SANDIA REPORT

SAND2000-3200
Unlimited Release
Printed February 2001

Probabilistic Approach to Site Characterization: MIU Site, Tono Region, Japan

Sean A. McKenna

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/ordering.htm

SAND2000-3200
Unlimited Release
Printed February 2001

Probabilistic Approach to Site Characterization: MIU Site, Tono Region, Japan

Sean A. McKenna
Geohydrology Department
Sandia National Laboratories
PO Box 5800
Albuquerque, New Mexico 87185-0735

Abstract

Geostatistical simulation is used to extrapolate data derived from site characterization activities at the MIU site into information describing the three-dimensional distribution of hydraulic conductivity at the site and the uncertainty in the estimates of hydraulic conductivity. This process is demonstrated for six different data sets representing incrementally increasing amounts of characterization data. Short horizontal ranges characterize the spatial variability of both the rock types (facies) and the hydraulic conductivity measurements. For each of the six data sets, 50 geostatistical realizations of the facies and 50 realizations of the hydraulic conductivity are combined to produce 50 final realizations of the hydraulic conductivity distribution. Analysis of these final realizations indicates that the mean hydraulic conductivity value increases with the addition of site characterization data. The average hydraulic conductivity as a function of elevation changes from a uniform profile to a profile showing relatively high hydraulic conductivity values near the top and bottom of the simulation domain. Three-dimensional uncertainty maps show the highest amount of uncetainty in the hydraulic conductivity distribution near the top and bottom of the model. These upper and lower areas of high uncertainty are interpreted to be due to the unconformity at the top of the granitic rocks and the Tsukyoshi fault respectively.

Acknowledgements

This report benefited from thoughtful reviews by Shinji Takeuchi, Hiromitsu Saegusa, Katsushi Nakano, Susan Altman and Erik Webb. This work was funded by the Japan Nuclear Cycle Development Institute through a contract with Sandia National Laboratories. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Table of Contents

List of Figures

List of Tables

Introduction

Site characterization is the process by which information is obtained regarding the physical properties of a site. This process is most often conducted by drilling holes into the subsurface and obtaining samples of the material at different depths. The goals of site characterization can be to simply learn more about the site (e.g., defining the local stratigraphy), to map or define areas of special interest within the subsurface (e.g., regions of hydraulic conductivity below $1x10^{-12}$ m/s), and/or to define regions of uncertainty with respect to a site performance measure. This latter goal is often used to optimize any additional characterization efforts with the objective of reducing the amount of uncertainty.

The amount of site characterization that can be accomplished is constrained by several factors. The main difficulty in site characterization activities is that there are never enough resources to collect all the data that are necessary to eliminate uncertainty in the spatial distribution of properties within the subsurface. The amount of material that is sampled in site characterization boreholes is typically only $1x10^{-6}$ to $1x10^{-9}$ of the total site volume (Journel and Alabert, 1989). This limited amount of sampling leads to uncertainty in the spatial distribution of material properties and to uncertainty in decisions that must be made with respect to those properties.

One approach for modeling the uncertainty in the spatial distribution of rock properties is geostatistical simulation. Geostatistical simulation defines the spatial uncertainty in rock properties by creating multiple, equiprobable realizations of the rock properties. Each realization is conditioned to the available characterization data and each realization is a possible description of the actual distribution of rock properties in the subsurface. Consideration of multiple realizations makes it possible to define the uncertainty in the distribution of rock properties. The geostatistical simulation approach to modeling rock properties has previously been applied to both site characterization and performance assessment problems at nuclear waste repositories in the United States (Rautman and McKenna, 1997; Lavenue and RamaRao, 1992).

The main objective of this report is to apply geostatistical simulation techniques to define the spatial distribution of hydraulic conductivity in the subsurface at the MIU site. Analysis of the geostatistical simulations will also be conducted to examine the uncertainty in the spatial distribution of the hydraulic conductivity distributions. Specific goals of this study are to examine the changes in the mean (average) estimate of the hydraulic conductivity and in the uncertainty in the hydraulic conductivity values as a function of increasing amount of characterization data. These specific goals are accomplished by creating six sets of geostatistical simulations using six different subsets of characterization data from both the region surrounding the MIU site and from within the MIU site boundaries.

Site Description

The MIU site is located near the town of Mizunami in central Japan. The site has been chosen as the location for the construction of an underground research facility by JNC. Site characterization for the underground research facility has begun and construction of the main shaft of the research facility is scheduled to begin in 2001. The MIU site will be constructed in the Toki granite that lies below the Mizunami Group, a series of sedimentary rocks. There is an unconformity separating the sedimentary and granitic rocks. More details of the geology in the vicinity of the MIU site can be found in Yusa, and others (1992).

For the geostatistical modeling efforts described in this report, it was necessary to define a three dimensional, rectangular volume that would surround the MIU site. This volume and the number of gridblock cells necessary to discretize this domain in the geostatistical modeling is shown in Figure 1. The total number of gridblocks is 712,472. Each gridblock is 8.0 by 8.0 by 6.5 meters. The gridblock size was chosen to best represent the volume of aquifer examined by a single packer test. The majority of the packer tests were conducted with a packer interval of 6.5 meters. The origin of the domain (lower, front-left corner) is located at the coordinates: 5256.0, -69008.0, -786.0. The z coordinate is the elevation with respect to mean sea level.

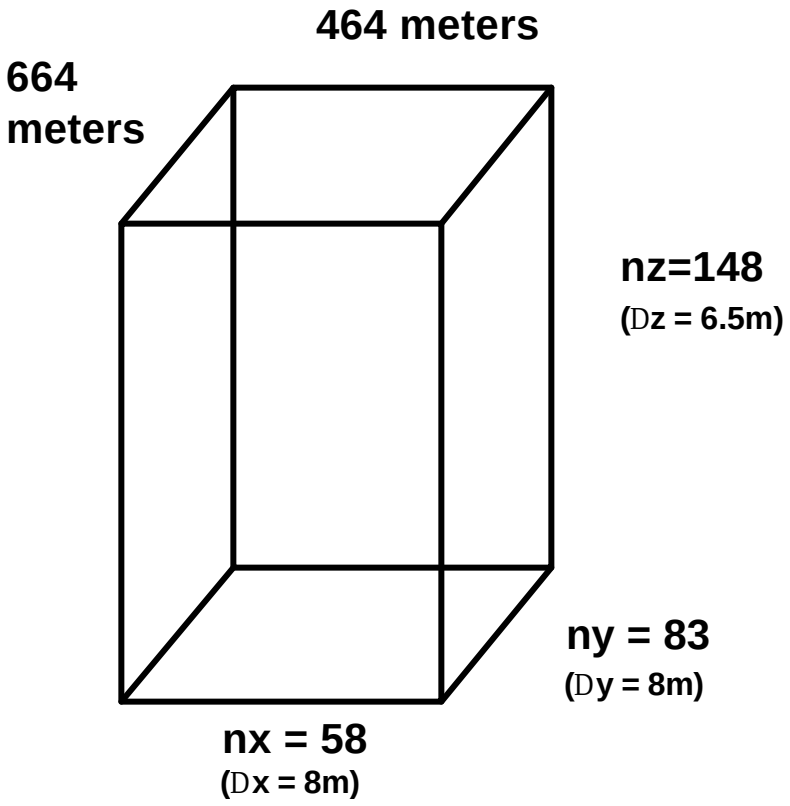

Figure 1. Description of the three-dimensional domain surrounding the MIU site used for the geostatistical modeling done in this report.

The upper boundary of the site is the unconformable contact between the Cretaceous granite and the overlying Miocene sedimentary rocks. The elevation of this contact varies across the site. For the modeling done in this report, the elevation of this contact is defined using data from a number of uranium exploration boreholes. The contact information in these boreholes is interpolated onto a regular grid as shown in Figure 2.

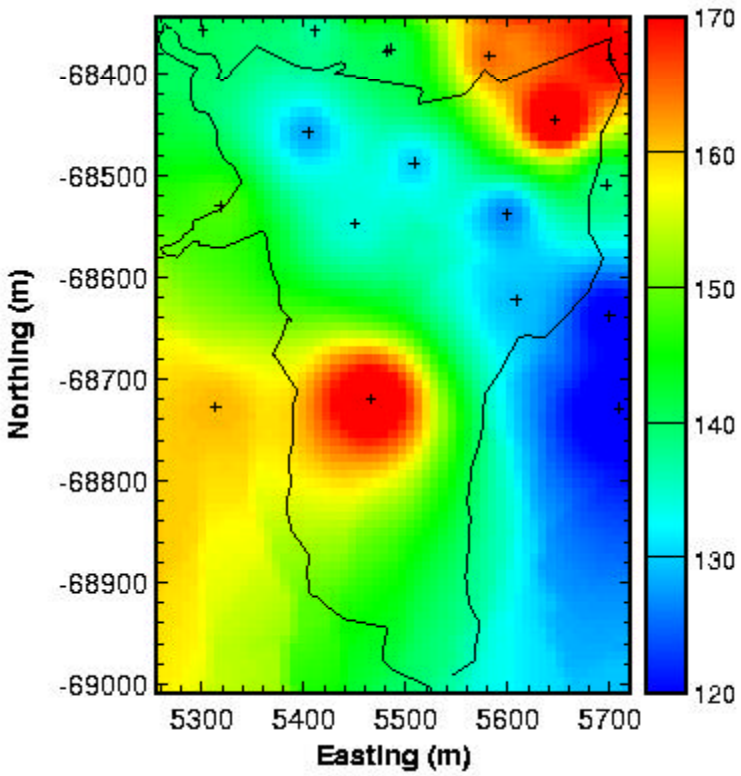

Figure 2. Elevation (in meters above sea level) of the contact between the granite and sedimentary rock in the area surrounding the MIU site. The "+"'s denote the locations of the uranium exploration boreholes used to create this map. The boundary of the MIU site is shown as a solid line.

The contact between the two formations provides the upper boundary of the MIU site for modeling purposes. The lateral extent of the MIU site is defined by the boundary lines provided by JNC and shown in Figure 2. In order to use this boundary in the geostatistical modeling, it is necessary to redefine this boundary in terms of the gridblocks in the geostatistical simulations. The lateral extent of the MIU site as defined on these gridblocks is shown in Figure 3.

A goal of this modeling exercise is to examine the changes in the hydraulic conductivity estimates and the uncertainty about those estimates as a function of increasing site characterization data. In order to achieve this goal, six different data sets are defined.

The first data set includes information from only the 14 boreholes outside of the MIU site boundary. This original data set includes information from TH-2, TH-3, TH-4, TH-5, TH-6, TH-7, TH-8, AN-6, SN-4, DH-5, DH-6, DH-7, DH-8, and DH-9. Each additional data set contains the original 14 boreholes, plus an additional borehole from within the MIU site boundary. The definition of these data sets is given in Table 1.

