R-99-03

Long-range spatial dependence in fractured rock

Empirical evidence and implications for tracer transport

Scott Painter

Commonwealth Scientific and
Industrial Research Organization
Melbourne, Australia

February 1999

Svensk Kärnbränslehantering AB
Swedish Nuclear Fuel
and Waste Management Co
Box 5864
SE-102 40 Stockholm Sweden
Tel 08-459 84 00
+46 8 459 84 00
Fax 08-661 57 19
+46 8 661 57 19

SKB

R-99-03

Long-range spatial dependence in fractured rock

Empirical evidence and implications for tracer transport

Scott Painter

Commonwealth Scientific and
Industrial Research Organization
Melbourne, Australia

February 1999

Svensk Kärnbränslehantering AB
Swedish Nuclear Fuel
and Waste Management Co
Box 5864
SE-102 40 Stockholm Sweden
Tel 08-459 84 00
+46 8 459 84 00
Fax 08-661 57 19
+46 8 661 57 19

SKB

ISSN 1402-3091
SKB Rapport R-99-03

Long-range spatial dependence in fractured rock

Empirical evidence and implications for tracer transport

Scott Painter

Commonwealth Scientific and
Industrial Research Organization
Melbourne, Australia

February 1999

Abstract

Nonclassical stochastic continuum models incorporating long-range spatial dependence are evaluated as models for fractured crystalline rock. Open fractures and fracture zones are not modeled explicitly in this approach. The fracture zones and intact rock are modeled as a single stochastic continuum. The large contrasts between the fracture zones and unfractured rock are accounted for by making use of random field models specifically designed for highly variable systems. Hydraulic conductivity data derived from packer tests in the vicinity of the Äspö Hard Rock Laboratory form the basis for the evaluation. The Äspö $\log K$ data were found to be consistent with a fractal scaling model based on bounded fractional Levy motion (bfLm), a model that has been used previously to model highly variable sedimentary formations. However, the data are not sufficient to choose between this model, a fractional Brownian motion model for the normal-score transform of $\log K$, and a conventional geostatistical model. Stochastic simulations conditioned by the Äspö data coupled with flow and tracer transport calculations demonstrate that the models with long-range dependence predict earlier arrival times for contaminants. This demonstrates the need to evaluate this class of models when assessing the performance of proposed waste repositories. The relationship between intermediate-scale and large-scale transport properties in media with long-range dependence is also addressed. A new Monte Carlo method for stochastic upscaling of intermediate-scale field data is proposed.

Contents

1 Introduction

The large contrast in hydraulic conductivity between fractures and intact rock is one of the key challenges in modeling fractured geological media. An equally important issue is the long-range spatial correlation in hydraulic conductivity implied by spatially extensive fracture zones. While these issues present no problem for deterministic continuum-level descriptions, they do present problems for the stochastic continuum models that form an important part of the strategy for dealing systematically with unobserved geological heterogeneities. Specifically, large contrasts in conductivity and long-range spatial correlations are inconsistent with the classical random field models underlying many stochastic continuum descriptions. These limitations have created significant interest in Monte Carlo methods based on discrete fractures as opposed to more traditional continuum-level descriptions. While these methods are clearly useful for many applications, they can be computationally intensive and are difficult to extend to treat mass exchange between the fractures and rock matrix.

However, models for complex subsurface heterogeneities have advanced considerably beyond the classical models of subsurface hydrology and are now able to treat long-range dependence and high spatial variability more effectively. Given the well known advantages of stochastic continuum models over discrete fracture models (most notably their smaller computational requirements and the ease at which they can be extended to treat mass exchange with the rock matrix) there is considerable motivation for revisiting stochastic continuum models of fractured rock while making use of recent theoretical models incorporating long-range dependence.

In this report, stochastic continuum representations incorporating long-range spatial dependence are evaluated as models for fractured crystalline rock. The discussion is based on hydraulic conductivity data from the area of Äspö Island in southern Sweden, the site of the SKB Hard Rock Laboratory. The data are first analyzed for fractal structure (long-range dependence). Stochastic simulations based on two fractal models and conditional on the hydraulic conductivity data are then combined with flow and particle tracking calculations to explore the consequences for non-reacting tracer transport. The two fractal models employed include fractional Brownian motion (fBm) (Mandelbrot & Ness, 1968; Feder, 1988) and fractional Levy motion (fLm) (Mandelbrot & Ness, 1968; Taqqu, 1987; Kumar & Foufoula-Georgiou, 1993), a generalization of fBm which has been used successfully in the past to model subsurface formations with high spatial variability (Painter & Paterson, 1994; Painter, 1995; Painter, 1996). For comparison purposes, a conventional multiGaussian model with short-range correlation structure is

also considered.

The issues addressed here are related to geostatistical *model selection* and its impact on uncertainty in performance assessment of high-level waste repositories. Geostatistical methods such as conditional stochastic simulation are often advocated as general approaches for quantifying uncertainties created by incomplete sampling of subsurface heterogeneities. However, the uncertainty calculated in the geostatistical approach is not the true uncertainty, but the uncertainty internal to the particular choice of random field model. More difficult to deal with is the *modelization uncertainty*, the uncertainty associated with the choice of geostatistical model. Because data from potential repository sites are rarely sufficient to distinguish between competing geostatistical models, it is important to have a broad understanding of how the choice of geostatistical models affects the predicted performance of the repository geosphere. Although similar issues have been addressed in a recent publication (Tsang *et al.*, 1996), the focus there was on a single non-parametric geostatistical methodology (indicator simulation (Journel & Huijbregts, 1978)). Here the focus is on a class of long-range dependent (fractal scaling) models that have been used with considerable success in other subsurface applications (Molz *et al.*, 1997). The aim is not to advocate one methodology over another but to illustrate how the choice of geostatistical model can have a profound effect on the predicted geosphere performance.

