Evaluation of Uncertainty Associated with Contaminant Migration in Groundwater - A Technically Feasible Approach

William L. Wingle and Eileen Poeter

Department of Geology and Geological Engineering
Colorado School of Mines, Golden, Colorado 80401

Abstract

Currently, a primary problem in identifying contaminant plumes in the subsurface and predicting future pathways of plumes, is the uncertainty in defining the heterogeneities of the subsurface. Identification and characterization of heterogeneity is critical to predicting migration of contaminants and designing remediation facilities. If problem areas can be accurately identified, remediation cost will be reduced because remedial action will be focused on the problem areas and the remedial methods can be specifically designed for site conditions. It is economically infeasible to characterize heterogeneity in sufficient detail through direct measurements in boreholes, so alternative approaches must be developed. A promising approach is to conduct conjunctive interpretation of all available data to quantify the uncertainty in the interpretation of subsurface heterogeneity. Results of such an analysis are probabilistic inputs for risk assessments which aid in making decisions regarding prioritization of sites for characterization and remediation, conduct of further characterization, and selection of remedial methods. If it is decided further characterization is necessary, then the analysis results can be used to identify the most cost effective plan for further data collection. Once it is decided to proceed with remediation, the techniques developed in the project can be used to select the most efficient and economical remedial techniques given the uncertainty associated with knowledge of the subsurface at the site.

Public domain, user friendly software is being developed which uses all available site data (soft data and hard data) to make a conjunctive interpretation of data describing the subsurface. Indicator and Bayesian theory are utilized to characterize the data uncertainty, and conditional stochastic indicator simulations are conducted to represent the range of subsurface configurations based on the available data. A mesh or grid designed to represent these subsurface configurations is automatically developed. Groundwater flow and contaminant transport are simulated in the various configurations and compiled to produce a frequency distribution of head, velocity, or concentration at selected locations and times. Ultimately, inverse techniques will be employed to thoroughly integrate known field conditions in estimates of flow and transport parameters within each stochastic realization, thus, further constraining interpretation of the system. Graphical interfaces are essential to proper analysis, consequently they are an integral part of the software package.

Introduction

In evaluating contaminant migration, hydrogeologists need a detailed description of subsurface materials so the flow system (saturated or unsaturated) may be modeled as accurately as possible. However, the amount of drilling that can be undertaken to characterize the site is always less than desired, whether because of the expense of drilling, or the potential for creating new pathways for contaminant migration. With limited drill hole data, perhaps geophysical data and outcrop measurements, groundwater modelers attempt to describe the subsurface. Generally, attempts yield reasonable but simplified representations of the actual geology. Simplified representations work well for water supply and drainage modeling but are not sufficient for modeling contaminant migration. Also, once the model is calibrated (model heads and concentrations approximately match field measurements), that description of the geology and material properties (hydraulic conductivity, porosity, etc.) often is the only description used for further modeling. This

approach fails to recognize that many alternative descriptions of the subsurface, which also honor geologic/geophysical data can be calibrated to match hydraulic/contaminant conditions observed in the field equally well, and depending on which description is used, the nature and extent of contaminant migration will vary. This is not to say that the hydrogeologist's model is incorrect, but that it is only one of many possible scenarios for the site. To suggest that a single model or a number of similar related models (such as the small model variations that are often made to determine model sensitivity) can define the nature and extent of contamination, is an inappropriate simplification of the problem.

The effort required to consider the many possible subsurface configurations is not trivial. Although a modeler can define many realizations of the subsurface, doing so is a time consuming and tedious task, and preconceptions by the modeler will likely bias the resulting models to a preferred pattern. In this paper, we present a procedure for making an objective, streamlined evaluation of uncertainty in the subsurface. A computer data analysis, visualization, and modeling package which is currently under development is described which will guide and aid modelers through the process.

Simulations of the geohydrologic units are accomplished either discretely using stochastic, conditional, multiple indicator simulations or continuously using more conventional kriging techniques. The discrete approach is preferred for evaluating contaminant migration because it allows the modeler to determine the probability of occurrence of a parameter value at a particular location within a model based on hard data and soft data. Each time this process is repeated one "realization" (i.e. interpretation) of the subsurface is created. By repeating this process using conditional, stochastic simulation techniques, multiple alternative realizations of the subsurface are automatically generated, with each realization honoring the known data and their statistics equally well.