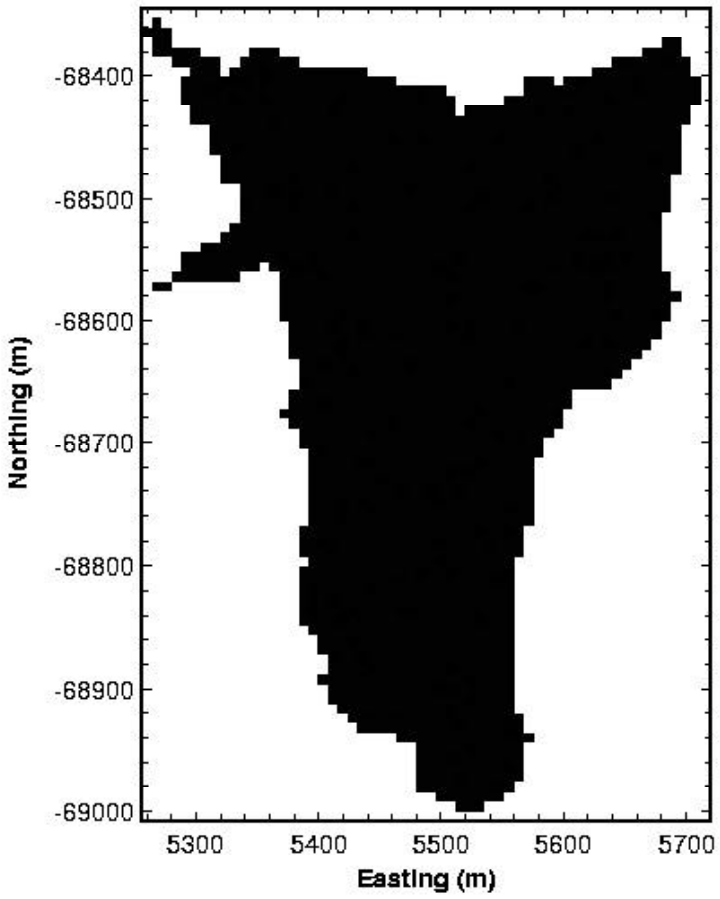

Figure 3. Lateral extent of the MIU site as discretized within the geostatistical-modeling domain.

Table 1. Definition of data sets used in site characterization exercise.

Data Set	Additional Borehole
14 holes	Original Data Set
15 holes	AN-1
16 holes	AN-3
17 holes	MIU-1
18 holes	MIU-2
19 holes	MIU-3

Available Data

There are three types of data that can be used in the geostatistical descriptions of the hydraulic conductivity at the MIU site: The locations of the packer test intervals, the values of hydraulic conductivity measured in the test intervals and geophysical logs obtained in the boreholes. Ideally, all of these different pieces of subsurface information can be combined into a coherent picture of the subsurface hydraulic conductivity distribution. Each of these different data sets is described below.

Location of Test Intervals

The choice of hydraulic testing intervals within each different borehole are based on several different factors including loss of circulation during drilling, the results of borehole flowmeter surveys, temperature logs, borehole televiewer logs, geological judgement, and, to a lesser degree, the geophysical logs. At this time, there is not a completely objective process for determining the locations of the hydraulic test intervals. The locations of the test intervals, based on the above information, are used in the indicator geostatistical simulations to define the locations of high and low hydraulic conductivity granite.

For each of the six data sets, the fraction of the total borehole length that is contained within the hydraulic test intervals is shown in Table 2. The values in Table 2 show that the fraction of the boreholes within test intervals is just less than 25 percent of the total borehole lengths. Extension of these one-dimensional data to the three dimensional subsurface indicates that the hydraulic conductivity values measured within these intervals applies to 25 percent or less of the subsurface material. The remaining 75 percent of the material has not been tested.

Table 2. Fraction of the data sets within hydraulic conductivity test intervals.

Data Set	Fraction of Total Length within Test Intervals	Fraction of Total Length outside Test Intervals
14 holes	0.232	0.768
15 holes	0.243	0.757
16 holes	0.235	0.765
17 holes	0.218	0.782
18 holes	0.217	0.783
19 holes	0.254	0.746

Hydraulic Conductivity Measurements

The full hydraulic conductivity data set provided by JNC was examined to determine the number of single-hole hydraulic conductivity tests performed within the granite. All tests performed within the sedimentary rocks overlying the granite were removed from the data set. A total of 208 single-hole slug tests were conducted in packed off intervals

within the granite both within the MIU site and in the region surrounding the MIU site. The probability density function (histogram) of these data (the 19 hole data set) is shown in Figure 4. The parametric distribution that best defines these hydraulic conductivity data is a lognormal distribution with a mean of –8.19 and a standard deviation of 1.66 in log10 space.

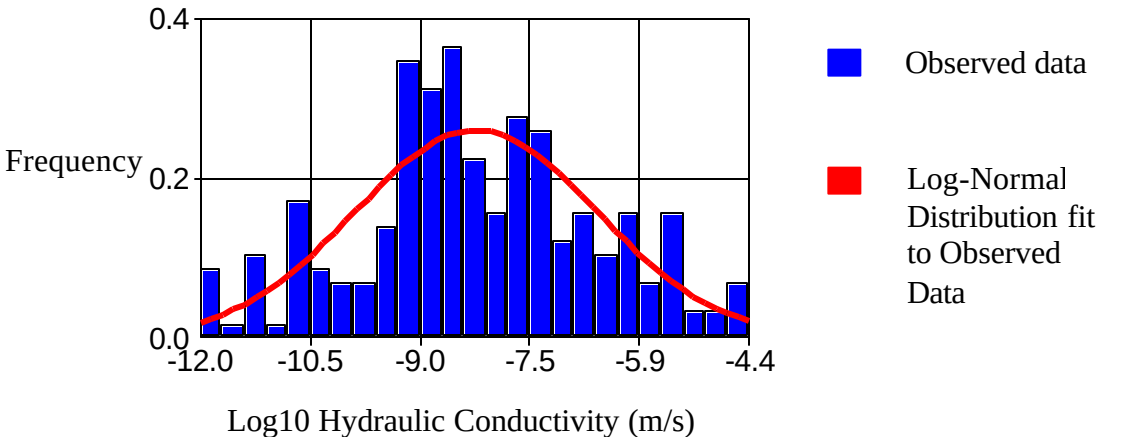

Figure 4. Probability density function (histogram) of the 208 hydraulic conductivity values measured with slug tests in the Toki granite.

The distribution of the hydraulic conductivity data changes as a function of the number of boreholes in the data set. The parameters describing the distribution of hydraulic conductivity for each data set are given in Table 3. It is interesting to note that as the number of boreholes increases, the mean of the hydraulic conductivity data also increases. This increase is approximately one-half order of magnitude from the 14 hole data set to the 19 hole data set. This increase may be due to the site characterization team becoming more proficient at selecting zones of high hydraulic conductivity as the site characterization progressed. The standard deviation of the hydraulic conductivity data also increases with increasing boreholes, but the 14 hole and 19 hole data sets have the same standard deviation values. The distribution of hydraulic conductivity shown in Figure 4 is only representative of the granite that was tested, not the entire subsurface.

Geophysical Logs

A full suite of geophysical logs has been collected in DH boreholes and in the boreholes within the MIU site. These geophysical logs include gamma, resistivity, neutron, single-point resistivity, density and sonic logs as well as temperature logs. Borehole televiewer logs have also been collected and these can be used to count different classes of fractures within the boreholes.

Table 3. Parameters describing the hydraulic conductivity data for each data set.

Data Set	Mean	Median	Standard Deviation	Number of Data
14 Holes	-8.68	-8.62	1.66	94
15 Holes	-8.53	-8.61	1.68	126
16 Holes	-8.35	-8.46	1.71	150
17 Holes	-8.27	-8.43	1.72	170
18 Holes	-8.21	-8.38	1.72	190
19 Holes	-8.19	-8.33	1.66	208

At this point, the geophysical logs are not used to constrain the geostatistical simulations of hydraulic conductivity within the MIU site. In the future, it would be desirable to extract as much information as possible from the geophysical logs regarding the hydraulic conductivity and use this information to constrain the hydraulic conductivity simulations. A proposed approach to extract this information, using neural networks, is outlined in the Discussion section of this report.

Geostatistical Simulations

A thorough introduction to geostatistical techniques is beyond the scope of this report. This report assumes that the reader has some basic knowledge of geostatistics including the concepts of variograms and stochastic simulation. Additional information on the basic theory and practice of geostatistics can be found in Journel and Huijbregts (1978), Isaaks and Srivastava (1989) and Goovaerts (1997). This report uses the geostatistical algorithms coded into the GSLIB software. The GSLIB book (Deutsch and Journel, 1998) provides an excellent introduction and user's manual for these algorithms.

For each data set, it is necessary to calculate and model variograms for both the indicator data, defining the locations of the high and low hydraulic conductivity granite, and the hydraulic conductivity data. The variograms are calculated using the equation:

$$g(h) = \frac{1}{2n(h)} \sum_{i=1}^{n} (x_i - (x_i + h))^2$$

where x and $x+h$ are two hydraulic conductivity values separated by the distance h. The $g(h)$ value on the left-hand side of the equation can be stated as: "one-half the average squared difference between all data values located a distance h away from each other". The values of $g(h)$ are plotted as a function of h. This type of plot is known as the "experimental variogram". In order to use the variogram information within the geostatistical simulation software, it is necessary to fit a model to the experimental variogram.

For each data set defined in Table 1, a total of 100 geostatistical simulations are created. 50 of these realizations are indicator realizations describing the spatial distribution of the

granite that is of a high enough hydraulic conductivity to be tested. These simulations are conditioned to the locations of the test intervals as measured in the different boreholes. It is necessary to divide the granite into two "facies" through this indicator transform because only 21 to 25 percent of the rock is tested hydraulically and the hydraulic conductivity values derived from the tests only apply to that fraction of the rock. The lower hydraulic conductivity facies is not tested and therefore the hydraulic conductivity values from the tests do not apply to this portion of the granite. In this report, it is assumed that all high hydraulic conductivity facies that intersect a borehole were tested hydraulically. The other set of 50 simulations for each data set defines the spatial distribution of the hydraulic conductivity as measured in the test intervals.

The two sets of simulations are combined to produce 50 three-dimensional models of the spatial distribution of hydraulic conductivity within the MIU site for each of the 6 data sets. It is these 300 final hydraulic conductivity models that are analyzed to determine the mean hydraulic conductivity maps and the uncertainty maps for each data set. The process of creating the geostatistical simulations and combining them together into the final models of hydraulic conductivity is described in the following pages.

Indicator Simulations

The presence of the test intervals along each borehole is defined using an indicator transform. At each location along the borehole, the location is assigned a “1” if that location is within a test interval or a “0” if that location is outside of a test interval. This indicator classification process can be seen as defining two “facies” within the granitic rock: the presumably higher conductivity facies that is tested hydraulically and the presumably lower conductivity facies that is not chosen for hydraulic tests. The distribution of the two facies within the boreholes of the MIU site is shown in Figure 5.

Variogram analysis is done on the indicator data for each of the six different data sets (see example variograms in Figure 6). A short initial range of approximately 20 -50 meters in both the vertical and horizontal directions characterizes the indicator variograms. The vertical variogram then shows a slow climb to a final range (roughly 1500 meters) that is greater than the vertical dimension of the MIU site domain. The horizontal variogram has a final variogram range value in the 100 to 400 meter range (see example in Figure 6). These variograms are fit with a nugget value of 0.05 and two nested spherical models. The first model defines the short-range variability and the second model defines the larger range and the final sill value. The spherical model was chosen as it is the easiest variogram model to interpret and it has proven to be a robust model for small data sets in previous studies. All variograms in this report were calculated and modeled using the *vario* and *variofit* packages in the UNCERT software (Wingle et al, 1999). The *variofit* package allows for several automatic variogram-fitting options; however, all variogram models shown in this report were fit by manually adjusting the variogram model parameters.

A typical example of these models fit to experimental indicator variograms in the vertical and horizontal directions is shown in Figure 6. A single isotropic horizontal variogram was used. It is possible that the hydraulic conductivity data have some directions of preferential correlation (anisotropy), but there are not enough boreholes to determine whether or not this horizontal anisotropy exists.

These variogram models are used to create conditional stochastic realizations of the distribution of high and low permeability facies throughout the three-dimensional domain shown in Figure 1. An example realization of the facies distribution is shown in Figure 7. This realization was created with the 19 hole data set. Note that the proportion of high hydraulic conductivity facies in Figure 7 is 25 percent and the proportion of the low hydraulic conductivity facies is 75 percent. The somewhat random distribution of the high hydraulic conductivity facies is a result of the relatively short range in the horizontal variogram.