2 Stochastic models with long-range spatial dependence

The standard approach in stochastic subsurface hydrology is to model the log conductivity field ($Y = \log K$, where K is the hydraulic conductivity) as a multivariate Normal (multiGaussian) random field with exponentially decaying correlation. There is, however, increasing evidence suggesting that several of the assumptions inherent in this choice of random field may be violated by subsurface data. It has been suggested by direct analysis of subsurface data (Hewett, 1986; Todoeschuck *et al.*, 1990; Molz & Boman, 1993; Molz & Boman, 1995; Painter & Paterson, 1994; Painter *et al.*, 1995; Painter, 1996; Crane & Tubman, 1995) and indirectly by observed scale-dependence in dispersivities (Neuman, 1990) that the assumption of a finite correlation range (integral scale) may not be met. This implies that fluctuations in subsurface properties exist over a wide spatial range. Further, the assumption of stationarity and Gaussianity in the Y field have also been questioned (Painter, 1995), suggesting a higher degree of spatial variability.

Random field models with long-range spatial dependence provide alternatives that may better mimic the complex non-classical fluctuations observed in subsurface datasets.

2.1 Fractal noises

The fractal noises are the simplest class of long-range dependent (LRD) models. The signature for this class is a power-law decay in the correlation function

$$\langle Y(u)Y(u+l)\rangle \propto |l|^{-\gamma} \tag{1}$$

for large lag $|l|$ (see e.g. (Feder, 1988)). Here u is the spatial coordinate and $\langle\cdot\rangle$ denotes an ensemble average. The most famous of these models, fractional Gaussian noise (fGn), results from a particular choice of the correlation near the origin. Although the fractal noises have long-range correlation, the degree of spatial variability is not much different from standard model. This class of LRD models in not considered here for this reason. Instead, the focus is on LRD models with more spatial variability.

2.2 Fractional Brownian motion

Fractional Brownian motion (fBm) is perhaps the best known LRD model. The signature for this model is a power-law variogram

$$\langle [Y(u+l) - Y(u)]^2\rangle \propto |l|^{2H} \tag{2}$$

or, equivalently a power-law Fourier spectrum.

FBm is a non-stationary model. In the absence of conditioning data or other constraints it makes no prediction about the univariate statistics. Instead the focus is on the stationary increments, which in the case of fBm have a Gaussian distribution

$$Pr[Y(u+l) - Y(u) < \zeta] = G(\zeta\ ;\ 0, \sigma_0^2|l|^{2H}) \tag{3}$$

here $G(\cdot\ ;\ \mu, \sigma^2)$ is the univariate Gaussian cumulative distribution function (CDF) with mean μ and variance σ^2. The two parameters appearing in Eq. 3 are σ_0^2, which is the variogram at unit lag, and the Hurst parameter $H \in (0,1)$, which quantifies the degree of spatial dependence in the random field. The situation $H = 1/2$ corresponds to classical Brownian motion, while $H > 1/2$ implies positive correlation in increments referred to as persistence. For subsurface data H is typically less than $1/2$ implying antipersistence (positive increments tend to be followed, on average, by negative increments). Example traces of two fBm processes are shown in Figure 1.