Once these realizations of the subsurface have been created, they can each be input into groundwater flow and/or contaminant transport models to determine which of the realizations produce hydraulic or contaminant concentration conditions similar to those found in the field. Many will poorly represent the site and can be discarded, others will be reasonable and will show the likely nature and extent of contaminant transport. These later models can be used to determine a frequency distribution of site conditions. From this, extreme and most likely case scenarios can be determined. The probability of these scenarios occurring can be quantified (e.g. there is a 10% probability that the contaminant levels will be more than four times the federal drinking water standard in 10 years at point B, a 5% probability that levels will be below the standard, and an 85% probability (associated with a distribution that may be skewed) of being between the two levels). The operation of potential remediation facilities can be simulated within these realistic site representations to select and optimize design of such facilities based on quantitative assessments of the probability of satisfactory performance under the range of possible subsurface configurations.

The discrete representations are based on multiple indicator, conditional, stochastic simulations. Because this procedure is so valuable to the analysis and is new to many hydrogeologists, we discuss the method in relatively detailed but simple terms herein.

Data Types - Hard and Soft Data

In evaluating a site many types of data are available to the hydrogeologist, but integrating the data in a quantitative manner can be difficult. Data may be "hard" (i.e. data that can be physically measured or observed for the value of a particular parameter). Hard data might come from bore hole logs and outcrops where the material type is identified and physical parameters may have been measured. "Soft" data can take several forms; data can be imprecise because the measurements are of properties from which desired parameter values are inferred (Alabert, 1987). For example, seismic profiles measure the velocity of wave propagation. These velocities can be correlated to material types, but there are problems; the same material will have different velocities if dry or saturated, and different materials may have the same or similar velocity. Seismic studies are therefore considered soft data because they measure a property from which the desired property must be inferred. Soft data may define a possible range of values (e.g. the hydraulic conductivity is in the "sand" range ~ 10^{-4} to 10^{0} cm/sec). Soft data may be even less specific, for example, professional experience, though difficult to quantify, is valuable to site interpretation. Also data may be inferred by lack of observation or trends in observations (Journel, 1986). For example: if a drill hole is drilled to fifty feet and the water-table is not encountered, then, although the water-table has not been observed, data indicate the water-table must be deeper than fifty feet. This is fairly obvious, but not necessarily easy to quantify using simple kriging or automated contouring techniques. Through the use of Bayesian theory and indicator kriging these types of data can be utilized.

Geostatistics and Indicator Theory

For a computer to automatically create descriptions of the subsurface, rules have to be defined on how data will be treated and honored. The use of geostatistics, and in particular indicator kriging suit this process well. These techniques can be used to describe the spatial variation of data, define the probability of a parameter value at a given location, and assess the error associated with the parameter estimate at each location. Alabert (1987) gives a thorough discussion of the topic. Indicator kriging is a variant of ordinary kriging, where field measured parameters are transformed to indicator values at several cutoffs before they are kriged (e.g., cutoffs could be defined for hydraulic conductivities in a clay, sand, gravel system as: $K < 10^{-6}$ cm/sec (clay) as 1 indicator, $K < 10^{0}$ cm/sec (clay and sand) as another indicator, and $K < 10^{2}$ cm/sec (clay, sand, and gravel) as a third indicator. Indicators are defined as "1" for measured values below the cutoff, and "0" for values above the cutoff. Advantages of indicator kriging are: 1) it allows the modeler to evaluate continuity of the units of specified properties (particularly important is the ability to evaluate continuity of extreme values of parameters, such as a thin but continuous unit of high hydraulic conductivity), yet the most probable realization is not affected unduly by outlier data (Isaaks, 1989); and 2) qualitative data can be used to constrain the model; for example, relations such as gravel/sand/silt/clay or contaminated/non-contaminated can be kriged.

To use this method, the first step is to produce indicator semivariograms for each cutoff considered. To develop the semivariogram for each cutoff, the data are transformed into a series of 1's and 0's through the relation:

$$i_j = \begin{cases} 1 \text{ if } v_j \le v_c \\ 0 \text{ if } v_j > v_c \end{cases}$$

where i_j = the indicator value, v_j = the sample value, and v_c = the cutoff value (Isaaks, 1989). These values are then evaluated to determine the nugget, range(s), and sill(s) terms for given directions, window angles, and bandwidths. Once the semivariograms have been developed, the indicator data are kriged at each cutoff. The process determining the weights of sample data around a point being estimated is identical to that used in ordinary kriging, whether blocks or points are being evaluated:

$$F^*(v_c) = \Sigma w_j i(v_j); \text{ for indicator kriging estimates.}$$

$$Z^*(v) = \Sigma b_j v_j; \text{ for ordinary kriging estimates.}$$

w_i and b_i are weights, $F^*(v_c)$ and $Z^*(v)$ are kriged estimates of the indicator and the parameter respectively, and the summations are from 1 to the number of data points. To determine the indicator value at the prescribed point, a cumulative distribution function (cdf) is developed. In Figure 1, a simple example is shown for defining the cdf for an individual block. In this case, five samples are equally distant from the block (and within the range of influence), and therefore the weights are equal ($w_1 = w_2 = w_3 = w_4 = w_5 = 0.20$). Cutoffs were set at 0.02, 0.10, 0.13, and 0.26 cm/sec. Only one point is less than or equal to the first cutoff (0.02) so there is a 20% probability the value at the point is less than 0.02, 40% probability of being less than 0.10, 60% probability of being less than 0.13, and an 80% probability of being less than 0.26.