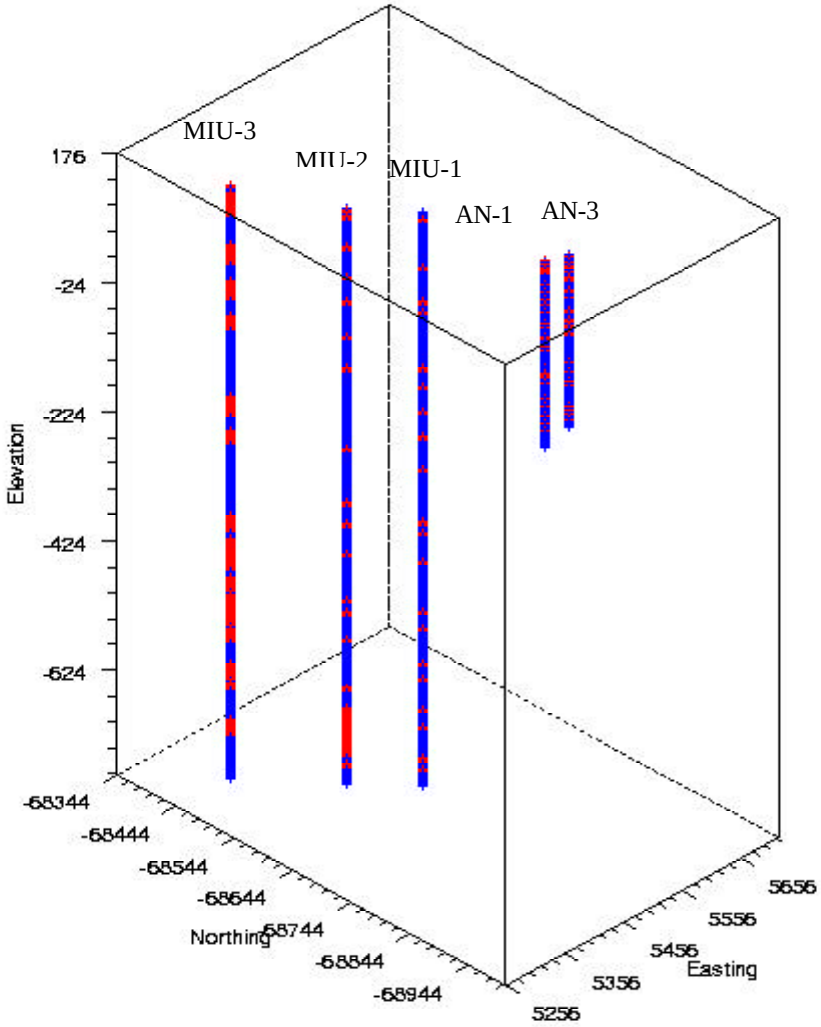

Figure 5. Distribution of the test intervals within the MIU site model domain. Red indicates a test interval and blue indicates zones that were not tested. These data provide the basis for the indicator simulations of high and low conductivity facies.

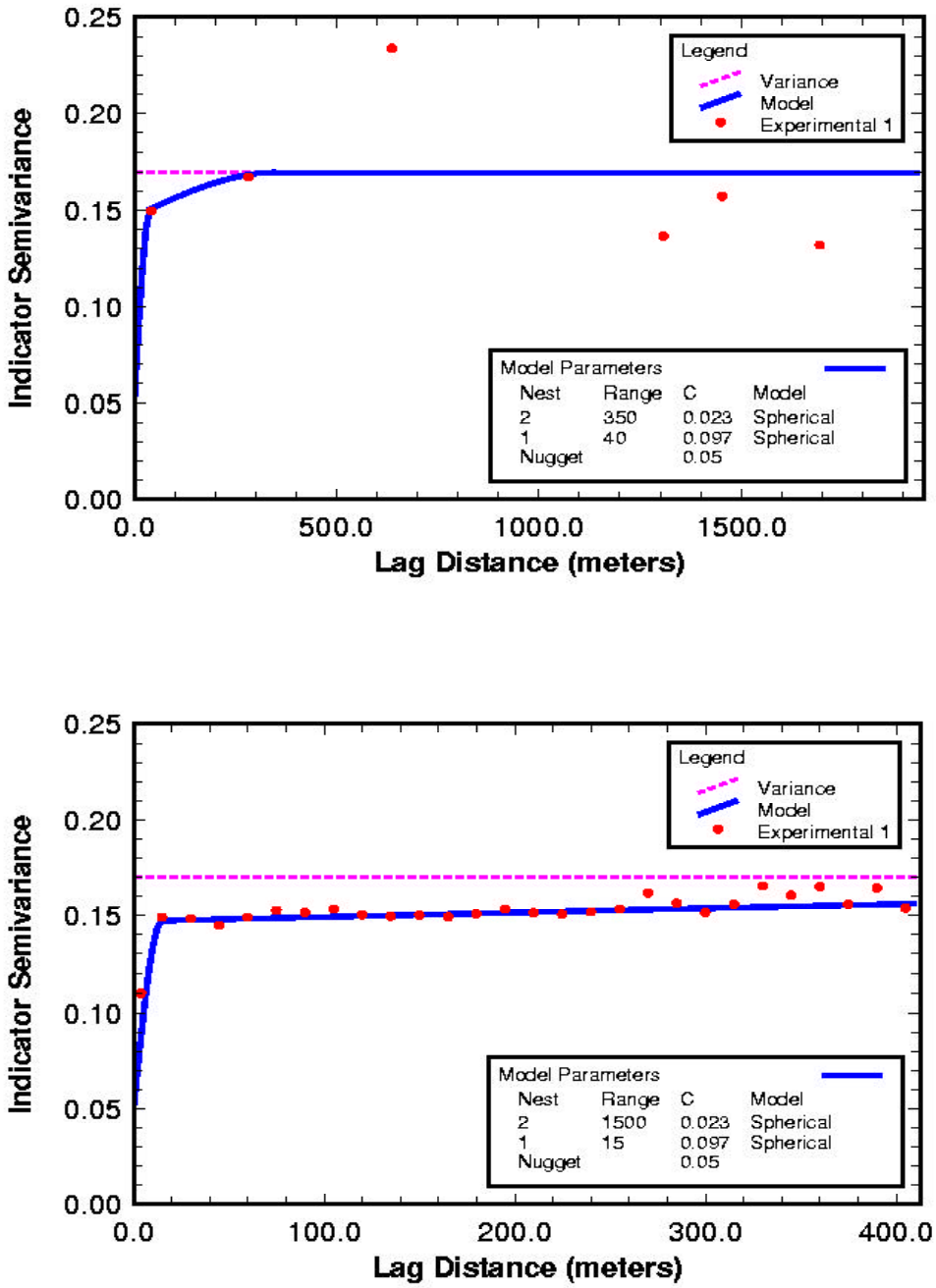

Figure 6. Indicator variograms for the vertical (upper image) and horizontal (lower image) directions as calculated on the 17 hole data set. These variograms are shown as typical examples of all the indicator variograms. Each experimental variogram (red dots) is fit with a double-nested spherical variogram model (blue line).

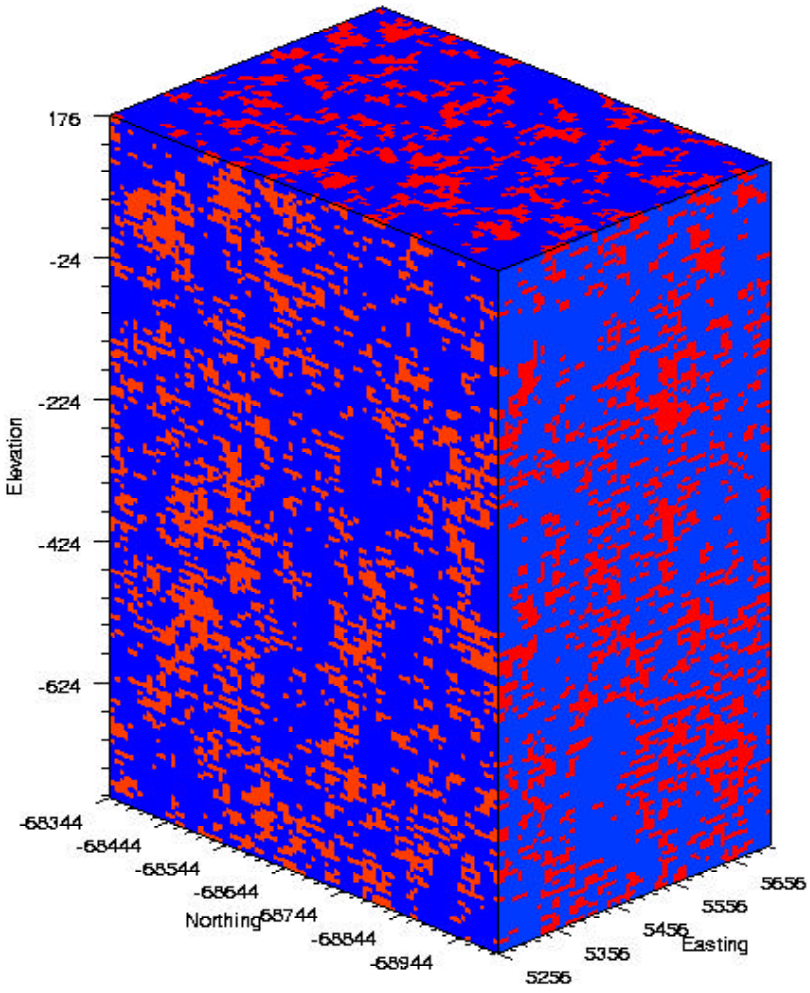

Figure 7. Example indicator simulation showing the distribution of high conductivity (red) and low conductivity (blue) facies within the granite. This example is realization number 1 created with the 19 hole data set.

Hydraulic Conductivity Simulations

Single-hole hydraulic conductivity tests were completed in a total of 208 test intervals within the Toki granite in the area around and within the MIU site. The spatial distribution of these hydraulic conductivity measurements for only the boreholes within the MIU site is shown in Figure 8. It is important to keep in mind that these measurements only define the hydraulic conductivity within the class of rock that has been subject to testing—the high hydraulic conductivity facies. There is also uncertainty in the actual values of the hydraulic conductivity due to interpretation of the slug tests; however, for the geostatistical simulations conducted in this report, this uncertainty is not taken into account.

Prior to the variogram analysis, it is necessary to transform the hydraulic conductivity data distribution into a standard normal distribution. The standard normal distribution has a mean of 0.0 and variance of 1.0. This transform is necessary in order to use the Guassian geostatistical simulation algorithm, *sgsim*, in the GSLIB software library. This transformation is done for each of the six different data sets using the *nscore*-transform software included in GSLIB (Deutsch and Journel, 1998). The simulation algorithm, *sgsim*, completes the geostatistical simulation in the standard-normal space and then transforms the data back to the raw hydraulic conductivity space. The entire transformation process is transparent to the user with the exception of the requirement to model the variograms in standard-normal space.

The normal-score transformed hydraulic conductivity variograms are similar to the indicator variograms in that they are characterized by a short initial range (the first spherical model) in both the vertical and horizontal directions. The vertical variogram then shows a slow climb to a final range that is approximately equal to the vertical dimension of the MIU site domain (approximately 1000 meters). The horizontal variogram has a final variogram range value in the 100 to 300 meter range. These variograms are fit with a nugget value of 0.10 and two nested spherical models. The first model defines the short-range variability and the second model defines the larger range and the final sill value. A typical example of these models fit to the normal-score experimental variograms in the vertical and horizontal directions is shown in Figure 9. A single isotropic horizontal variogram was used. It is possible that the hydraulic conductivity data have some directions of preferential correlation (anisotropy), but there are not enough boreholes to determine whether or not this horizontal anisotropy exists.