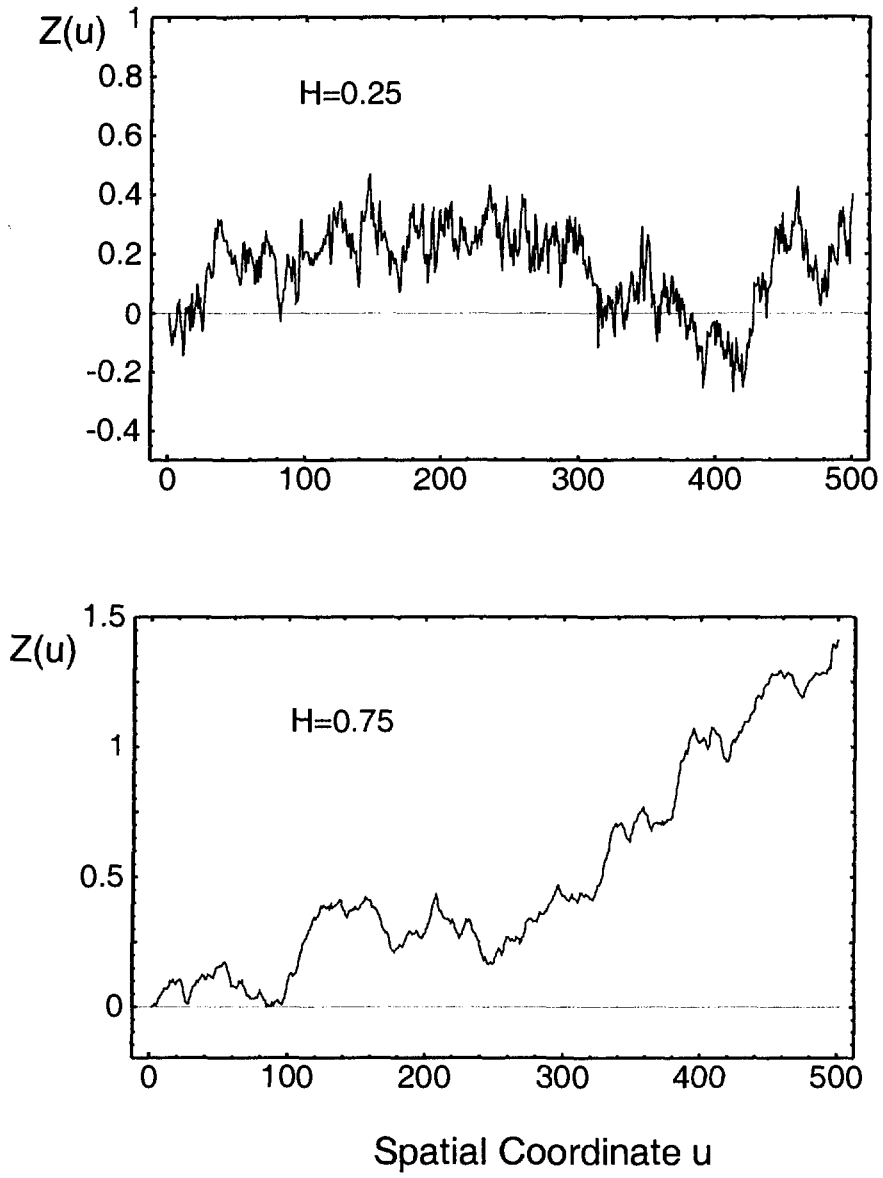

Figure 1: Realizations of a fractional Brownian motion (fBm) process for two values of the Hurst parameter H. These were generated using the LevySim fractal random field generator. The fBm process is the simplest model with long-range dependence and has been used extensively to model geophysical time series and spatial fields.

Although fBm is a considerable improvement over the standard model for many applications, the fact that it is Gaussian-based means that it also has a limited degree of spatial variability. This can be an important limitation for subsurface formations that have large contrasts in hydraulic conductivity. In this case it is necessary to turn to generalizations with more flexibility to describe large contrasts in hydraulic conductivity.

2.3 Fractional Levy motion

Fractional Levy motion (fLm) is a class of models that generalize fBm. This class of random field models is based on the Levy-stable family of probability distributions (Levy, 1937; Zolotarev, 1986; Feller, 1971). These distributions have a central role in mathematical statistics similar to the Gaussian distribution in that they result from sums of large numbers of independent random variables. When the individual components in the sum have a distribution with finite variance the distribution of the sum tends to a Gaussian distribution according to the central limit theorem. When the individual components in the sum have a power-law distribution, the variance is not defined and the sum tends to a Levy-stable distribution (generalized central limit theorem). Levy-stable distributions have slowly decaying power-law tails, which makes them useful for modeling systems with a high degree of spatial variability. Densities are shown in Figure 2. Levy-stable distributions are parameterized by the width parameter C the Levy index α and the location parameter d. The location parameter simply centers the distribution, similar to the mean in the Gaussian distribution, C measures the spread in the distribution about its center, similar to the standard deviation in the Gaussian distribution. The Levy index quantifies the decay in the tails of the distribution. It runs between 0 (exclusive) and 2 (inclusive) with the singular case $\alpha = 2$ corresponding to the Gaussian distribution.

The most easily observable feature of fLm is a Levy-stable distribution for the incremental values

$$Pr[Y(u+l) - Y(u) < \zeta] = \psi(\zeta\ ;\ 0, C_0|l|^H, \alpha) \tag{4}$$

Here $\psi(\cdot\ ;\ d, C, \alpha)$ is the univariate Levy-stable CDF with location parameter d, width parameter C, and Levy index α. The situation $\alpha = 2$ corresponds to the Gaussian distribution with variance $C_0/2$. In this situation the model reduces to the fBm model,

$$\psi(\zeta\ ;\ 0, C_0|l|^H, 2) = G(\zeta\ ;\ 0, \frac{C_0}{2}|l|^{2H}). \tag{5}$$

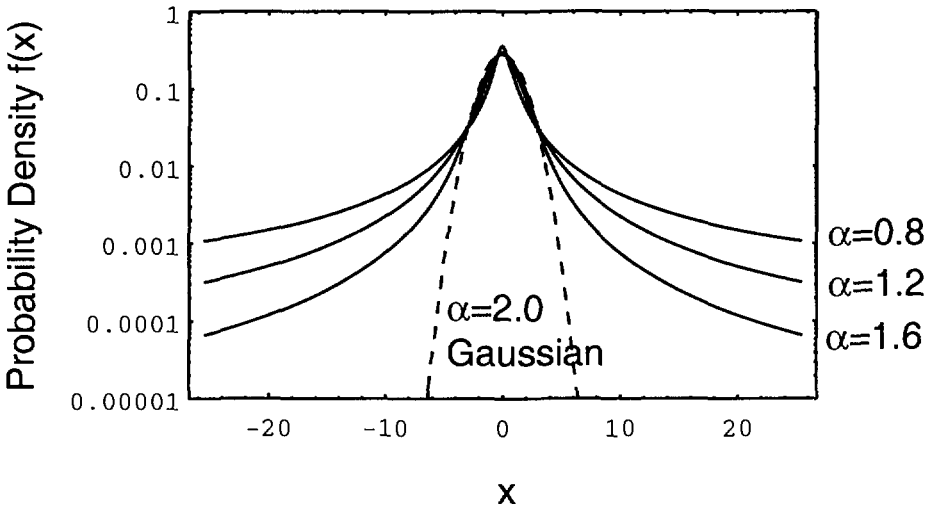

Figure 2: Probability densities for four members of the Levy-stable family of probability distributions. They have a central role in mathematical statistics and have been observed to model accurately subsurface data. The Levy-stable distributions are parameterized by the Levy index $\alpha \in (0, 2]$, with $\alpha = 2$ corresponding to the Gaussian distribution. The other members in this family have slowly decaying tails which makes them useful for modeling systems with enhanced variability.