From the cdf, several tacks may be taken in evaluating the indicator data; 1) maps can be made to indicate the best estimate of parameter values (value defined equal the value equal to the 50% probability), 2) maps can illustrate the probability that the value of a parameter is above or below some specified value, 3) maps can indicate the parameter value above or below given a specified probability, or 4) realization maps can be made where the values are determined by randomly selecting the indicator for each location from the cdf determined from hard data and previously simulated points; this last option is a stochastic simulation.

Stochastic Simulation

The process of stochastic simulation, described by Gómez-Hernández and Srivastava (1990), takes advantage of the cdf's determined by indicator kriging, and Monte-Carlo techniques. The process is most clearly described using an example, in this case, the simulation of the occurrence of three geologic facies (clay, sand, and gravel). To generate an individual realization, a grid is defined. Starting with the first indicator range (e.g. clay), grid blocks at hard data locations are defined as either "1" (clay) or "something else" (e.g. "2"

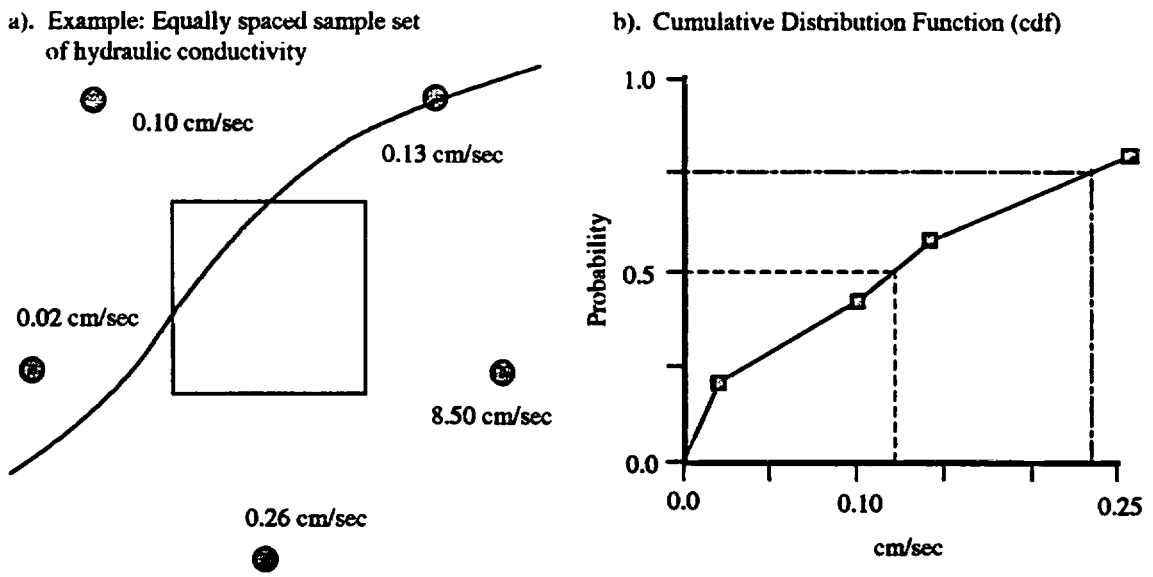

Figure 1. Using indicator kriging, with the block equidistant from five sample locations, the best indicator estimate (50% probability), determined from the cdf, for the block hydraulic conductivity (K) is 0.12 cm/sec. If ordinary kriging were used, the best estimate would be 2.21 cm/sec. This shows how indicator kriging does not allow high-value outlier data to overwhelm more plentiful, low-value data. The cdf may also be used to approximately indicate the probability the block K is below a certain value; here the cdf indicates that there is a 75% probability the K of the block is below 0.24 cm/sec. Use of stochastic simulation would yield a small percentage of realizations with a high K for the block. If enough blocks have the potential to be defined as high K, some realizations will exhibit a continuous zone of high K. This is how indicator kriging allows high value outlier data to be included in interpretations, without unduly biasing the results. The number of realizations which exhibit this condition, divided by the total number of realizations is the probability that such a condition exists.