These variogram models are used to create conditional stochastic realizations of the distribution of hydraulic conductivity values throughout the three-dimensional domain shown in Figure 1. Each realization is conditioned to the measured hydraulic conductivity values at their respective locations. An example realization of the hydraulic conductivity distribution is shown in Figure 10. This realization was created with the 19 hole data set. The somewhat random distribution of the hydraulic conductivity is a result of the relatively short range in the horizontal variogram.

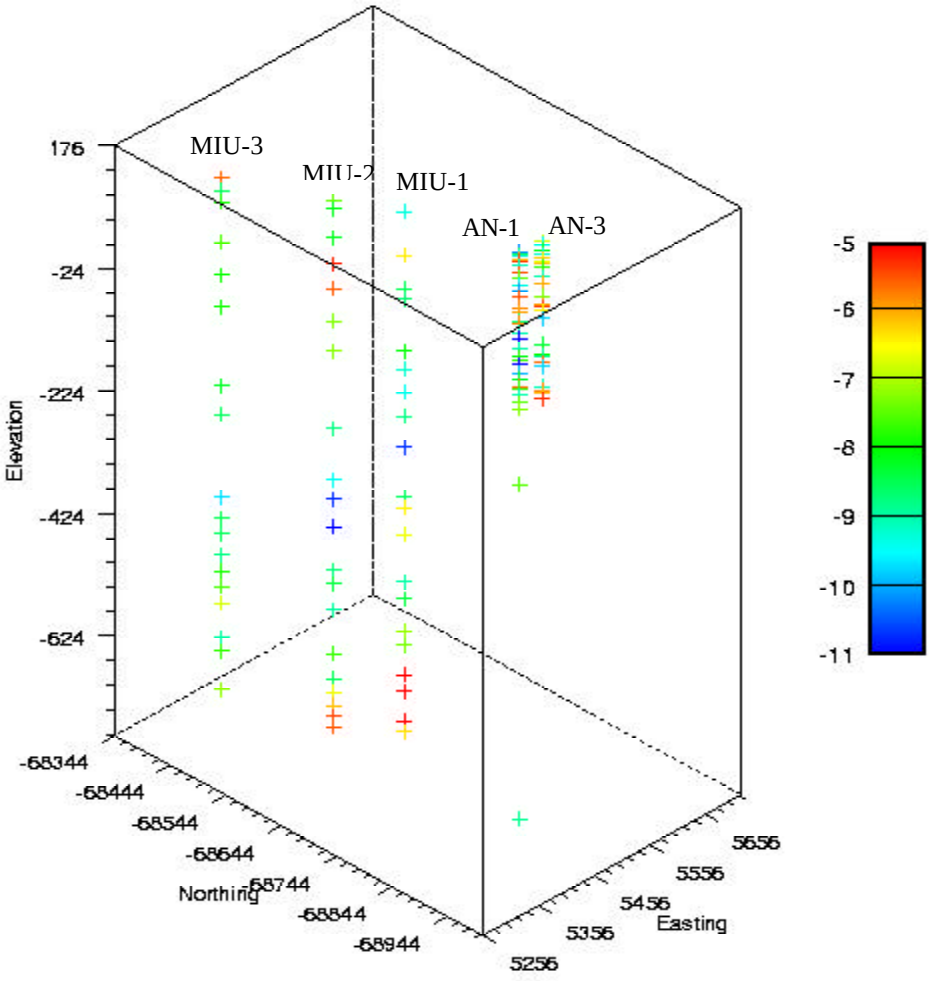

Figure 8. Distribution of the hydraulic conductivity measurements within the MIU site model domain. The color legend indicates the log10 hydraulic conductivity in m/s. These data provide the basis for the hydraulic conductivity simulations.

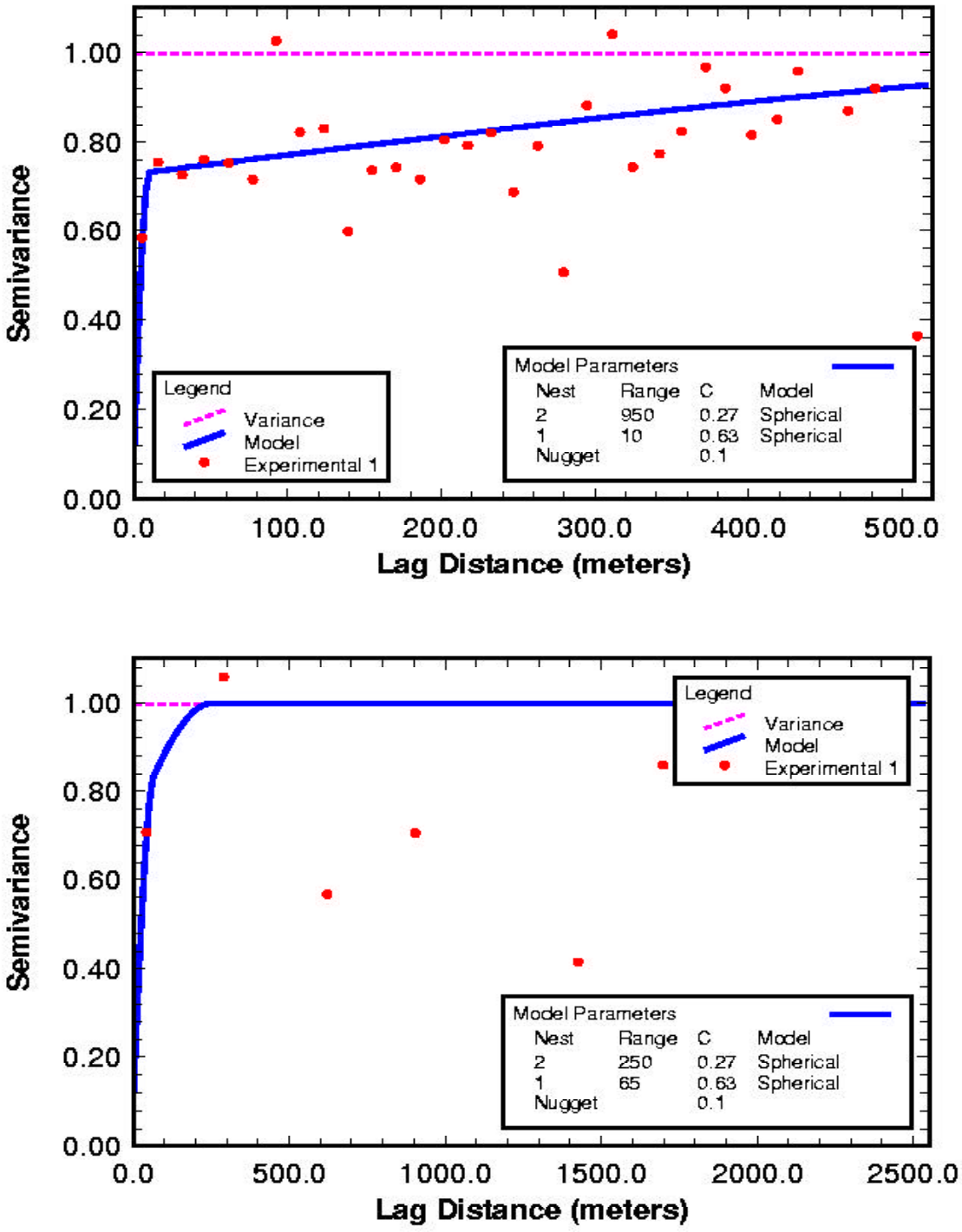

Figure 9. Normal-score variograms for the vertical (upper image) and horizontal (lower image) directions as calculated on the 17 hole data set. These variograms are shown as typical examples of all the normal-score transformed hydraulic conductivity variograms. Each experimental variogram (red dots) is fit with a double-nested spherical variogram model (blue line).

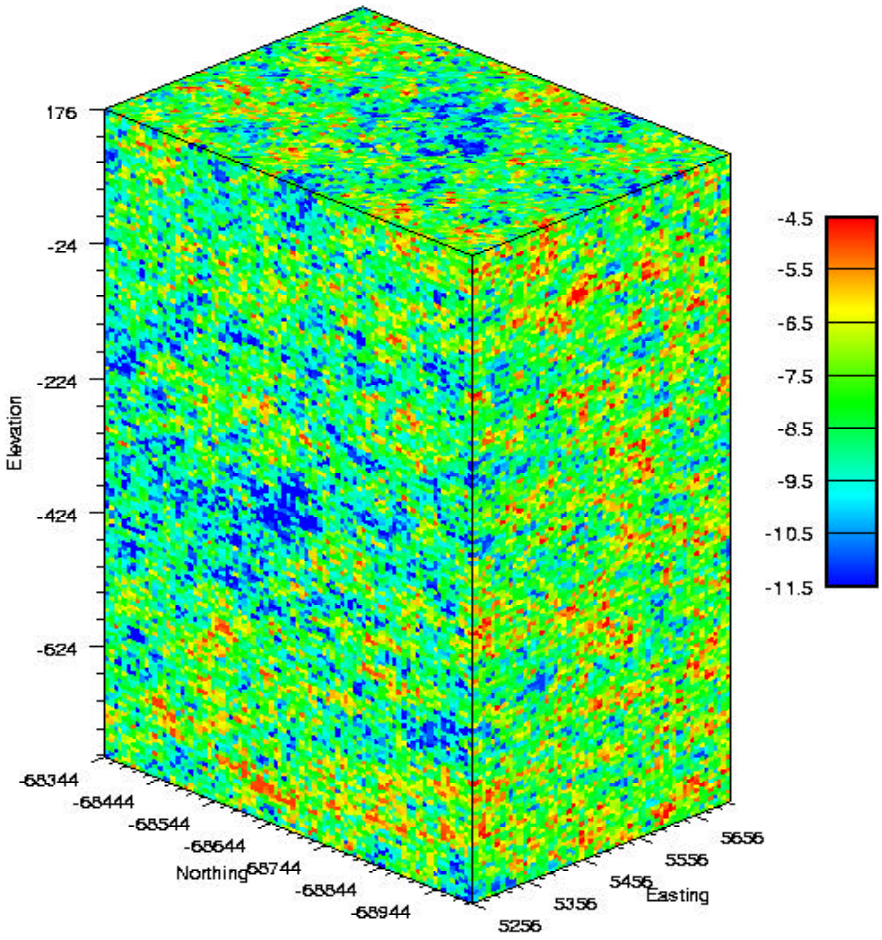

Figure 10. Example log10 hydraulic conductivity simulation showing the spatial distribution of hydraulic conductivity throughout the model domain based on the test interval hydraulic conductivity data. This example is realization number 1 created with the 19 hole data set. The color legend defines the value of log10 hydraulic conductivity in m/s.

Combined Simulations

The final step in the creation of the hydraulic conductivity simulations is the combination of the indicator facies model with the Gaussian simulation of hydraulic conductivity. For the high-hydraulic-conductivity facies (indicator value = 1), the hydraulic conductivity is obtained from the hydraulic conductivity simulation at the same grid cell location. For the low-hydraulic-conductivity facies (indicator = 0), the hydraulic conductivity value is drawn randomly from an assumed low-hydraulic-conductivity distribution (defined below). This combination is done as a simple mapping operation as outlined in pseudocode in Figure 11.

```
Loop over all grid cells
        If the indicator value = 1 (the high conductivity facies)
                Assign the simulated K value to the final model
        Else (the indicator value is a 0 representing the low K facies)
                Draw a random number from 0 to 1
                Draw the corresponding K value from the low K distribution
                Assign this K to final model
End loop
```

Figure 11. Pseudocode representation of the combination process used to create the final hydraulic conductivity values.

