CHAPTER 6

UNCERT: GEOSTATISTICAL, GROUND WATER MODELING, AND VISUALIZATION SOFTWARE

UNCERT is an uncertainty analysis, geostatistical, ground water modeling, and visualization software package. It was developed for evaluating the uncertainty associated with the characterization and prediction of subsurface geology, hydraulic properties, and the migration of hazardous contaminates in groundwater flow systems. The package is well suited for evaluating hazardous waste sites and evaluating remediation methods, but it also includes general modules which are usable by researchers from a wide range of disciplines.

6.1: Introduction

UNCERT is a collection of software program modules designed to work together to aid ground water modelers through the data analysis, site modeling, and site evaluation processes. A flow chart of the basic UNCERT processes is shown in Figure 6.1. Many of the tools are also applicable to scientists and engineers from many other fields such as mining, oil exploration, meteorology, and criminology.

6.2: Previous Work

Much of the programming in UNCERT is original work, some of which has been described in previous chapters, but some of the codes are taken from or based on previous work. Software code from other sources, was either 1) public domain, 2) offered with unrestricted use such that copyright notice remained intact and was referenced, 3) transferred to me with permission from the author or the authors agent, or 4) a previous code/algorithm was used as a reference, and original code was written. These resources are described in Appendix G of the UNCERT User's Manual (Appendix A, CD-ROM) and are summarized below:

Project Design: Indicator Kriging and Stochastic Simulation Analysis Flowchart

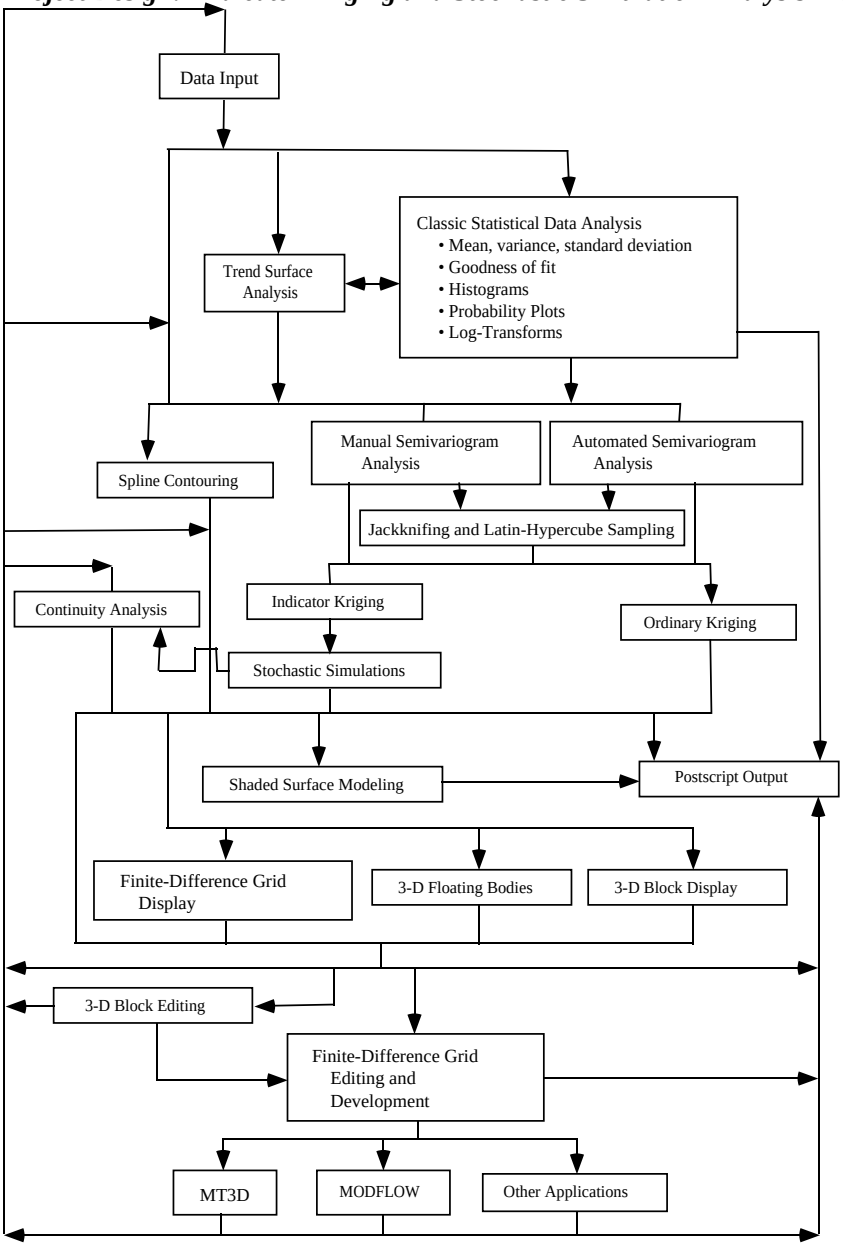

FIGURE 6-1. Detailed flow chart of uncertainty analysis software package.

- Three-dimensional rotations and transformations for 2D and 3D visualization. The algorithms are based on work by Foley, et. al. (1984).
- MODFLOW (McDonald and Harbaugh, 1988).
- MT3D(Zheng, 1990).
- Simple and ordinary kriging algorithm ktb3d (Deutsch and Journel, 1992).
- Irregularly spaced data semivariogram algorithm gamv3 (Deutsch and Journel, 1992).
- Regularly gridded data semivariogram algorithm gam3 (Deutsch and Journel, 1992).
- Indicator conditional simulation program sisim3d (Gómez-Hernández and Srivastava, 1990; McKenna, 1994).
- Contouring and spline algorithm (Wessel and Smith, 1991).
- Rotated text (Richardson, 1993).
- Text editor/viewer program editor (Heller, 1991).
- HTML help browser (NCSA, 1993; Punin, 1994).
- Linear algebra routines from LINPACK (Dongarra, Bunch, et al., 1984).

6.3: Platform Support

In selecting a development platform for UNCERT, two main issues were of concern; portability and processing power. To balance these two issues, UNIX computers, ANSI-C compilers, the Postscript printer language, and the X-windows/Motif graphical user interface was selected. UNIX computers were selected because they 1) had the processing power required by many of the tasks in UNCERT, 2) supported true multi-tasking, and 3) at the beginning of the project, offered one of the best, high resolution graphical work environments. ANSI-C was selected for its 1) computational efficiency, 2) structured programming capability, and 3) portability between platforms. Some programs and code segments were left in FORTRAN, because translating these sections would serve little purpose, cost time, and potentially introduce software errors. Postscript was selected for printer output, because it is 1) a non-hardware dependent printer language and 2) a standard in the UNIX environment. X-windows (developed at MIT) and Motif (developed by the OSF (Open Software Foundation)) were selected as the window manager interface because they have become an industry standard on UNIX computer systems. As a standard, software developed on one system, is easily ported between platforms.

UNCERT has currently been tested on, and is running on eight different 32-bit UNIX platforms (Data General, Dec, IBM RISC-6000, HP, Linux, SGI, Solaris, and SUN) using native ANSI-C and FORTRAN compilers, and gcc (a public domain ANSI-C compiler, (GNU, 1995)) and f2c (a public domain FORTRAN77 to ANSI-C converter (Feldman, Gay, et al., 1990)).

6.4: UNCERT Modules

Below are descriptions of each of the modules in UNCERT and a brief discussion of the mathematical techniques used. For a complete description of each module, refer to the UNCERT User's Manual (Appendix A: Attached Tape). This is an HTML (Hyper-Text Markup Language, NCSA (1993)) document, viewable on a wide range on freeware and commercial World Wide Web (WWW) browsers available on Microsoft-Windows, UNIX, and Macintosh platforms. Current versions of UNCERT and the User's Manual may also be viewed or downloaded from the WWW, from:

http://uncert.mines.edu/

or by using anonymous ftp from:

ftp://uncert.mines.edu/pub/uncert/manual/

Listed below is a brief description and summary of the features of each UNCERT module.