= sand, "3" = gravel; in kriging calculations, these values are treated as "0" if another indicator is being evaluated or as "1" if it is the indicator currently under consideration). This ensures the simulation honors the locations where hard data indicate clay. At soft data locations the indicator is estimated from the field data and the probability of correct estimation is entered. Because of the many simulations at this location, clay will be present in about 70% of the realizations, sand 20%, and gravel 10%. Then a random starting location within the model grid is selected and indicator kriging is used to determine the indicator cdf for that location. The indicator value for that simulation is then selected from the cdf using a random number generator (e.g. if the location has a 70% probability of being clay, 20% sand, and 10% gravel, and a random number of 0.87 is generated, the location is defined as sand, "2"). The cdf, at this point, is based only of the hard and soft data (the soft data locations are now treated as hard because a value has been specified for that location in this simulation). If the random number is less than the probability the indicator value exists a "1" is assigned (i.e. clay is present at the grid location), otherwise a "0" is defined (another type of material is present). Next, another grid location, which has not been simulated, is selected at random and its indicator value is determined based on the hard and soft data, and the previously simulated indicator values (now considered hard data values for the rest of this realization). This process of selecting random grid locations and simulating an indicator value from the cdf created by the indicator kriging, based on the hard, soft, and previously kriged data, is continued until all grid locations are defined and a map of "1's" (clay) and "0's" (not clay) is created. The next indicator range is then selected (sand) and all the locations still containing "0's" are revisited and re-kriged (here the cdf is based only on the possibility the parameter value is sand or gravel). This revisiting and re-kriging process is repeated until all the indicator ranges have been evaluated and the map is composed of all "1's". An example of a stochastic simulation based on hard data from nine wells is shown in Figure 2. Note that the example shows a great deal of detail which reflects the natural character of material distributions. It is unlikely a modeler would generate a model with similar complexity without the aid of indicator kriging (this figure was created with the 3-D block visualization tool kit). The distribution and continuity of units can be quickly

Geostatistics and Indicator Theory

For a computer to automatically create descriptions of the subsurface, rules have to be defined on how data will be treated and honored. The use of geostatistics, and in particular indicator kriging suit this process well. These techniques can be used to describe the spatial variation of data, define the probability of a parameter value at a given location, and assess the error associated with the parameter estimate at each location. Alabert (1987) gives a thorough discussion of the topic. Indicator kriging is a variant of ordinary kriging, where field measured parameters are transformed to indicator values at several cutoffs before they are kriged (e.g., cutoffs could be defined for hydraulic conductivities in a clay, sand, gravel system as: $K < 10^{-6}$ cm/sec (clay) as 1 indicator, $K < 10^{0}$ cm/sec (clay and sand) as another indicator, and $K < 10^{2}$ cm/sec (clay, sand, and gravel) as a third indicator. Indicators are defined as "1" for measured values below the cutoff, and "0" for values above the cutoff. Advantages of indicator kriging are: 1) it allows the modeler to evaluate continuity of the units of specified properties (particularly important is the ability to evaluate continuity of extreme values of parameters, such as a thin but continuous unit of high hydraulic conductivity), yet the most probable realization is not affected unduly by outlier data (Isaaks, 1989); and 2) qualitative data can be used to constrain the model; for example, relations such as gravel/sand/silt/clay or contaminated/non-contaminated can be kriged.

To use this method, the first step is to produce indicator semivariograms for each cutoff considered. To develop the semivariogram for each cutoff, the data are transformed into a series of 1's and 0's through the relation:

$$i_j = \begin{cases} 1 \text{ if } v_j \le v_c \\ 0 \text{ if } v_j > v_c \end{cases}$$

where i_j = the indicator value, v_j = the sample value, and v_c = the cutoff value (Isaaks, 1989). These values are then evaluated to determine the nugget, range(s), and sill(s) terms for given directions, window angles, and bandwidths. Once the semivariograms have been developed, the indicator data are kriged at each cutoff. The process determining the weights of sample data around a point being estimated is identical to that used in ordinary kriging, whether blocks or points are being evaluated:

$$F^*(v_c) = \Sigma w_j i(v_j); \text{ for indicator kriging estimates.}$$

$$Z^*(v) = \Sigma b_j v_j; \text{ for ordinary kriging estimates.}$$

w_i and b_i are weights, $F^*(v_c)$ and $Z^*(v)$ are kriged estimates of the indicator and the parameter respectively, and the summations are from 1 to the number of data points. To determine the indicator value at the prescribed point, a cumulative distribution function (cdf) is developed. In Figure 1, a simple example is shown for defining the cdf for an individual block. In this case, five samples are equally distant from the block (and within the range of influence), and therefore the weights are equal ($w_1 = w_2 = w_3 = w_4 = w_5 = 0.20$). Cutoffs were set at 0.02, 0.10, 0.13, and 0.26 cm/sec. Only one point is less than or equal to the first cutoff (0.02) so there is a 20% probability the value at the point is less than 0.02, 40% probability of being less than 0.10, 60% probability of being less than 0.13, and an 80% probability of being less than 0.26.