The low-hydraulic-conductivity distribution is chosen such that the values in the distribution lie below the lowest log10 measured hydraulic conductivity value of –12.0. For this work, the mean and standard deviation of the low-hydraulic-conductivity distribution are set to be –13.0 and 0.5 respectively. The measured hydraulic-conductivity distribution (the high hydraulic conductivity distribution) and the assumed low hydraulic conductivity distribution for the untested portion of the granite are shown as cumulative distributions in Figure 12. Only small amounts of groundwater flow are expected to occur in the low hydraulic conductivity material. For this reason, the unknown spatial correlation of hydraulic conductivity within the low hydraulic conductivity facies was not modeled. All of the low hydraulic-conductivity-values were drawn randomly, no spatial correlation, from the specified distribution. If necessary, it would be possible to determine the importance of the spatial correlation of the low hydraulic conductivity values to groundwater flow through a groundwater flow modeling sensitivity analysis, but such an analysis is beyond the scope of this study.

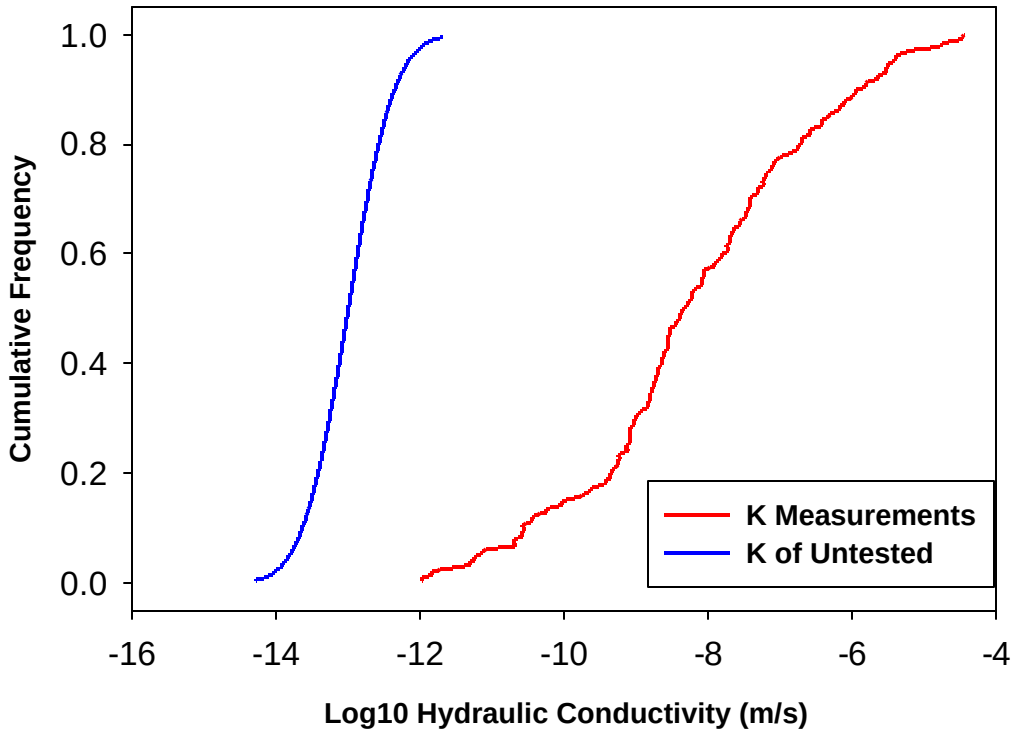

Figure 12. Cumulative distributions of hydraulic conductivity used in the creation of the final hydraulic conductivity models. The red line is the distribution of the 208 measured hydraulic conductivity values used for the high-conductivity facies and the blue line is the distribution assigned to the low-conductivity facies.

A total of 50 final hydraulic conductivity models were created from each of the six data sets. Each of these 50 models results from the combination of 50 different indicator models and 50 different hydraulic conductivity models as added together using the process outlined in Figure 11. For the entire study, 300 indicator simulations, 300 hydraulic conductivity simulations and 300 final, combined, models of hydraulic conductivity were created. These 300 final hydraulic conductivity models are analyzed in the next section to determine the effects of the additional data on the reduction of uncertainty and on the estimate of the mean hydraulic conductivity.

An example result of the combination process is shown in Figure 13. This final hydraulic conductivity image is created through a combination of the indicator and Gaussian realizations shown in Figures 7 and 10 respectively. From examination of Figure 13, it can be seen how the spatial distribution of the facies controls the final spatial distribution of the hydraulic conductivity.

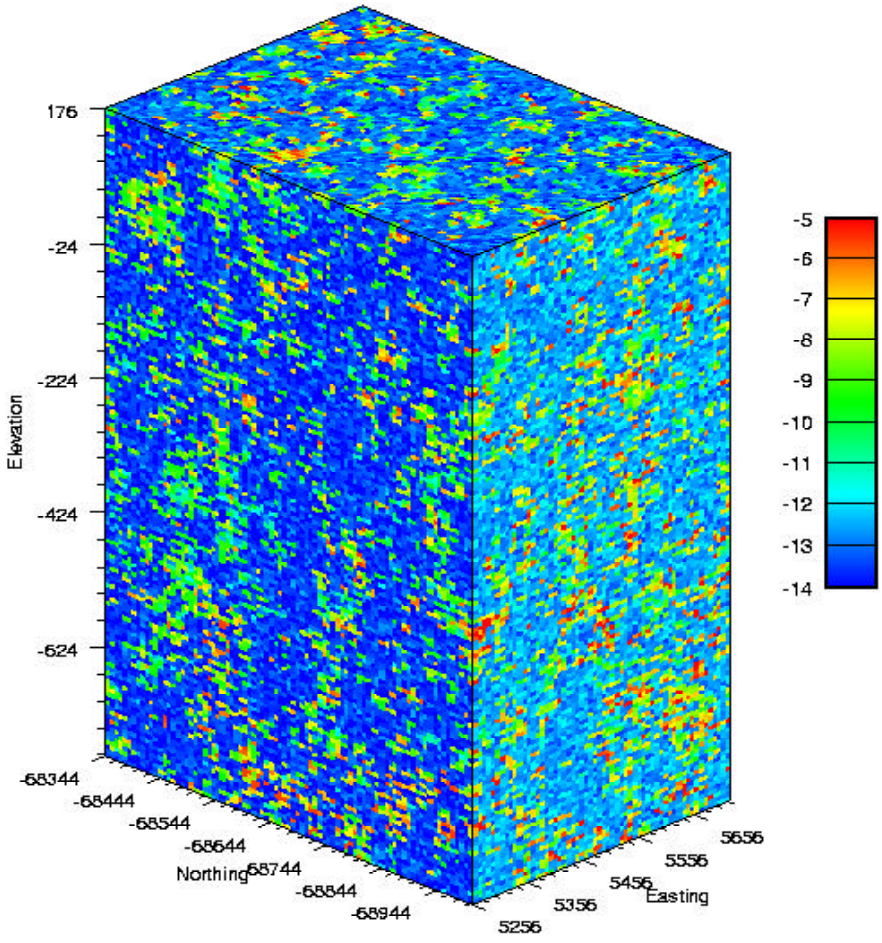

Figure 13. An example realization of the final hydraulic conductivity models. This example was created with the 19 hole data set. The facies and hydraulic conductivity realizations that were combined to create this final model are shown in Figures 7 and 10. The color legend defines the log10 hydraulic conductivity values in m/s.

Analysis of Models

The two specific goals of this study were to determine how increasing the amount of site characterization data changes the estimate of the mean hydraulic conductivity and how this increase in site characterization data changes the uncertainty surrounding the estimates of the spatial distribution of hydraulic conductivity. These two objectives are met by analyzing the 50 realizations created with each data set and comparing the results

Estimate of Mean Hydraulic Conductivity

The first comparison is the change in the mean estimate of the hydraulic conductivity as a function of the increasing amount of characterization data. This comparison is made by calculating the average (geometric mean) hydraulic conductivity for each of the 148 layers in the model and then plotting these average values as a function of the elevation. It is noted that each average is a double average: first, the average hydraulic conductivity value is calculated across all 4814 cells within the layer. Then the average value for each layer is averaged across all 50 realizations. The results of these average calculations are shown in Figure 14.

There are several interesting points to notice in Figure 14. First of all, the average values range between –12.1 and –11.4. This range is considerably lower than the average value, -8.19, of the measured hydraulic conductivities. These values are due to inclusion of the assumed values for the low-hydraulic conductivity regions. An estimate of the final average hydraulic conductivity, in log10 space, for the entire model domain can be calculated as a linear average of the two mean hydraulic conductivity values weighted by the fraction of each facies in the model:

$$K_{avg} = Fraction_{low} * -13.0 + Fraction_{high} * -8.19$$

Using the fractions of the two facies in Table 2, this calculation results in a mean hydraulic conductivity range from –12.0 to –11.8. This simple calculation provides a good single-value match to the results shown in Figure 14 for the data sets created with 14, 15, and 17 boreholes, but it does not describe the vertical variability seen in the average values constructed with the 17, 18 and 19 borehole data sets (lower image, Figure 14). These final three data sets show relatively higher hydraulic conductivity values near the top of the model and the bottom of the model. The middle elevations in these models are still within the range of –12.0 to –11.8.

The reasons for the vertical variability in the mean estimate of the hydraulic conductivity were not initially clear. However, discussions with the staff at the TGC indicated that the Tsukyoshi fault may have some influence on the vertical variation in the hydraulic conductivity. The average hydraulic conductivity profiles calculated from the realizations created with 14, 15 and 16 boreholes indicate a relatively constant average hydraulic conductivity (upper image, Figure 14). These profiles come from models created with the 14 boreholes outside the MIU site, and the same 14 boreholes with the addition of AN-1 and AN-3. None of these boreholes intercept the Tsukyoshi Fault.

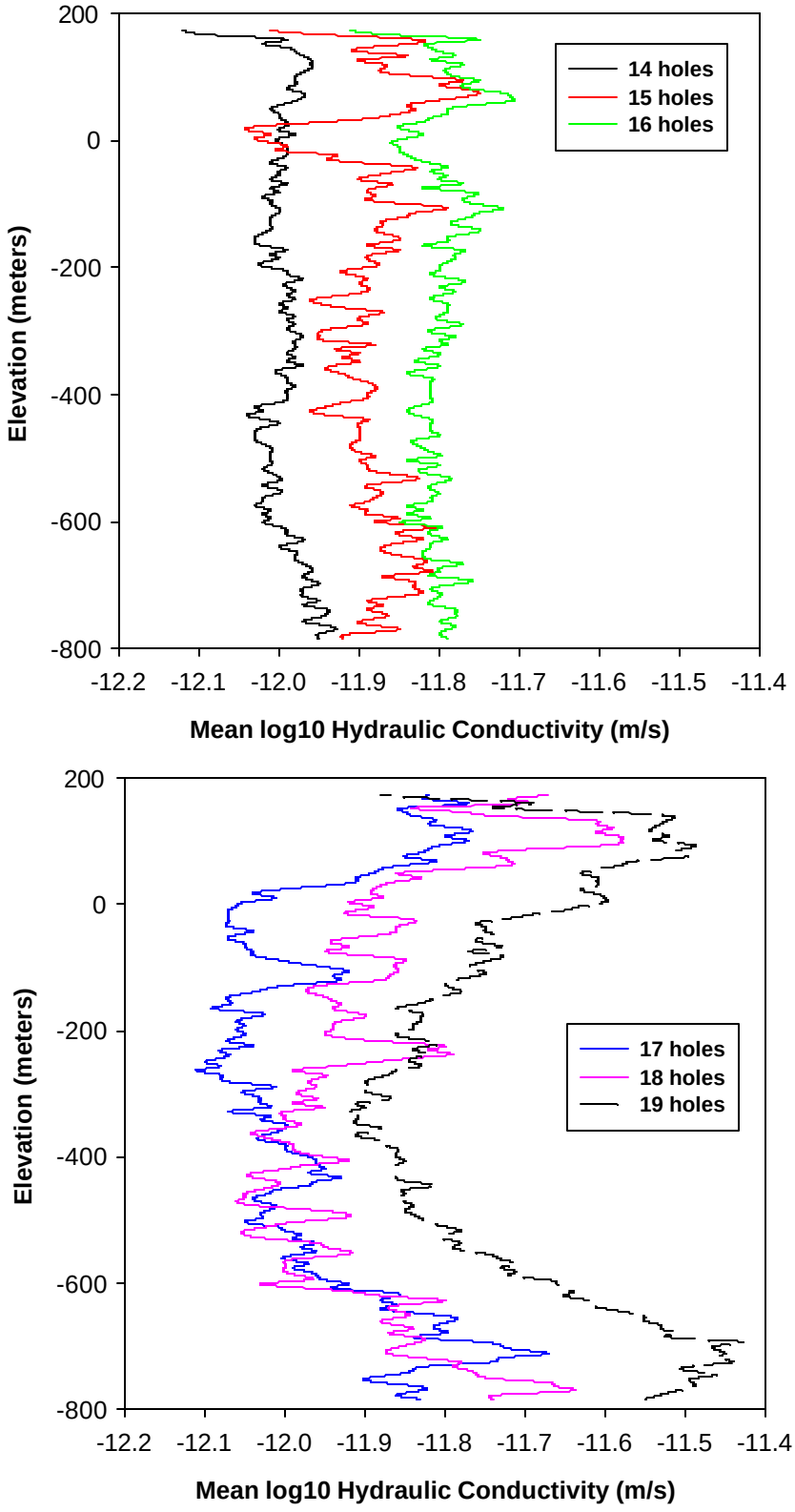

Figure 14. Mean hydraulic conductivity in each layer as a function of elevation for the size different data sets.