6.4.1: Mainmenu

The mainmenu module is a simple user interface to execute the different modules in the UNCERT software package. It is designed to be a user friendly interface, so that user's can progress through the software to evaluate their field data, and to model the site of concern.

Currently mainmenu is a very simple interface which allows the user only to execute the different software modules within UNCERT. As it stands now, mainmenu is used mainly as a convenience in executing software which the user may not be familiar with, and allows the user to minimize working in the UNIX command line environment. It is a simple attempt to bring the entire UNCERT package together into a unified, windows based environment. It is not recommended that the user try to use this interface exclusively. A great deal of functionality in the software would be lost, trying to do so.

6.4.2: Plotgraph

The plotgraph application is used for plotting two-dimensional X-Y graphs. The application allows the user to plot lines, points with various symbols, and calculate regression lines (up to a tenth order polynomial). The data can be plotted using normal, semi-log, and log-log axes.

6.4.3: Histo

The histo application is used to calculate and display univariant statistics for different data sets. For a sample population, histo can be used to calculate basic statistics such as the mean, standard deviation, and variance. It may also be used to display the behavior of several different populations at once using stacked histograms, cumulative distribution plots, probability plots, and box and whisker plots.

6.4.4: Distcomp

The distcomp application calculates all of the statistical information calculated by histo, but focuses on how sample populations vary between different data sets. For a multiple sample population distcomp can be used to calculate basic statistics such as the mean, and variance. It may also be used to display the behavior of several different populations at once using stacked histograms, cumulative distribution plots, probability plots, P-P (probability vs. probability) plots, and Q-Q (quantile vs. quantile) plots.

6.4.5: Vario

The vario application is used for calculating one- and two-dimensional experimental semivariograms for scattered and regularly gridded data. The package is not limited to the classic semivariogram but will also calculate covariances, madograms, rodograms, cross-semivariograms, etc. Jackknifing the sample data set is also an option. Three types of soft indicator data can be used with hard data to calculate spatial continuity. The application displays the measure of covariance ($\gamma(h)$) versus lag.

6.4.6: Variofit

The variofit application is used to fit model semivariograms to experimental and jackknifed experimental semivariograms (generated by vario). This can be done manually or automatically using least-squares regression or latin-hypercube sampling techniques. Ergodic variations of the model semivariogram from simulation series may also be evaluated.

6.4.7: Grid

The Grid module interpolates parameter values at locations were there are no physical data. This is done using various interpolation algorithms (inverse-distance, kriging, trend-surface analysis) based on irregularly spaced data. Sometimes it is of interest to estimate what is occurring between data locations. For other applications, for convenience, or for clarity, irregularly spaced data must be interpolated onto a regular grid. For example contour, surface, and block require that the data being viewed be gridded with a rectangular pattern. These programs then allow the user to visually view the interpolated estimate of the field data. Grid is used to interpolate values at locations of convenience based on field data.

Within grid there are several gridding algorithms; inverse-distance, simple and ordinary kriging, and trend-surface analysis. Inverse-distance is a relatively simple method which estimates the value of a location based on the distance and value of surrounding sample data points. Kriging does much the same thing as inverse-distance, except kriging also considers spatial statistics describing how the field data vary spatially. Kriging is often referred to a the best unbiased estimator for evaluating a value at a given location. Trend-surface analysis is a least-squares regression technique which assumes the data values are a function of a “regional” trend with minor “local” variations. The calculated trend-surface attempts to describe the “regional” component.