From the cdf, several tacks may be taken in evaluating the indicator data; 1) maps can be made to indicate the best estimate of parameter values (value defined equal the value equal to the 50% probability), 2) maps can illustrate the probability that the value of a parameter is above or below some specified value, 3) maps can indicate the parameter value above or below given a specified probability, or 4) realization maps can be made where the values are determined by randomly selecting the indicator for each location from the cdf determined from hard data and previously simulated points; this last option is a stochastic simulation.

Stochastic Simulation

The process of stochastic simulation, described by Gómez-Hernández and Srivastava (1990), takes advantage of the cdf's determined by indicator kriging, and Monte-Carlo techniques. The process is most clearly described using an example, in this case, the simulation of the occurrence of three geologic facies (clay, sand, and gravel). To generate an individual realization, a grid is defined. Starting with the first indicator range (e.g. clay), grid blocks at hard data locations are defined as either "1" (clay) or "something else" (e.g. "2"

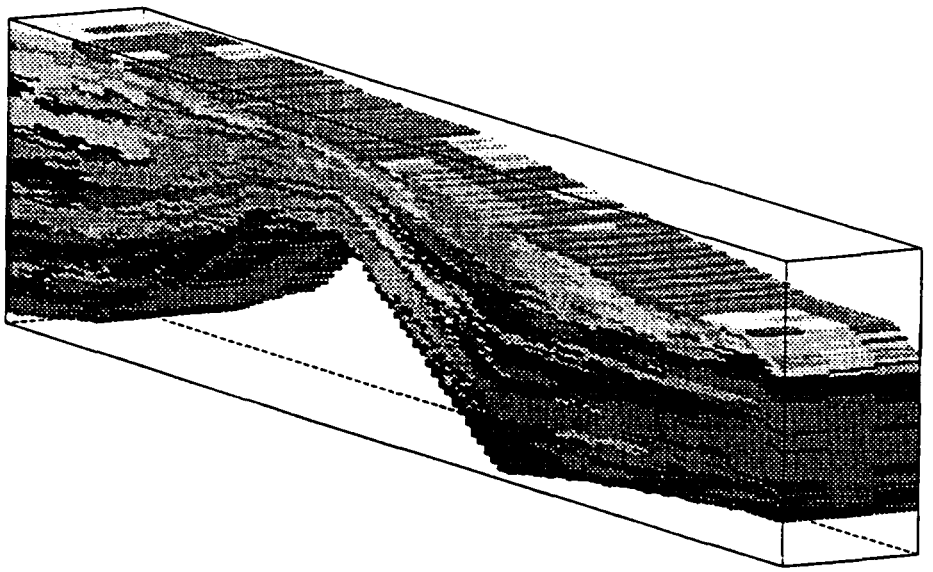

Figure 2. A three dimensional block diagram created by the software package shows the distribution of clay, sand, sandy-gravel, and gravel indicators (lighter colored blocks are the coarser materials). This is one stochastic simulation based on nine drill holes of hard data. The distributions are quite detailed and more complex than a modeler would define without the aid of indicator kriging. Ordinary kriging would smooth out the estimated values omitting thin but continuous units of high or low conductivity which may be critical to the nature and extent of contaminant migration.