The addition of the MIU-1 borehole to the data set (17 hole data set, lower image, Figure 14) begins to change the shape of the average hydraulic conductivity profile. The 17 hole data set shows significantly higher hydraulic conductivity at the top and the bottom of the model domain with lower conductivity in the middle of the domain. This pattern is repeated in the 18 and 19 hole data sets (lower image, Figure 14).

The reason for the increase in permeability at the top of the model is unknown. It is speculated that this is due to some alteration or weathering of the granite near the unconformity. Another possibility is that the fracture frequency increases near the top of the granite due to stress unloading during the period of erosion prior to the deposition of the sedimentary rocks on top of the granite. However, this trend is not seen with the 14, 15 and 16 hole data sets.

The increase in hydraulic conductivity near the bottom of the model domain appears to be directly related to the presence of the Tsukyoshi Fault. Observations of fracture frequency and hydraulic conductivity in boreholes at the MIU site and in the Tono mine show that the Tsukyoshi Fault follows the fault model developed by Caine and others (1996). In this fault model, a higher permeability damage zone surrounds a lower permeability fault core. The fault core is the center of the fault where the alteration of the host rock has been the most extreme. The presence of fault gauge and chemical alteration has reduced the hydraulic conductivity of the fault core. The damage zone surrounding the fault core is the area of greatest stress relief. The damage zone contains a dense, well-connected network of fractures that enhance the permeability of this zone relative to the host rock. In the specific case of the Tsukyoshi Fault, the core is roughly 10 meters thick and the damage zone may extend out up to 35 meters on either side of the fault core.

The profiles of average hydraulic conductivity in the lower image of Figure 14 suggest that the MIU-1 borehole (17 hole data set) just intercepts the top of the damage zone creating the higher hydraulic conductivity values in the bottom of the model. Addition of the MIU-2 borehole (18 hole data set), increases the hydraulic conductivity values at the bottom of the model domain even further than did the MIU-1 borehole. This additional increase is attributed to the MIU-2 borehole intercepting even more of the Tsukyoshi Fault damage zone. Addition of the MIU-3 borehole (19 hole data set) creates a hydraulic conductivity profile with an increase in hydraulic conductivity to a maximum value of approximately –11.45 and then a decrease in hydraulic conductivity at the bottom of the model domain. It is possible that this increase and then decrease in the average hydraulic conductivity is the result of the borehole intercepting the damage zone and then the fault core.

Uncertainty Mapping

Geostatistical simulation results in one hydraulic conductivity value at every location within the model domain for each realization created. Through the creation of multiple realizations, a distribution of hydraulic conductivity values is constructed for every location in the model. The average of this distribution is the mean estimate of hydraulic conductivity at that location. The spread of this distribution, as defined by the range of

the distribution (maximum-minimum simulated value), the interquartile range (75th percentile – 25th percentile), or the standard deviation is a measure of the uncertainty in the estimate of hydraulic conductivity at each location. For this study, the standard deviation of the hydraulic conductivity distribution at each location in the model domain, as calculated over 50 realizations, is used as a measure of the uncertainty on the hydraulic conductivity value. These standard deviation values can be displayed as a two or three-dimensional map to identify areas of greatest uncertainty.

In order to examine the change in the standard deviation as a function of the increasing amount of characterization data, a vertical cross-section of the model was chosen along the X = 30 gridblock. This cross-section contains the AN-3 and MIU-1 boreholes. The values of the standard deviation of the hydraulic conductivity across this cross-section are shown in Figures 15 and 16 for the different data sets.

Figures 15 and 16 exhibit an interesting change in the uncertainly of the hydraulic conductivity as a function of increasing site characterization data. In general, uncertainty should be the lowest at the borehole locations and the greatest at locations that are the furthest away from the boreholes. However, this is not always the case in these simulations of the MIU site. Figure 15 shows the uncertainty maps for the case of the 14 boreholes outside the MIU boundary and then for the case of these same 14 boreholes and the addition of the AN-1 borehole, which does not fall along this cross-section. The uncertainty is distributed evenly in the left image of Figure 15. The right image of Figure 15 shows some layering in the standard deviation values near the top of the model. This layering is a direct result of the incorporating the AN-1 borehole into the data set.

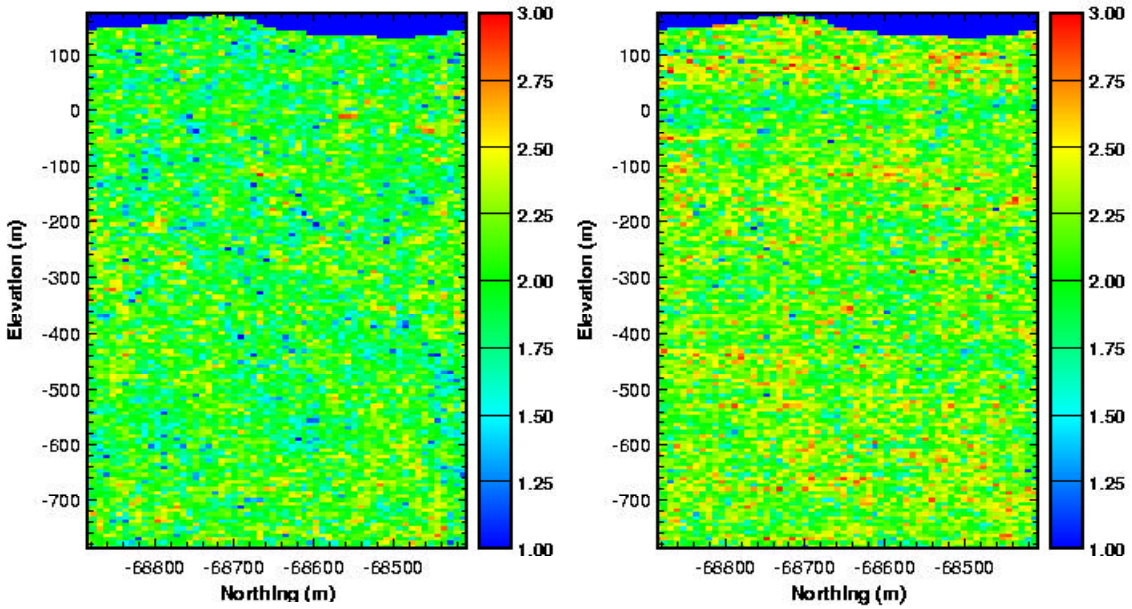

Figure 15. Cross-section of the model at the x =30 grid block showing the standard deviation in hydraulic conductivity values for the 14 (left) and 15 (right) hole data sets.

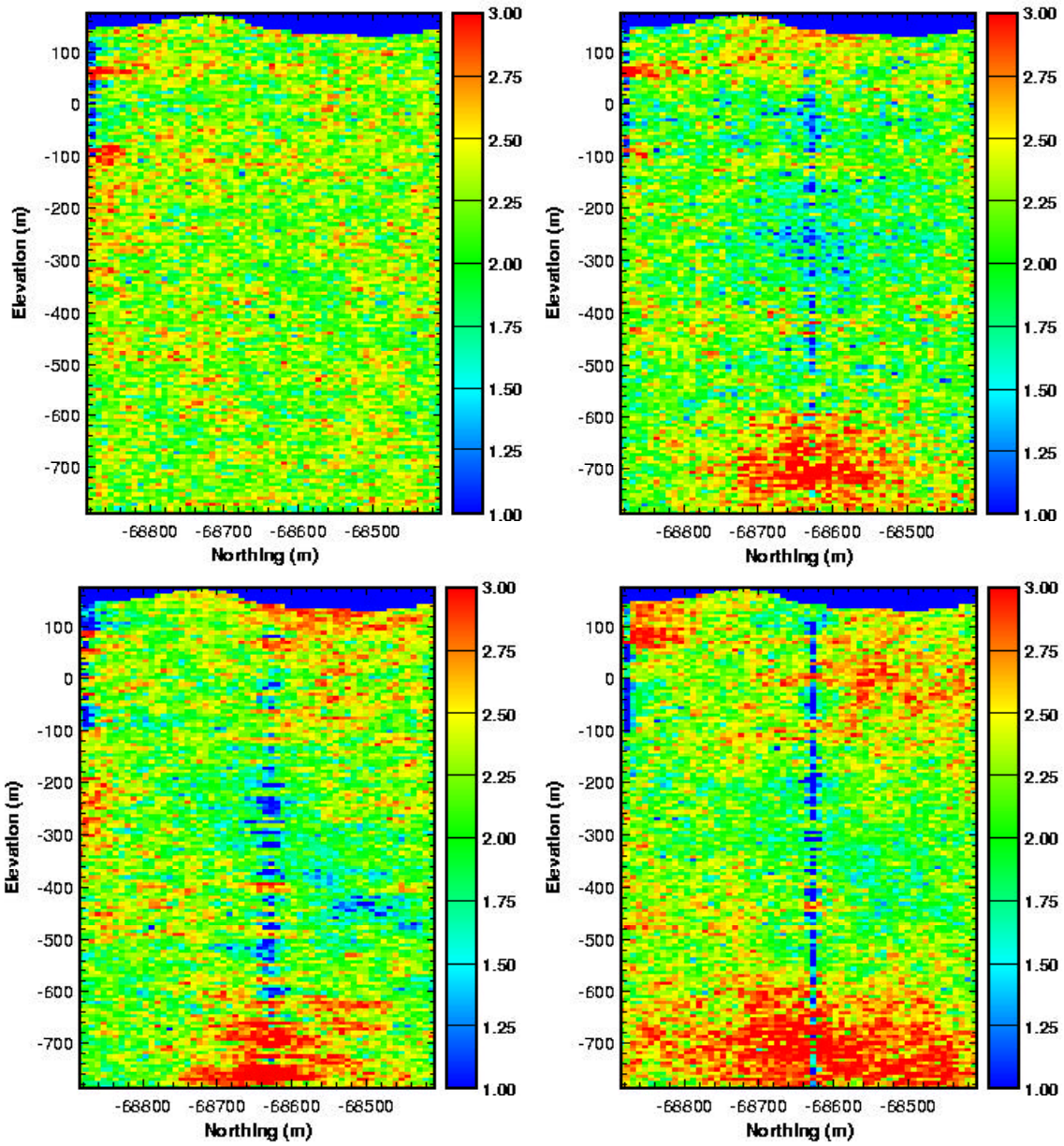

Figure 16. Cross-section of the model at the x =30 grid block showing the standard deviation in hydraulic conductivity values for the 16 (upper left), 17 (upper right), 18 (lower left) and 19 (lower right) hole data sets.