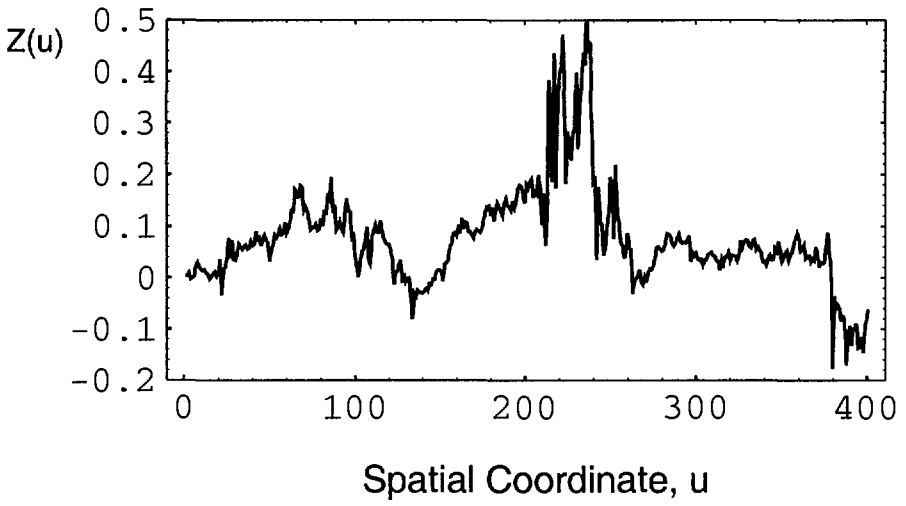

Figure 3: Example trace produced by the LevySim random field generator using the bounded fractional Levy motion (bfLm) model. This model is similar to fBm but is based on the non-Gaussian Levy-stable distributions and has enhanced spatial variability.

In the particular construction of fLm used here, the Hurst parameter H has a similar meaning as in the fBm situation.

As a final modification to the fLm model, explicit bounds (b_u and b_l) are imposed on the simulated variables. This is necessary as an unbounded fLm model would imply non-zero probability density for non-physical values. This bounded fLm (bfLm) model better reproduces features observed in hydraulic conductivity data. For example, the bounds force a return to Gaussianity at large separation distances, a feature which Liu and Molz (1997) have observed in the MADE dataset (Boggs *et al.*, 1992). They also eliminate the diverging theoretical moments characteristic of Levy models. In the bfLm model, the distribution of incremental values depends on the lag separation, which is consistent with multifractal rather than monofractal properties. By multifractal, we mean that the media is characterized by a generalized variogram of the form

$$\langle |Y(u+l) - Y(u)|^q \rangle \propto |l|^{p(q)} \tag{6}$$

with multifractal spectrum $p(q)$ that deviates from the monofractal result $p(q) = qH$.

An example trace of the bfLm model is shown in Figure 3. This trace was generated using the sequential simulation algorithm described by Painter (1998).

3 LRD stochastic continuum representation of the Äspö region

3.1 Hydraulic conductivity data from the Äspö site

The data forming the basis of this study are 1300 measurements of hydraulic conductivity inferred from injection tests in 3-m packered-off sections (Rhen *et al.*, 1997). These measurements come from 8 boreholes in the vicinity of Äspö island ranging in depth from 600 m to 900 m. A top view and side view of the trajectories for the 8 boreholes are shown in Figure 4. Four of these boreholes have dip angles near 60 degrees while the other four are close to vertical.

Although the packer intervals are roughly the same for these measurements, the scale of support may vary somewhat because the range of influence for each packer test is related to the local hydraulic conductivity. This possible variation in the scale of support is neglected here. Hydraulic anisotropy is also neglected, although it is recognized that fractured rock often has significantly different conductivity in the horizontal and vertical directions. In situations where sufficient data on vertical and horizontal conductivites are available, both could be included in analyses similar to those described below. This would involve treating the vertical and horizontal conductivites as distinct but highly correlated random fields.

A histogram of $\log K$ is shown in Figure 5, where K is hydraulic conductivity in m/s and log refers to the base 10 logarithm. The K distribution spans more than 8 orders of magnitude. The $\log K$ distribution has a mean of -9.08 and a standard deviation of 2.99. Further, it is highly skewed to the right, suggesting that the K distribution is not well approximated by a log-normal distribution.

One geostatistical simulation recipe for dealing with a univariate distribution that is not Gaussian or log-normal is to simply transform the data so that the resulting empirical distribution is Gaussian. A sequential Gaussian simulation (SGS) (Journel & Huijbregts, 1978) is then be employed and the resulting realization back-transformed to obtain a realization of the conductivity field. While this undoubtedly improves the performance of the SGS method, the normal-score transform does not correct the multipoint joint distribution, which is presumed to be multivariate Gaussian in the SGS method. The consequence of this is that aburbt changes in conductivity fields over short distances, an undeniable feature of many geological data sets, is not necessarily preserved in the geostatistical simulations. Abrupt changes in hydraulic conductivity such as those associated with the boundaries of