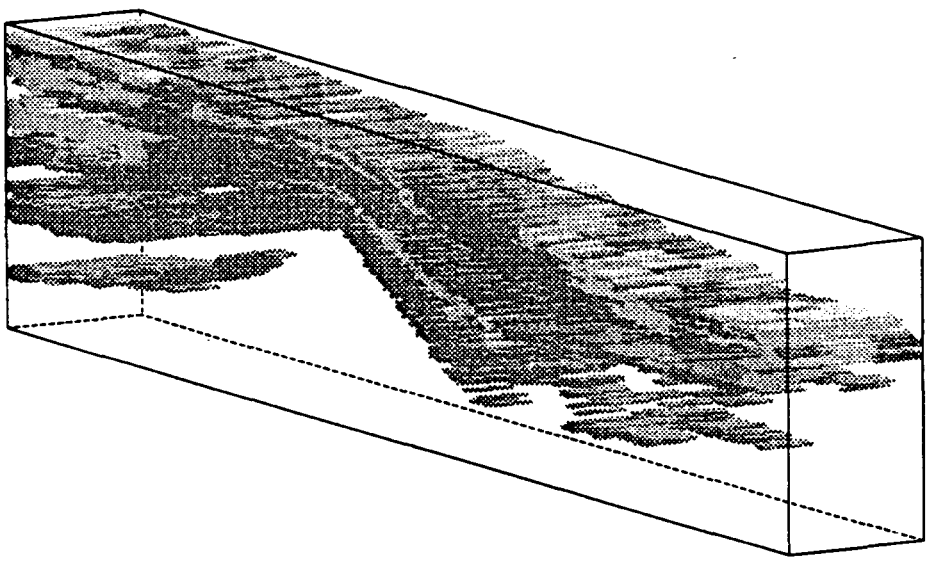

Figure 3. A three dimensional floating body diagram created with the software package shows the distribution of the high hydraulic conductivity (K) sandy-gravel and gravel indicators from the block diagram in Figure 2. With this tool, connectivity of units can be easily assessed. When concerned with contaminant transport, such a tool can aid in rapid assessment of continuous and discontinuous units. Zones of high K that are not connected to main pathways can be identified as the discontinuous gravel facies shown. If contamination enters these zones and they are surrounded by low K material, transport will not proceed rapidly because velocities will be controlled by the low K material. A thin slice of the site is shown. Simulation to a greater distance in the third dimension is required to determine if the gravel facies is continuous via connections in the third dimension.

determined by "peeling" the diagram to reveal a specified range of parameters such as the discontinuous gravel facies presented in Figure 3 or by using a simple program which calculates connected paths.

To create another realization this process is repeated, but a new random pattern is selected to cover the grid and cdf's are calculated and are utilized in a Monte-Carlo selection of an indicator at each grid location to be simulated. Alternative realizations can be created following this process until the desired number of simulations is obtained.

Each realization honors the statistics of the original data, and has equal probability of existing. These realizations can then be used as maps of parameters for use in modeling the site.

The Software Package

A computer software package is being developed to meet the design goals described above. The package will be public domain and portable to many computer systems, consequently easily incorporated into private companies and government agencies. The package is designed to be easy to use for groundwater modelers familiar and unfamiliar with geostatistics. Many practicing hydrologists do not have an understanding of advanced geostatistics so many of the processes must be automated if the package is to be useful; on the other hand, modelers familiar with the techniques will desire the ability to guide the solution using non-standard assumptions (i.e. fitting a semivariogram model with a nested spherical and exponential model instead of just a single linear or spherical model). Although the automated process must yield reasonable results, an expert must have the ability to manipulate the tools to refine the results. The package must also be able to import and export data in various formats, therefore the package will use ASCII data file formats, but locations where data are input and output within the source code will be identified so end users can easily modify the software to meet their data file format requirements. Eventually database support will be provided.

The Tool Kit

The software package is designed to run on UNIX[1] workstations using X-windows[2] and Motif[3]. UNIX work stations were chosen because the operating system can freely address all RAM on the computer (MS-DOS computers present programming problems when more than 64K and 640K of memory is needed, the latter is non-trivial to surmount) and if more memory is needed, the program may page[4] off the hard disk. Currently there are few UNIX workstations in small consulting firms, but with the growth of computer technology, and the ever decreasing hardware costs, they are becoming more common. Many workstations currently available are not much more expensive than Intel 80386/486[5] based IBM-compatible computers. Also IBM-compatible 80386/486 and Macintosh[6] computers can be setup to run under UNIX with X and Motif, though this solution is not ideal. Within the UNIX environment, X-windows is becoming a standard window manager and Motif is a common tool kit available from most computer hardware vendors. Therefore, this is a good design environment because from a cost view point it is, or soon will be, well within the reach of small consulting firms, and it is already a common environment in larger consulting firms, government agencies, and universities.

Features of the Software

This software package has a number of modules which allow a modeler to start with raw field data, or data prepared by other products, and guide the modeler through statistical analysis of the data, generation of multiple realizations of each data property, development of model grids to accommodate the geometry of each realization, kriging data properties into model grids, generation of data input files for flow and contaminant transport models, execution of models, and visualization of model results (Figure 4).

Classical statistical tools allow the modeler to examine the mean, median, and variance of the data set, create ANOVA[7] tables, examine frequency and probability distributions, and evaluate the normality or

---

[1] UNIX: Operation system initially developed by AT&T.

[2] X-windows: Window manager developed at MIT.