The upper left image of Figure 16 shows the uncertainty along this cross-section with the addition of the AN-3 borehole (the 16 hole data set). The location of the AN-3 borehole is obvious in the upper left side of this upper left image. In the immediate vicinity of the AN-3 borehole the uncertainty is reduced. This reduction in uncertainty remains relatively localized due to the short range of the variograms.

The upper right image of Figure 16 shows the uncertainty map after the addition of the MIU-1 borehole to the site characterization data set. The location of the MIU-1 borehole within the cross-section is obvious as the vertical line of relatively lower standard deviation values down the center of the cross-section. It is interesting to note that at the bottom of the cross-section, the uncertainty is highest in the vicinity of the MIU-1 borehole (upper right image, Figure 16). This area of higher standard deviation remains in the same location and becomes even larger with the addition of the MIU-2 and MIU-3 boreholes to the data set (lower images, Figure 16).

The increase in the uncertainty along the cross-section with the addition of the MIU boreholes is due to the conceptual model used in the geostatistics not being consistent with the hydraulic conductivity data collected near the bottom of the site domain. The addition of boreholes MIU-1, MIU-2 and MIU-3 to the data set provides increasing amounts of information on the hydraulic conductivity structure of the Tsukyoshi Fault. The geostatistical model employed in this study does not explicitly include the Tsukyoshi Fault and this leads to large uncertainty in the simulated hydraulic conductivity models in the vicinity of the fault. For example, in the damage zone surrounding the fault, the hydraulic conductivity is high and this zone should be included in the high conductivity facies defined from the locations of the packer tests. However, the geostatistical model is not defined to include all of this zone as a high hydraulic conductivity facies in every simulation. The reality of the hydraulic conductivity values being increased by the Tsukyoshi Fault is in conflict with the simpler geostatistical model that assumes statistical stationarity across the simulation domain. This conflict between the data and the model leads to increased uncertainty in these areas.

The increased uncertainty in the area of the Tsukyoshi Fault can be viewed as an important site characterization tool. The increased uncertainty is caused by uncertainty in the conceptual model for the lower portion of the simulation domain. This is not the usual cause of uncertainty in geostatistical models, but it demonstrates how a geostatistical approach can be used to check the validity of a conceptual model. If the location and the structure of the Tsukyoshi Fault was not known prior to this modeling, these areas of high uncertainty would be a warning that something outside the conceptual model was occurring and more characterization should be focussed in this area. A three-dimensional map of areas with the highest uncertainty (standard deviation) values is shown in Figure 17. The high uncertainty near the bottom of the domain is due to the Tsukyoshi Fault not being included in the conceptual basis for the geostatistical models. The high uncertainty near the top of the model is interpreted as being due to alteration or increased fracturing due to unloading during erosion, which is also not included in the conceptual model.

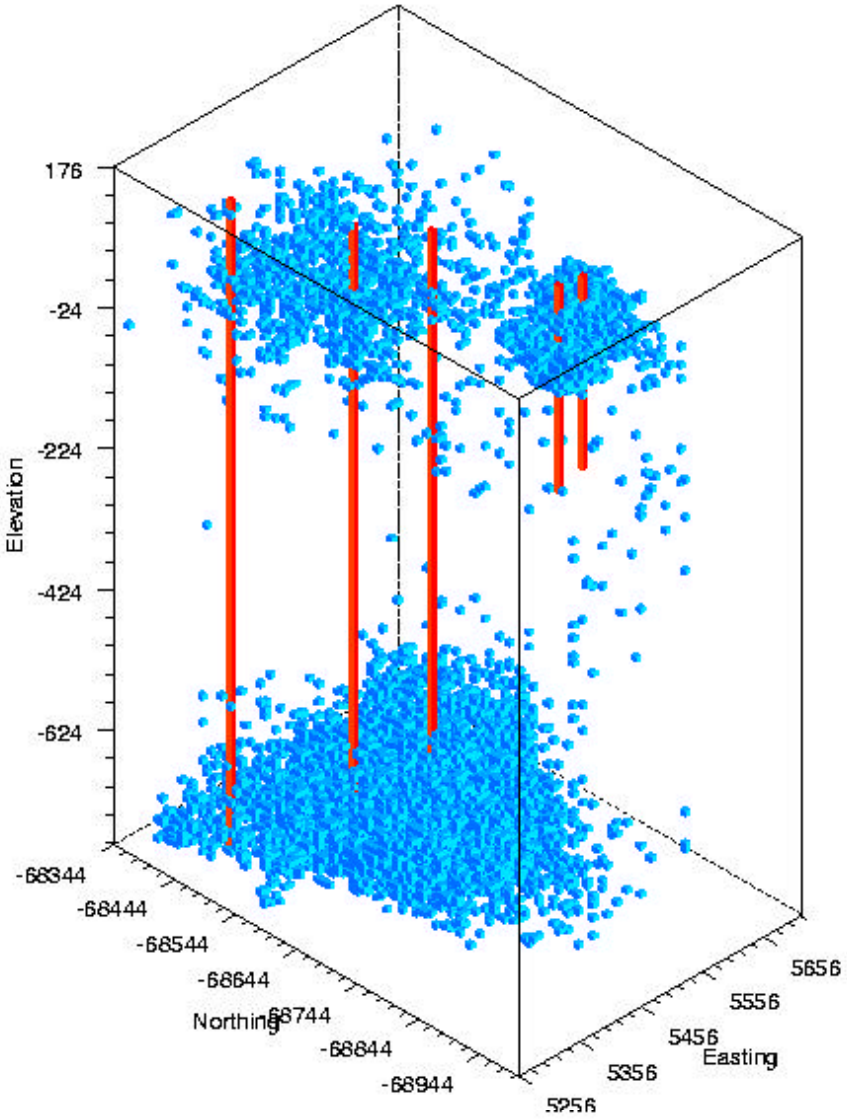

Figure 17. Three-dimensional view of the highest standard deviation values in the simulation domain. Standard deviation values greater than 3.0 are shown in blue, the borehole locations are shown in red. This map was created from the geostatistical realizations using the 19-hole data set.

Discussion

Several issues were identified during this work that require further discussion. One finding of this work is that the spatial correlation of both the test intervals and the hydraulic conductivity measured in these intervals has a very short range. Two reasons are discussed that may contribute to these short ranges. There are three additional issues that need to be covered in this work:1) Plans for construction of a second-phase of models; 2) a proposal for extracting information from the geophysical logs that can be more readily used in the site characterization process and 3) further discussion of approaches to optimizing the locations of additional boreholes.

Variogram Ranges

Figures 6 and 9 show typical variograms created with the indicator and the hydraulic conductivity data respectively. These variograms are distinctive in that the majority of the total variability occurs at relatively short lag spacings. For the variograms calculated in the horizontal directions (lower images of Figures 6 and 9) it is difficult to define the variogram at short lag spacings due to the lack of experimental variogram points between 100 and 300 meters. Additionally, the experimental variogram points that do exist at 100 and 300 meters have a small number of data pairs defining the variogram at those locations. A possible solution for this problem is discussed below in the Borehole Optimization section.

Another reason for high variability in the hydraulic conductivity variograms (Figure 9) may be that the hydraulic conductivity values are not from a single population. Examination of the cumulative distribution of the log10 hydraulic conductivity data (Figure 12) shows a distinct break in the slope of the distribution at a hydraulic conductivity value of approximately -9.3. A break in the slope of a cdf often indicates two different statistical populations combined into a single data set. In the case of the log10 hydraulic conductivity for the MIU site, it appears that the cdf follows a log-normal distribution above values of -9.3 and a uniform distribution for the values below -9.3. A uniform distribution of values is often indicative of a random process behind the generation of those values. If random values are included in the data set, then variability in the variogram will be increased.

It is interesting to compare the vertical variograms for the indicator and hydraulic conductivity data (upper images of Figures 6 and 9). The variograms are very similar in overall shape, but the experimental variogram for the hydraulic conductivity (upper image Figure 9) shows considerably more scatter than the indicator variogram. This additional variability may be due to combining two populations of hydraulic conductivity data into a single variogram, especially if one of the distributions is generated from a random process. Prior to any additional models being created, the potential of two hydraulic conductivity populations will have to be examined.

Second-Phase Models

Discussion between SNL and JNC regarding construction of a second set of geostatistical simulations of hydraulic conductivity for the MIU site has already begun. These models will explicitly incorporate the conceptual model of the Tsukyoshi Fault. The conceptual model will be implemented as subjective, or "soft", information to better constrain the geostatistical simulations. One approach being considered is to constrain the mean values of the hydraulic conductivity within the fault core and the damage zone to values based on the packer tests in those areas and agreed upon by the site characterization team at TGC. Changing the mean value of hydraulic conductivity within the fault core and damage zone will constrain the simulated values to be near the specified mean values in those regions, but this approach will still allow for spatial variability in the simulated values.

An additional observation regarding the hydraulic conductivity simulations was the apparent periodic nature of the choice of intervals within the MIU site for hydraulic conductivity testing. Evidence for this periodicity, or layering, occurs in a vertically oriented indicator variogram calculated on the test interval data for the five boreholes within the MIU boundary (Figure 18). The notable feature of this variogram is the "hole-effect" or sine wave appearance of the experimental variogram as it approaches the sill. This characteristic of the variogram is indicative of layering in the data used to create the variogram. In this case, the data are the indicator transforms of the test interval locations (1 = inside interval; 0 = outside interval) and the layering is caused by the choice of locations for test intervals.

Application of the variogram equation to the test interval data results in comparison of data values within the same type of material (2 points within the same test interval) and relatively lower variability. As the variogram lag distance increases, the two points being compared will come from inside and outside the test interval creating higher variability. As the lag distance increases again, values from one test interval are compared with values from within another test interval (both indicator values of 1) and they are similar, thus reducing variability. This pattern continues and creates the sine-wave pattern in the variogram. The peaks in the variogram can be conceptualized as resulting from differences in values arising from comparison of data in the center of one layer with data in the center of a different layer. Therefore, one-half the distance between peaks (or troughs) is the average layer thickness.

The variogram in Figure 18 presents five well-defined peaks in the data for lag distances greater than 100 meters. These peaks are consistently 60-65 meters apart. This frequency in the peaks is the direct result of layering in the data where the layers are one-half the thickness of the spacing between peaks. For the test interval spacing within the MIU site, the average layering is 30-32 meters. In discussions at TGC in January, 2000, corroborating evidence for this layering in the form of layering in the fracture frequency within several of the boreholes was presented (Goto, personal communication). In the next iteration of geostatistical models of the hydraulic conductivity, this apparent layering will be explicitly modeled. The layering in the test intervals being a direct result of layering in the fracture frequency still needs to be verified by checking the data. If there

is a strong relationship between the fracture frequency and the hydraulic conductivity, then this relationship needs to be exploited in the next round of models. One way to do this is to create geostatistical models of the fracture frequency and then use these fracture frequency models to constrain the models of hydraulic conductivity. There is considerably more fracture frequency data than there is hydraulic conductivity data, and it would be worthwhile to try and use these data.

