurtosis, verbal descriptions, and the frequency or cumulative frequency percentages of the whole-phi fractions from 13 phi through  $-5$  phi. If the user has selected extrapolation, the two additional phi-fraction fields will be populated (Table 1B). If the user has selected inclusive graphics statistics, the verbal description field will be populated (Fig. 1, Table 1C).

The program will generate error messages if problems associated with file format, required parameters, or data content are encountered. Most of the error messages appear in a new drop-down window, but some warnings appear in the list view screen of the main program display window and(or) in the 34th field of the output file. A complete listing of error messages with common causes and suggested solutions is provided in the documentation. Blank records abort the program.

Selecting the “Information” button allows the user to view the documentation, output files, and five schematics including: a correlation chart of the phi classes with the Wentworth (1922) grade scale, equations for methods of moments statistics with graphical representations, the equations for inclusive graphics statistics with verbal descriptions, extrapolation options (none, linear, exponential, and averaged), and the Shepard (1954) classification system. On systems with no mouse, the user can hold down the “Alt” button key and press

the underlined character of the desired function. Successful completion of a required option changes the color of the button. The Add/Remove Programs option in the Windows control panel folder can be readily used to uninstall the program.

Pre-existent statistical (Schlee and Webster, 1967; Poppe et al., 2000) and extrapolation (Poppe and Eliason, 1999) software were used to check the accuracy of these functions in the GSTAT Program. There should be no upper limit to the number of records allowed in an input file; we have tested the program to over 400 records.

### Acknowledgements

We thank Dave Twichell, Ellen Mecray, and Jim Robb (all USGS) for their thoughtful reviews, and the Coastal and Marine Geology Program of the US Geological Survey for its financial support. Although this program has been used by the USGS and Eliason Data Services no warranty is expressed or implied. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

### References

- Collias, E.E., Rona, M.R., McManus, D.A., Creager, J.S., 1963. Machine processing of geological data. University of Washington Technical Report Number 87, 119pp.
- Folk, R.L., 1974. Petrology of Sedimentary Rocks Hemphill Publishing Co., Austin, TX, 182pp.
- Inman, D.L., 1952. Measures for describing size of sediments. *Journal of Sedimentary Petrology* 19 (2), 125–145.
- Kane, W.T., Hubert, J.F., 1963. Fortran program for the calculation of grain-size textural parameters on the IBM 1620 computer. *Sedimentology* 2 (1), 87–90.
- Kruger, C.J.C., 2003. Constrained Cubic Spline Interpolation. <http://www.korf.co.uk/spline.pdf>.
- Krumbein, W.C., 1934. Size frequency distribution of sediments. *Journal of Sedimentary Petrology* 4, 65–77.
- Krumbein, W.C., Pettijohn, F.J., 1938. Manual of Sedimentary Petrography Appleton-Century-Crofts, New York, 549pp.
- Poppe, L.J., Eliason, A.H., 1999. An interactive computer program to extrapolate the clay fraction distributions of truncated grain-size data. US Geological Survey Open-File Report 99–27, 44pp.
- Poppe, L.J., Eliason, A.H., Fredericks, J.J., Rendigs, R.R., Blackwood, D., Polloni, C.F., 2000. Grain size analysis of marine sediments: methodology and data processing. In: Poppe, L.J., Polloni, C.F. (Eds.), USGS East-coast sediment analysis: procedures, database, and georeferenced displays, US Geological Survey Open-File Report 00-358, CD-ROM.
- Royse, C.F., 1970. An Introduction to Sediment Analysis. Arizona State University, Tempe, 180pp.
- Schlee, J., Webster, J., 1967. A computer program for grain-size data. *Sedimentology* 8 (1), 45–54.
- Shepard, F.P., 1954. Nomenclature based on sand–silt–clay ratios. *Journal of Sedimentary Petrology* 24 (3), 151–158.
- Wentworth, C.K., 1922. A scale of grade and class terms for clastic sediments. *Journal of Geology* 30, 377–392.