[3] Motif: A X-windows tool kit developed by the Open Software Foundation (OSF).

[4] Paging off the hard disk is a method used to increase RAM space, and is referred to as virtual memory.

[5] 80386/486 are computer processor microcopies used on many IBM-compatible computers, a product of Intel, Inc.

[6] Macintosh: Registered trademark of Apple Computers, Inc.

[7] ANOVA: ANalysis Of VAriance, In this table the regression sum of squares, error sum of squares, total sum of squares, regression mean square, error mean square, and degree of freedom terms are presented (Neter, Wasserman, and Kutner (1985)).

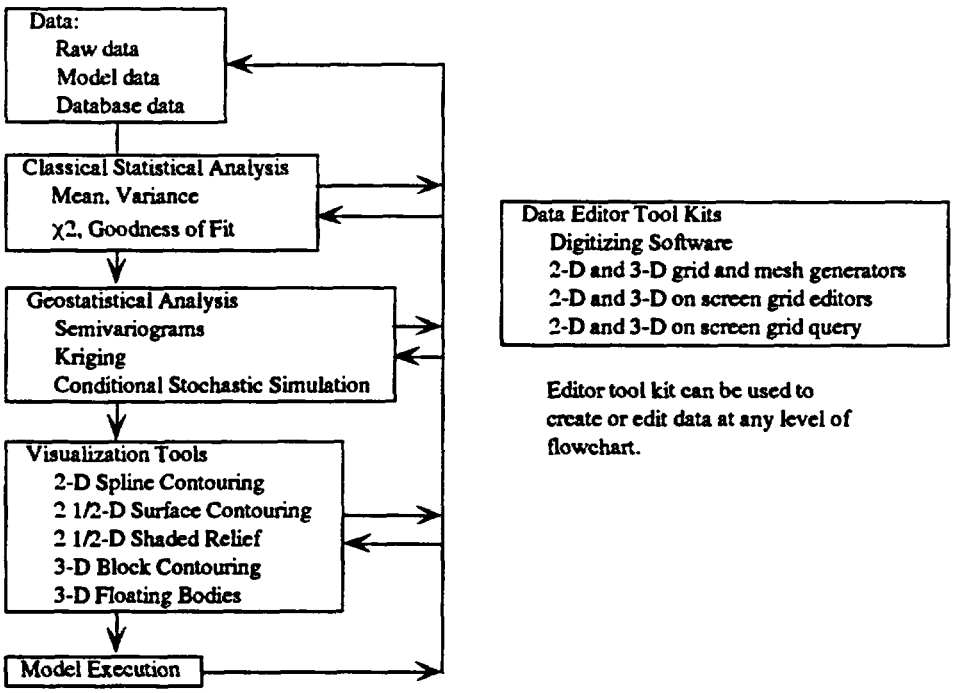

Figure 4. Flow chart of the analysis process from entering data, analyzing data, making stochastic simulations, and manipulating the data into formats acceptable for various groundwater and contaminant transport models. Once the models are executed, the output can be viewed by various methods, parts of the model can be interactively modified, and the models can be re-run.

log-normality of the data using a chi-squared test. Geostatistical tools include three-dimensional semivariogram and kriging analyses. The process of analyzing semivariograms can be automated and/or manually preformed using up to four nested structures. The automated process uses a least-squares or a constrained least-squares (forces sill to match variance and places greater weight on data points represented by more data pairs) algorithm to minimize the error of the model curve to the experimental semivariogram data. If the solution is not to the modelers liking, it can be manually modified. Solutions can be based on linear, linear with sill, spherical, exponential, logarithmic, Gaussian, or hole effect models. The nested models can be of different forms. Ordinary, universal, and indicator and Bayesian kriging options are available. Multiple indicator, conditional, stochastic simulations can be made to generate 10's to 1000's of realizations of the parameter being evaluated, which though different, all equally honor the measured data and their statistics. Kriging errors can be calculated so that locations of high uncertainty can be identified. This information can be used to guide further data acquisition or to temper dependence on results where simulations in those areas are critical.

Once the data have been analyzed, and the distributions defined, the parameter values may be merged into modeler defined finite-element meshes or finite-difference grids (options will be available to base the grids and meshes on the kriged data, or to krige the data directly into the grids and meshes). Once these grids are created, this data can be automatically converted into data sets for groundwater flow and/or contaminant transport models such as MODFLOW[8] or CFEST[9]. For flow and transport codes that are not supported by the software, guidelines will be available to modify the source code so that input to the new program can be automated. These programs can be executed for each realization of the subsurface. The results can then be examined individually using two-dimensional (contour maps), two and a half-dimensional (surface contour and

[8] MODFLOW: A Modular Three-Dimensional Finite-Difference Groundwater Flow Model by M.G. McDonald and A.W. Harbaugh, U.S. Geological Survey OFR 83-875, 1984.