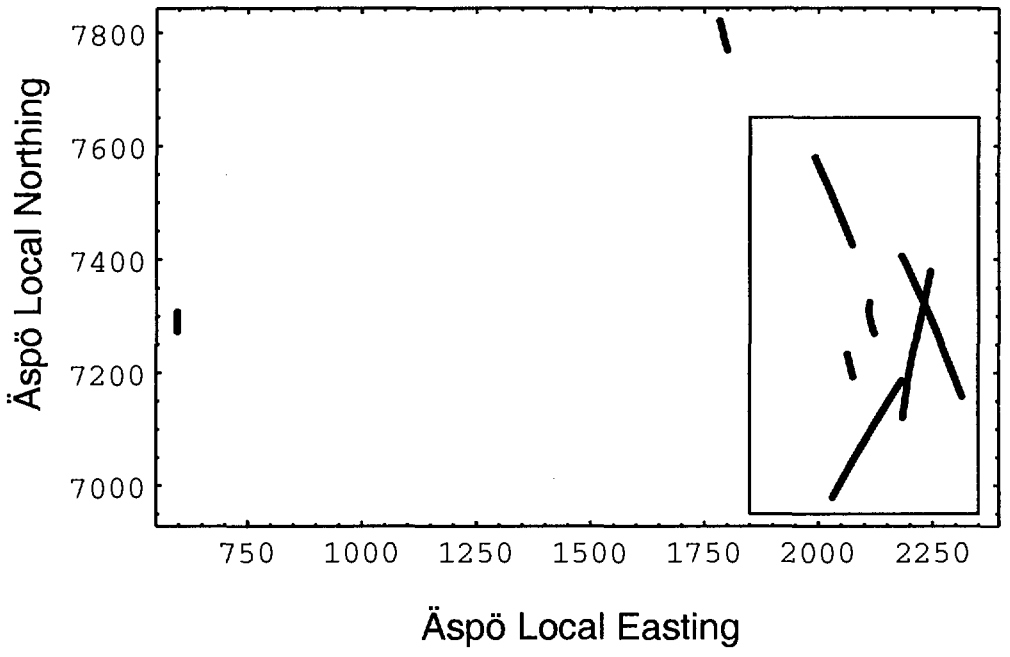

Figure 4: Locations of the boreholes that provide the hydraulic conductivity data for this study. Four of the eight boreholes were oriented at dip angles near 90 degrees. The other four have dip angles near 60 degrees. The smaller rectangle is the boundary of the simulation domain used in the stochastic simulations and tracer transport calculations.

fracture zones may have a profound influence on the flow properties of the region of interest. When large permeability contrasts are present alternative stochastic models that are able to mimic them should be considered.

Hydraulic conductivity values in the boreholes are shown in 3-D space in Figure 6. Each measurement is rendered as a disc centered at the measurement location. The radius of each disc is proportional to the value of logarithm of the hydraulic conductivity at that location. The color scale is inverted with warm colors representing low Log K.

Even a quick glance at the data suggests that they might not be adequately described by standard stochastic models. Each borehole shows large regions of relatively small log K, punctuated by sharp transitions to high values of log K. These regions of high K are presumably associated with fracture zones. This sort of behavior is problematic for Gaussian-based models, which tend to spread the disorder uniformly throughout the region. One method for dealing with this sort of heterogeneity is to manually place the fracture zones and then use a stochastic simulation for the residuals. This may be an effective strategy if geophysical or other data are available to unambiguously locate all fracture zones. Here it is presumed that this type of data is not available and the ability of fractal models to capture this sort of heterogeneity is evaluated.

3.2 Fractal analysis of the Äspö data

3.2.1 Variogram analysis

A conventional variogram analysis is considered as a starting point in the fractal analysis. A normal-score transform is first applied to the log K data. The transformed data thus have a standard normal distribution, which lessens the influence of extreme values and makes the variogram more robust.

Because data are available for only a limited number of boreholes, horizontal correlations are not expected to be detectable in this dataset. This has been confirmed by numerical experiments performed by Tsang et al. (Tsang *et al.*, 1996) using a dataset related to the present dataset. However, the fact that the boreholes are roughly divided into two sets according to their dip lends some hope that some understanding on vertical-to-horizontal anisotropy in the variogram can be obtained.

Two directional variograms are shown in Figure 7. In these plots, the increments have been divided into two direction classes, with dip angles of $75-90$ and 60 ± 15. Both variograms show a consistent rise over a wide range, consistent with fractal structure or long-range dependence. Beyond ≈ 200 m, the lag is approaching the limited range of the borehole data, and should

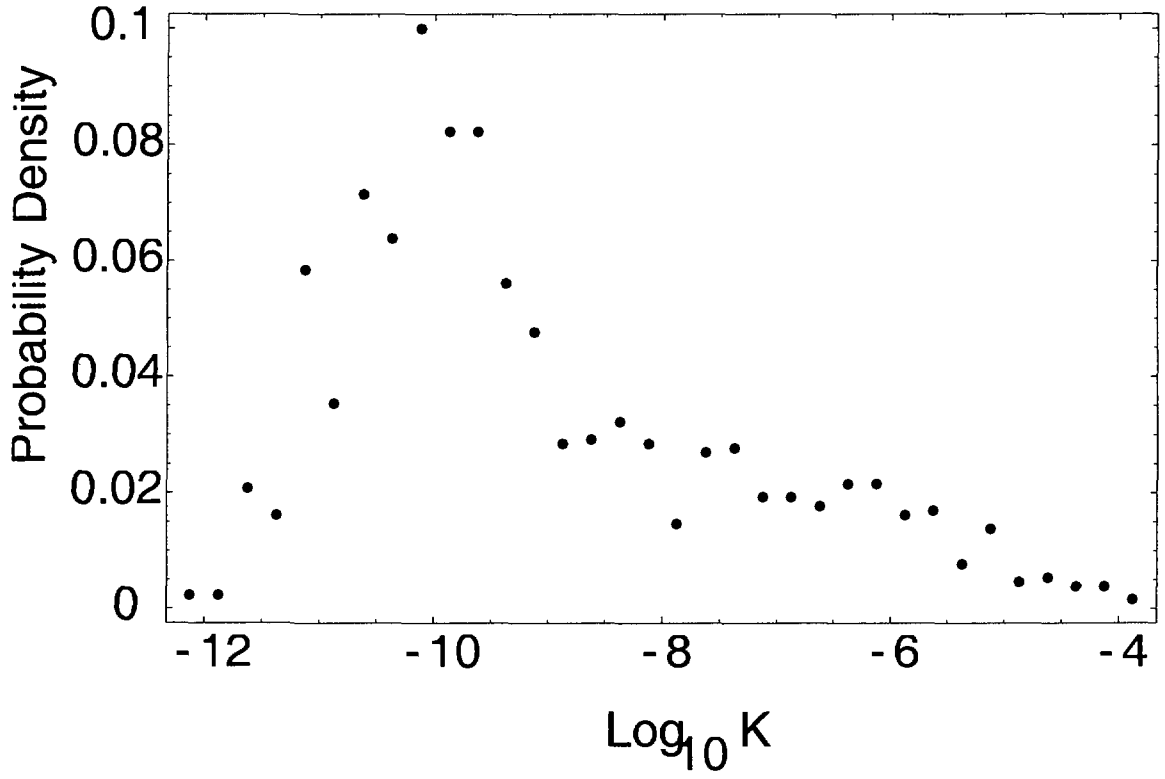

Figure 5: Histogram of logarithm (base 10) of hydraulic conductivity (m/s). The $\log K$ distribution has a mean of -9.08 and a standard deviation of 2.99. It is highly skewed to the right, suggesting that the K distribution is not well approximated by a log-normal distribution.