6.4.8: Contour

The contour application contours and performs gradient analysis for two- and three-dimensional, regularly gridded data. Only two-dimensional views are possible however. Three-dimensional data sets can be viewed along X-Y, X-Z, and Y-Z planes. Profile lines along the contoured surface can also be plotted.

6.4.9: Surface

Surface is a 2-1/2 dimensional visualization program for viewing regularly gridded data as a color contoured , gradient, or shaded relief surface. It is included in the UNCERT software as a tool to view two-dimensional grids as three dimensional surfaces. This is referred to as a 2-1/2 dimensional surface because for each X-Y grid location, there is only one Z value. In a true 3D model (see block) each X-Y location may have multiple Z values. This package is used to view gridded surface data generated from grid, or for examining layers or cross-sections from sisim, MODFLOW (McDonald and Harbaugh, 1984), and MT3D output files. Surface may also be used to display any regularly gridded data from other sources (some data file format manipulation maybe required); DEM's (Digital Elevation Model's) are an example.

6.4.10: Block

Block is a 3-dimensional visualization program for viewing regularly gridded data or scattered data points and lines (both cannot be viewed at the same time). It is included in the UNCERT software as a tool to view the values in three-dimensional grids as three dimensional blocks. This package is used to view gridded block data generated from grid or for examining output from sisim, modmain, and mt3dmain.

6.4.11: Sisim & Sisim3d

Sisim is a graphical user interface (GUI) for sisim3d, an indicator kriging and conditional stochastic simulation program for discrete data (non-continuous data: e.g. clay, sand, gravel) developed at Stanford University by Gómez-Hernández and Srivastava (ISIM3D, 1990) and modified at the Colorado School of Mines by McKenna (1994) to utilize soft data. Up to eight indicators can be modeled in a single simulation. In its basic form sisim3d can be awkward to use, particularly when many simulations are required based on varying semivariogram models. This interface assists the user in handling data files, input parameters, coordinating multiple simulations, tasking jobs to other computers, calculating simulation statistics, and visualizing results.

6.4.12: Modmain

Modmain is a graphical user interface for MODFLOW, the MODular three-dimensional, finite difference FLOW model developed by the United States Geological Survey (McDonald and Harbaugh, 1988). MODFLOW is a program designed to model ground water flow and heads (pressure and elevation) in confined and unconfined aquifer systems. In its basic form,

MODFLOW can be difficult, or awkward to use. The modmain program module is designed to simplify data entry, model editing, and analysis of results.

6.4.13: Mt3dmain

Mt3dmain is a graphical user interface for MT3D, a modular three-dimensional transport program (Zheng, 1990). MT3D is a program designed to model contaminant transport based on a pre-solved ground-water flow model (MODFLOW is often used to solve the ground water flow equations. MT3D uses the solution aquifer heads to base the transport results). In its basic form, MT3D can be difficult, or awkward to use. The mt3dmain program module is designed to simplify data entry, model editing, and analysis of results.

6.4.14: Array

The Array module is used to manipulate mathematically one, two, or a series of block, sisim, contour, or surface 2D or 3D grid files. Depending on the options selected, operations include addition, subtraction, multiplication, division, averaging, minimum, maximum, probability value within a range, reclassification, and basic statistics. These are basic grid tools similar to those used in Geographical Information System (GIS) software. This tool can be useful for data preparation, or for data and result analysis. For example, by reclassifying a contaminant plume map to a cost of remediation map, estimates can be made about site clean up costs.

6.4.15: Utilities

In addition to the main modules, there are also several utility modules: calc, lpr_ps, ps_merge, editor, and xhelp. These are very simple user-aid utilities. Calc is a simple RPN scientific calculator. Lpr_ps is used to print ASCII text files with variable margins, line numbers, and variable font sizes. Ps_merge is used to combine, translate, and scale two UNCERT Postscript files into a single Postscript file. Editor is a simple text editor. It is convenient for viewing wide (132 column) and extremely large files. Xhelp is a simple HTML viewer, though it does not display any graphics figures.