[9] CFEST: Coupled Fluid, Energy, and Solute Transport Model. Developed by Battelle Memorial Institute for the U.S. Department of Energy under the Civilian Radioactive Waste Management Program.

shaded relief maps), and three-dimensional (block and floating body (a block map with specified units removed) maps) visualization tools. The results from various simulations can also be combined to determine the probability that a particular condition will exist at a given location (e.g. the probability the contaminant will be greater than 10 ppm at location X at a given time), or to determine the range of results (e.g. all simulations resulted in contaminant concentrations between 5 ppm and 150 ppm at location Y at time t). These analyses can be used to examine best case, worst case, and most likely case scenarios. By going back several steps, assumptions about a particular parameter can be adjusted and the simulations easily rerun. This will aid in determining the sensitivity of the model to variations of input. The model may also be modified; "what if's" can be easily analyzed, e.g. the pumping rate is changed or well X is moved. "What if's" can be used to improve the design of the remediation facility. Many of the simulations may not yield the same hydraulic and/or contaminant conditions as measured in the field; these realizations can be removed from further analysis.

Inverse techniques can eventually be used to aid the modeler in calibrating each subsurface realization. By supplying the model with prior estimates of head conditions, recharge rates, distribution of materials, and hydraulic conductivity; programs such as CFEST-INV[10] and MODFLOWP[11] can be used to automatically calibrate the model by modifying estimates of hydraulic parameters. This technique can also be used to eliminate impossible or unlikely realizations; if the model doesn't converge within reasonable limits, it is unlikely the predicted kriged distribution of materials for that realization is appropriate.

In addition to the statistical tools, tools will be available for on-screen generation, editing, and query of model grids and meshes, and Postscript[12] printing of graphical displays.

Summary

The software package presented herein is needed for two reasons: 1) because groundwater professionals in government, industry, and academia need economic tool kits for quickly and conveniently visualizing subsurface data whether it be field data or model output, and 2) because tools which speed the process of site evaluation, address the uncertainties inherent in describing and modeling the subsurface, and allow the modeler a means of viewing the results are essential for properly modeling contaminant problems. The ability to view the spatial distribution of data and results is essential to the future of subsurface evaluation; 1) humans can generally understand a picture better than a series of numbers, though the raw numbers must be examined to see the fine details, and 2) pictures can be used effectively to make points or convey messages to people untrained in hydrogeology, contaminant remediation, or geology. This software gives hydrogeologists with varying degrees of geostatistical knowledge the ability to evaluate uncertainty associated with their modeling and their design of remediation facilities. A better result will be obtained by an individual more experienced in hydrogeology, geostatistics, and modeling. We do not intend this package to replace expertise, rather it is an aid to hydrogeologists and perhaps an inspiration for them to tackle the tedious chore of evaluating uncertainty and perhaps learn more about geostatistics or to work more closely with geostatisticians.

References

Alabert, F. G., 1987, Stochastic Imaging and Spatial Distributions Using Hard and Soft Information, Master's Thesis.

Freeze, R.A. and J.A. Cherry, 1979, Groundwater, Prentice-Hall, Englewood, New Jersey, pg. 604.

Gómez-Hernández, J.J. and R. M. Srivastava, 1990, ISIM3D: An ANSI-C Three Dimensional Multiple Indicator Conditional Simulation Program, Computers in Geoscience, Vol. 16, No. 4 pp 395-414.

Journel, A.G., 1986, Constrained Interpolation and Qualitative Information - The Soft Kriging Approach, Mathematical Geology, Vol. 18, No. 3, pp 269-286.

Isaaks, E.H. and R.M. Srivastava, 1989, An Introduction to Applied Geostatistics, Oxford University Press, New York.

Neter, J., W. Wasserman, and M.H. Kutner, 1985, Applied Linear Statistical Models, Second Edition, IRWIN, Homewood, Illinois.

---

[10] CFEST-INV: Coupled Fluid, Energy, and Solute Transport - Inverse Model. Developed by Battelle Memorial Institute for the U.S. Department of Energy under the Civilian Radioactive Waste Management Program.

[11] MODFLOWP: A Modular Three-Dimensional Finite-Difference Groundwater Flow Parameter Estimation Model by the U.S. Geological Survey, to be released in 1991.

[12] Postscript is trademark and a printer programming language developed by Adobe Systems, Inc. Postscript files may be printed on many different printers by different manufacturers without special printer drivers.