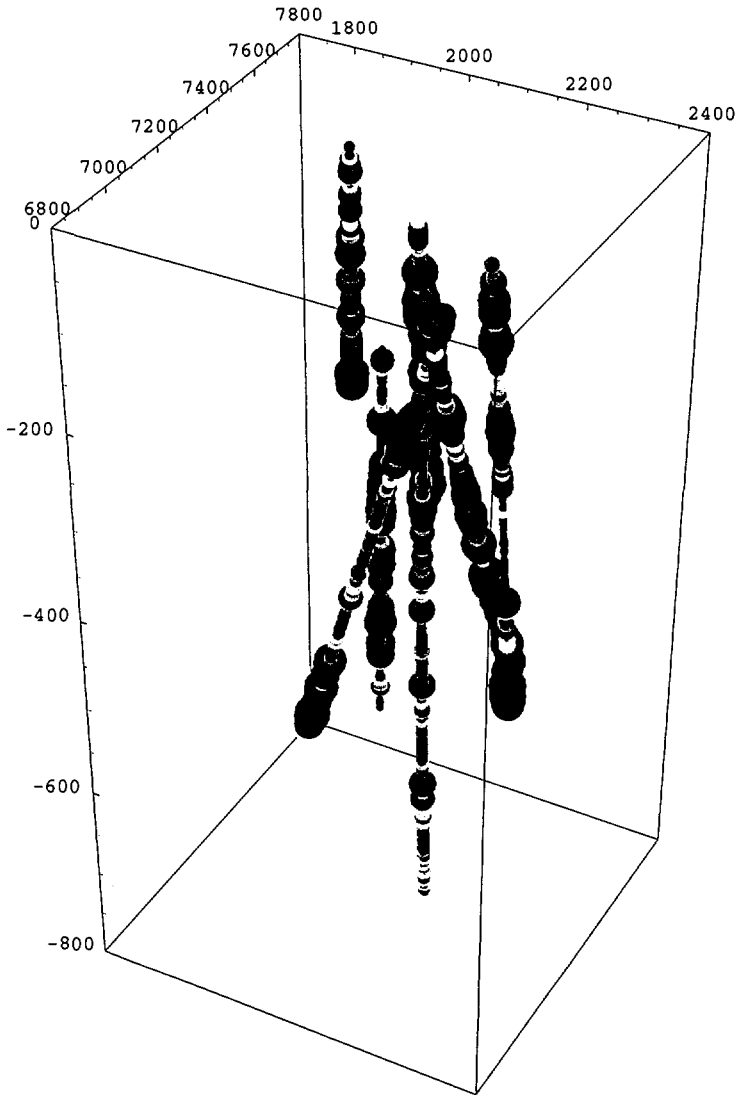

Figure 6: Hydraulic conductivity measurements in six of the eight boreholes used in this study. The radius of each disc is proportional to the value of Log K at that location. The color scale is inverted with warm colors representing low Log K.

be considered unreliable. Although it may be possible to fit a variogram model with limited range to this data with similarly good fit, the fractal model cannot be ruled out solely on this variogram. Indeed, the power-law fits to the two variogram, shown as the two solid lines on the log-log plot, are of good quality over a range of 6 m to more than 200 meters. There is a deviation from power-law behavior at lags of 3 m. However, this is near the scale-of-support for the 3 m packer tests which is expected to cause this sort of smoothing.

The main point to be observed from the data in Figure 7 is that the fractal power-law variogram fits the data at least as well as a more conventional model with finite range. This inability to distinguish between competing models is characteristic of many field investigations. If these variograms are to be input into performance assessment studies that require a conservative stance, it is thus necessary to evaluate both competing models or at least have an understanding of which leads to the more pessimistic prediction of repository performance.

A second point to be made from Figure 7 is that the directional variogram with dip of 60 degrees is consistently and unambiguously greater than the directional variogram with dip of 90 degrees. This implies stronger correlation in the vertical compared with the 30 degrees off-vertical. In other words, measurements at two points with a given separation distance differ less, on average, than those of two points separated horizontally by the same distance. This is consistent with open fractures or fracture zones being oriented primarily in the vertical direction.

The Hurst parameter H obtained from the slope of the fitted line on the log-log plot is 0.075 in Figure 7a and 0.13 in Figure 7b. Both of these values are much less than 0.5, indicating antipersistence. The different H values in the two directions is referred to as scaling anisotropy. Although models for scaling anisotropy are available (Lovejoy & Schertzer, 1985), conditional simulation codes are not currently available. For this reason, the two variograms are also fitted by power-laws with the same Hurst exponent (H=0.1), but different pre-factors (Figure 7c).