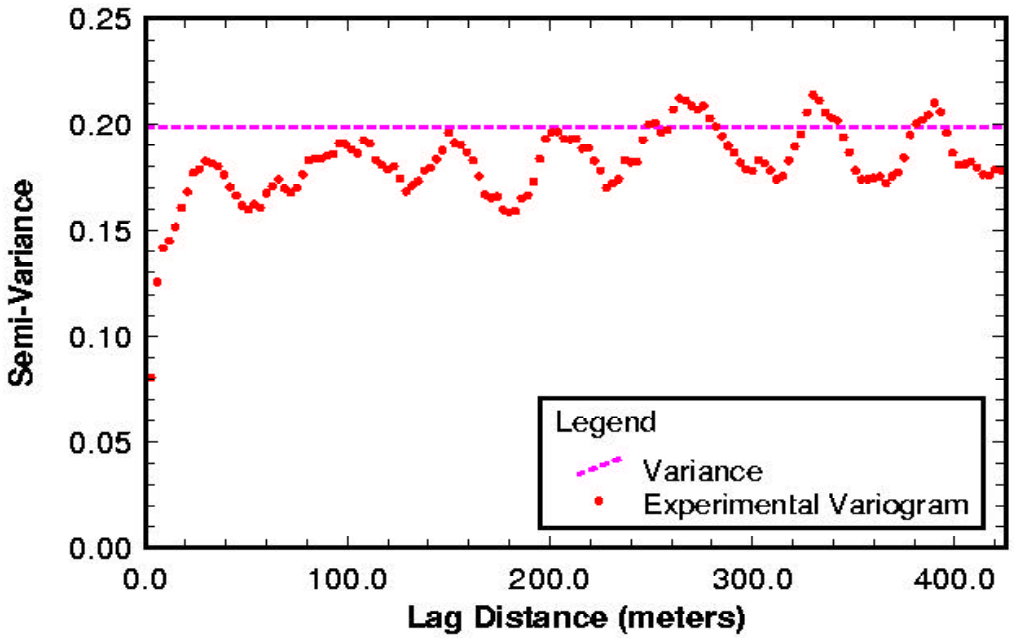

Figure 18. Indicator variogram of the test interval locations within the MIU site. This variogram was created using only data from the five boreholes within the MIU site. The prominent hole-effect indicates apparent layering in the test intervals.

Geophysical and Borehole Televiewer Logs

All of the DH boreholes and all of the boreholes within the MIU site have a full suite of geophysical logs. These logs have not yet been used in the geostatistical models of properties at the site. Additionally, it is felt that these logs could be used more efficiently in determining physical properties of the granite, especially hydraulic conductivity, and in locating hydraulic test intervals within new boreholes.

It is proposed to develop a flexible technique to extract the pertinent information from the geophysical logs and borehole televiewer logs for a number of applications. These applications may include identification of test intervals, prediction of relative hydraulic conductivity (high, medium, low), or prediction of absolute values of hydraulic conductivity. Neural networks appear to be the best way to create such a flexible tool.

Neural networks can be considered as nonlinear functions where multiple inputs to the function are used to derive one or more outputs. The function itself, the neural network, is adjusted by training the network with an existing data set. The existing data set must contain an adequate number of inputs and outputs to train the network as a robust estimator of the output for additional inputs that will be encountered outside of the training set. A number of authors have created neural networks to operate as facies classification functions in sedimentary environments (Baldwin, et al., 1989; Dowla and Rogers, 1995; Rogers et al., 1992). More recently, other authors have used neural networks to actually predict values of permeability in sedimentary formations (Mohaghegh et al., 1996; Wong and Shibli, 1998). It should be possible to extend these ideas to the problem facies classification and hydraulic conductivity estimation in fractured rocks at the MIU site. The previous works in sedimentary environments have used only geophysical logs as input to the neural network. For the MIU site, it will also be possible to extract information regarding fracture frequency from the borehole televiewer logs and information on the alteration of the rock from the geologists description of the core samples.

Borehole Optimization

This report has documented how a geologic property, such as hydraulic conductivity, can be simulated using geostatistical techniques. A direct result of the simulation process is that the uncertainty in the property can be readily determined from the distribution of simulated values at every location. This uncertainty can be used to guide further site characterization. Areas with the highest uncertainty should be targeted for the next round of characterization. The characterization should continue until the uncertainty is reduced to an acceptable level. Results in this report indicate that the uncertainty in the hydraulic conductivity values is highest in the vicinity of the Tsukyoshi Fault.

The work done in this report only considers uncertainty in the hydraulic conductivity at each location. In many cases, it is the uncertainty with respect to a performance-based question that is of most importance. As an example, the hydraulic conductivity realizations created in this report could be used as input to a groundwater flow and particle-tracking model. Running this model on each of the 300 hydraulic conductivity realizations would produce a distribution of groundwater travel times from a starting location in the center of the site to the site boundary. The goal of the characterization program has now changed from reducing uncertainty in the values of hydraulic conductivity at each location, to reducing the uncertainty in the groundwater travel time. The optimal location to drill the next borehole will not be the same for the two different characterization goals. This example points out the concept of performance-based, or question-specific, site characterization.

The results of any uncertainty estimates are only as good as the underlying model used to create those estimates. In the case of a geostatistical approach to site characterization, the variogram models must represent the spatial variability at the site. In many cases, there are not enough data to adequatly define the variogram. This is the case for the horizontal

variograms at the MIU site shown in this report. One solution to this problem of little data at short-range spacings is to arrange the site characterization boreholes in more efficient spiral pattern. A spiral pattern for the boreholes would provide more data points in the experimental variograms at short ranges. This approach has been suggested previously by Saegusa and others (1999), but has not been applied at the MIU site.

Summary

This report presents a geostatistical approach to modeling the spatial distribution of hydraulic conductivity at the MIU site. This approach is a stochastic approach that allows for the determination of the uncertainty associated with the estimated hydraulic conductivity at every location. A two-stage approach was used in this work. In the first step, the distribution of high and low conductivity regions, or facies, was simulated using indicator geostatistical simulation. These indicator simulations are conditioned to the locations of the test intervals and define the regions of the three-dimensional domain that contain "testable" rock. The second stage involved simulation of the hydraulic conductivity values conditioned to the values of hydraulic conductivity measured in the test intervals. These two types of simulations were then combined by retaining the simulated hydraulic conductivity values in the areas of the high conductivity rock as defined by the indicator simulation. In the areas defined as low conductivity rock, hydraulic conductivity values were drawn from a different hydraulic conductivity distribution with a lower mean and smaller variance than that of the measured distribution.

The effects of the data set on the final mean and uncertainty of the hydraulic conductivity models were examined by creating the models with six different data sets. For each of the six data sets, a total of 50 realizations of the hydraulic conductivity field were created. For each data set, the mean hydraulic conductivity of each layer in the model was calculated and graphed as a function of the elevation. These graphs show that with increasing amounts of data, the mean hydraulic conductivity of the model increases. As the data from the three MIU holes are added to the data set, the shape of the mean hydraulic conductivity profile changes to show increased hydraulic conductivity at the top and the bottom of the model domain. This change in the shape of the profile is attributed to increased fracturing near the top of the model due to alteration of the granite near the unconformity and/or unloading of lithostatic stress by erosion prior to deposition of the sedimentary units. The increase in mean hydraulic conductivity near the bottom of the model is attributed to the MIU boreholes intercepting the damage zone around the Tsukyoshi Fault. Both of these features that cause increases in hydraulic conductivity at the top and bottom of the model were not explicitly included in geostatistical models. This lack of explicit inclusion in the geostatistical model leads to large values of uncertainty, large standard deviations, in the hydraulic conductivity estimates at these locations. Results of this modeling indicate that the highest uncertainty is at the top and bottom of the model domain and based on these results, hydraulic conductivity testing should be focussed in these areas. Furthermore, these results indicate that more characterization is necessary to determine a stable estimate of the mean hydraulic conductivity within the MIU site.

Several ideas have been proposed for additional work involving geostatistical models of the hydraulic conductivity at the MIU site. Indicator variogram analysis of the test interval locations within the MIU site indicates that these intervals were chosen in a layered pattern where the average thickness of each layer is approximately 30 meters. Analysis of several of these boreholes by JNC staff indicates that fracture frequency may have a similar layered pattern and may be the cause of the layering in the test interval locations. This layered pattern will be incorporated into future models of the hydraulic conductivity within the MIU site. It is desirable to make better use of the geophysical and borehole televiewer logs for conditioning the geostatistical simulations. One idea is to extract more information from these logs by constructing a neural network to estimate values of hydraulic conductivity based on the geophysical and borehole televiewer logs.

References

Baldwin, J.L., D.N. Otte and C.L. Wheatley, Computer Emulation of Human Mental Processes: Application of Neural Network Simulators to Problems in Well Log Interpretation, *SPE 19619*, Presented at the 64th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, San Antonio, Texas, October 8-11,

Caine, J.S., J.P. Evans and C.B. Forster, 1996, Fault Zone Architecture and Permeability Structure, *Geology*, Vol. 24, pp. 1025-1028.

Deutsch, C.V. and A.G. Journel, 1998, *GSLIB: Geostatistical Software Library and User's Guide, 2nd edition*, Oxford University Press, New York, 361 pp.

Dowla, F.U. and L.L. Rogers, 1996, *Solving Problems in Environmental Engineering and Geosciences with Artificial Neural Networks*, MIT Press, Cambridge, Massachusetts, 310 pp.

Goovaerts, P., 1997, *Geostatistics for Natural Resources Evaluation*. Oxford Univ. Press, New-York, 483 pp.

Isaaks, E.H. and R.M. Srivastava, 1989, *An Introduction to Applied Geostatistics*, Oxford University Press, 561 pp.

Journel, A.G. and J-Ch. Huijbregts, 1978, *Mining Geostatistcs*, Academic Press, New York, 600 pp.

Journel, A.G. and F. Alabert, 1989, Non-Gaussian Data Expansion in the Earth Sciences, *Terra Nova*, Vol. 1, pp. 123-134.

Lavenue, A.M. and B.S. RamaRao, 1992, A Modeling Approach to Address Spatial Variability within the Culebra Dolomite Transmissivity Field, *SAND92-7306*, Sandia National Laboratories, Albuquerque, New Mexico, 116 pp.

Mohaghegh, S., S. Ameri and R. Arefi, 1996, Virtual Measurement of Heterogeneous Formation Permeability Using Geophysical Well Log Responses, *The Log Analyst*, March-April, pp. 32-39.

Rautman, C.A. and S.A. McKenna, 1997, Three-Dimensional Hydrological and Thermal Property Models of Yucca Mountain, Nevada, *SAND97-1730*, Sandia National Laboratories, Albuquerque, New Mexico, 322 pp.

Rogers, S.J., J.H. Fang, C.L. Karr and D.A. Stanley, 1992, Determination of Lithology from Well Logs Using a Neural Network, *AAPG Bulletin*, Vol. 76, No. 5, pp. 731-739.

Saegusa**,** H., S. Takeuchi, M.J. White, and H. Inaba, 1999, Modeling of Heterogeneous Hydraulic Conductivity Field Using Fractal Theory, in: Proceedings of the 7th International Conference Proceedings on Radioactive Waste Management and Environmental Remediation, ICEM '99, Nagoya, Japan, Sept. 26-30, 7 pp.

Wingle, W.L., E.P. Poeter and S.A. McKenna, 1999, UNCERT: Geostatistics, uncertainty analysis and visualization software applied to groundwater flow and contaminant transport modeling, *Computers and Geosciences*, Vol. 25, pp. 365-376.

Wong, P.M. and S.A.R. Shibli, 1998, Use of Interpolation Neural Networks for Permeability Estimation from Well Logs, *The Log Analyst*, November-December, pp. 18-26.

Yusa, Y., K. Ishimaru, K. Ota and K. Umeda, 1992, Geological and Geochemical Indicators of Paleohydrology in Tono Uranium Deposits, Japan, in: *Paleohydrogeological Methods and their Applications*, Proceedings of an NEA Workshop, Paris France, Nov. 9-12, pp. 117-146.