By comparing these two fitted model variograms, it is possible to arrive at an geometric anisotropy factor for use in the subsequent stochastic simulations. This does, however, require an assumption about the principal direction for the anisotropy ellipse. If the principal direction (direction of greatest correlation) is assumed to be vertical, then a geometric anisotropy factor of S=0.014 is calculated. In other words, the vertical axis is rescaled by a factor of 0.014 and an isotropic stochastic simulation is performed in the rescaled space. The value 0.014 was arrived at by calculating the degree of rescaling necessary for the vertical axis to

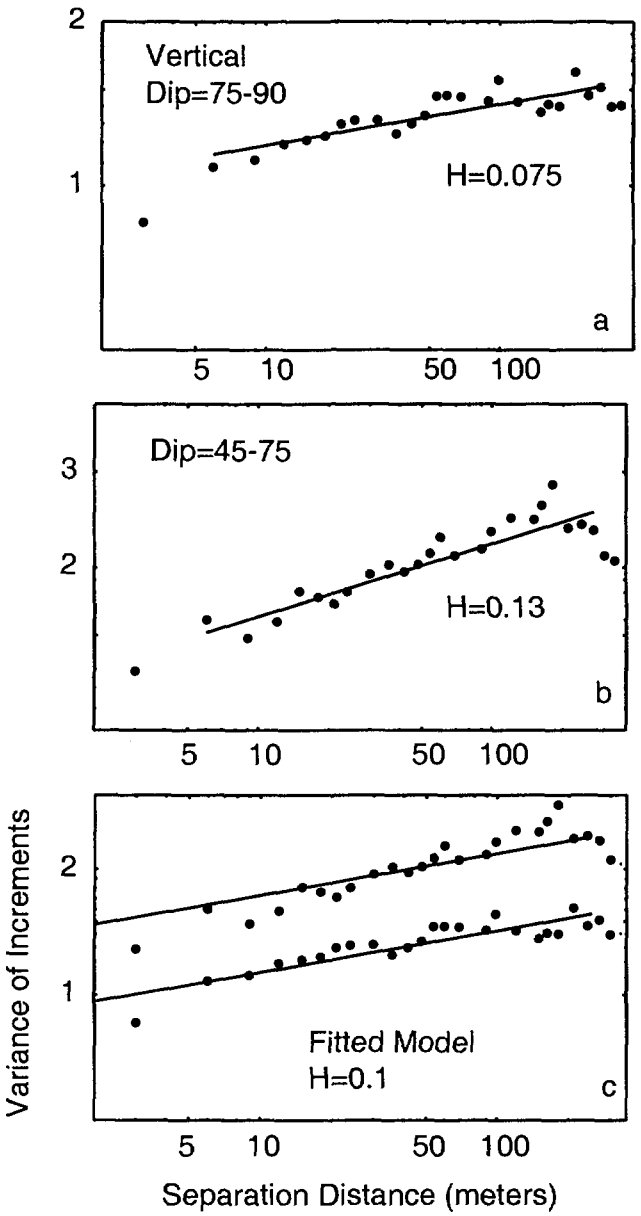

Figure 7: Directional variograms of Log K data after applying a normal-score transform. The top plot is for dip angles of 75-90 degrees while the middle one is for dip angles of 45-75 degrees. In both cases, the empirical variogram is well fitted by a power-law over a wide range, which is consistent with long-range spatial dependence. The bottom plot shows the two empirical variograms compared to the fitted model with H=0.1.

make two directional variograms coincide. Specifically, the vertical variogram is fitted by $\sigma_0^2(Sl)^{2H}$, and the variogram for the 60 degree boreholes by $\sigma_{60}^2[l_h^2 + (Sl_v)^2]^H = \sigma_{60}^2[\sin^2(30\,\text{deg}) + S^2\cos^2(30\,\text{deg})]^H l^{2H}$ where l_v and l_h are true separation distances in the vertical and horizontal directions and σ_{60}^2 and σ_0^2 are fitting parameters. The geometric anisotropy factor was then calculated by equating the two and solving for S.

3.2.2 Levy analysis

Levy analysis of the log K data is similar to the variogram analysis except for three important differences. First, the data are analyzed "as is" instead of after a normal-score transform. Second, the histogram of increments at short-lag are examined to see if they are fitted by a Levy probability density function. Third, the variogram is replaced by an analogous but more robust quantity.

Histograms of the increments in $\log K$ are shown in Figure 8. In Figure 8a the lag distance is 9 m, while in Figure 8b it is 30 m. Also shown in each is the best fit Levy pdf and a Gaussian pdf. The probability density is shown on a logarithm scale, so the Gaussian pdf has a downward parabolic shape. In both situations, the histogram of increments has slowly decaying tails. These are associated with abrupt changes in log K and are not modeled by the Gaussian distribution. The Levy PDF, in contrast, models this general shape of the histogram quite accurately. This sort of behavior has been identified previously in sedimentary rock, but this result is, to my knowledge, the first evidence for similar behavior in fractured crystalline rock.

The histogram of increments in Figure 8b is broader than the one shown in 8a. The fitted Levy distribution has $C = 0.77$ in Figure 8a as opposed to $C = 0.85$ in Figure 8b. Recall that the Levy width parameter measures the spread in the distribution about its center similar to the variance in a Gaussian distribution. The next step in the Levy fractal analysis is to examine the systematic change in this parameter with lag, similar to variogram analysis. The estimation of this C from data is based on a very robust quantile estimation procedure. The variation of C with lag, referred to here as the Levy structure function, is related to the generalized variogram of order 1 (referred to as the first order structure function in turbulence theory and as the madogram in geostatistics).

Two Levy structure functions and fitted power laws are shown in Figure 9. Here as in the variogram analysis the increments have been partitioned into two direction classes. The conclusions to be drawn from this analysis are similar to those in the variogram analysis: there is evidence for a power-law (fractal) structure over a wide range, the fitted Hurst parameters are

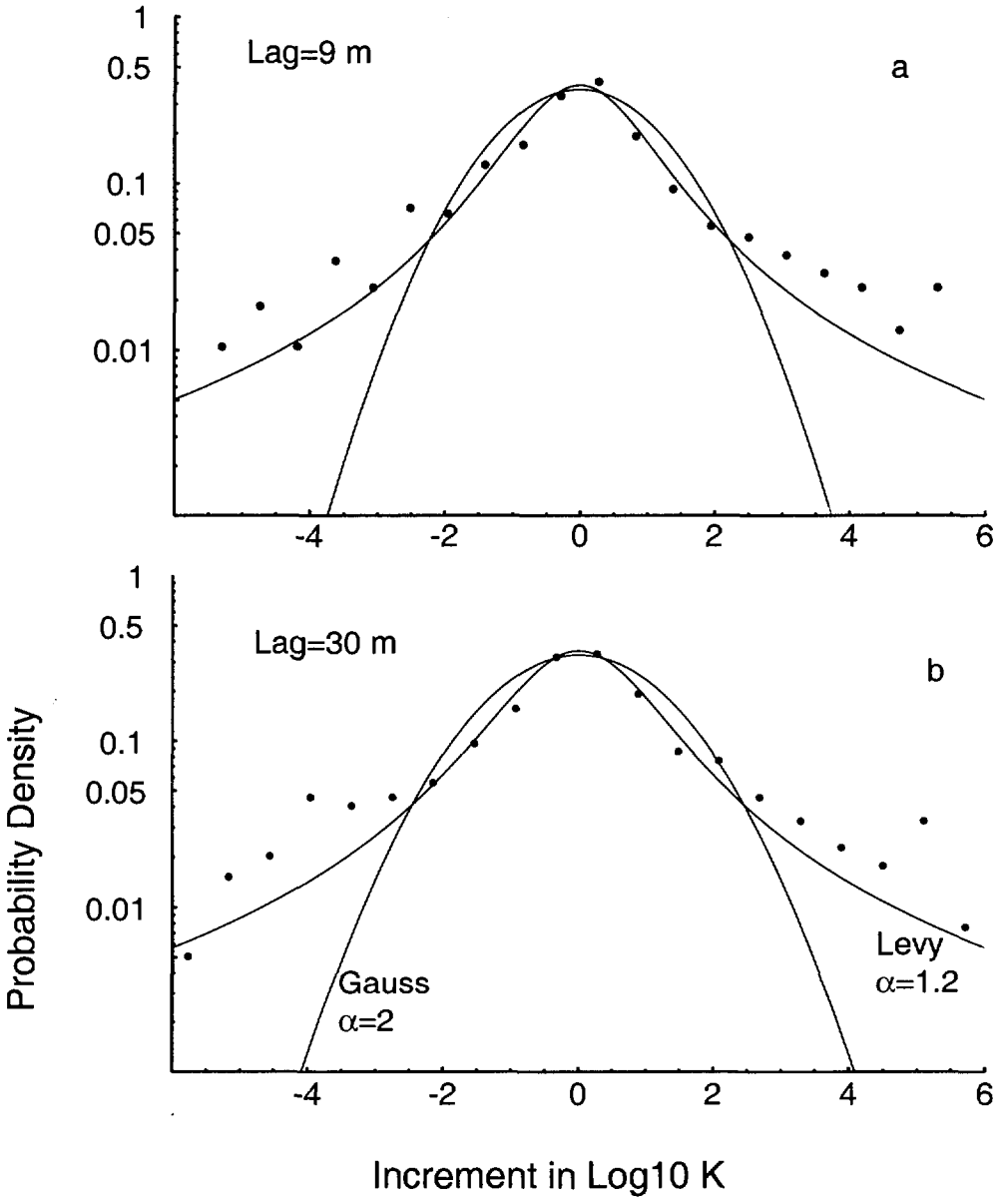

Figure 8: Levy PDFs fitted to histograms of increments in Log10 K for two values of the lag. The Levy PDF provides a better fit than the Gaussian distribution. This is consistent with a bfLm model for Log K.

consistent with antipersistence, and there is evidence for anisotropy with stronger correlation in the vertical direction.

Different Hurst parameters are found in the two directions, but in this case the differences are small. Fitting both Levy structure functions with the same power law with H=0.135 results in the following set of model parameters: $H = 0.135$, $\alpha = 1.2$, $C_0 = 0.88$ and $S = 0.028$, where S is the geometrical anisotropy factor. This fitted model is compared to the two directional structure functions in Figure 9c.

3.3 Conditional stochastic simulations

Having fitted an fBm and bfLm model to the log conductivity data, the next step is to perform stochastic simulations using the hydraulic conductivity data as conditioning data. The assumption is that the fractal characertistics observed in the conditioning data apply to the entire simulation domain.

A CSIRO internal simulation code (LevySim) was used for the conditional simulations. This code does conditional stochastic simulation of bfLm or fBm in 3D using regular or unstructured grids. The algorithms employed and validation studies have been published (Painter, 1998).

The simulation domain and the positions of the borehole conditioning data are shown in Figure 10. The domain was discretized into a $50 \times 70 \times 60$ grid with each gridblock being $10 \times 10 \times 10$ m^3. Conditioning data were relocated to the nearest gridblock center. When multiple data fell in the same gridblock, harmonic averaging was used. This is appropriate in the present situation because the boreholes and the flow direction in the subsequent simulations are oriented primarily in the same direction. No attempt to was made to correct for the differing scale-of-support between the conditioning data and the gridblock. Although it is expected that hydraulic conductivity in fractured rock will vary with scale-of-support, this is probably a minor correction in the present situation because the two scales do not differ by a large amount.

Sixteen stochastic simulations were performed for each of the bfLm and fBm models. For comparison purposes, sixteen simulations were also performed using a conventional multiGaussian model with an isotropic exponential covariance with 20 m range. The multiGaussian simulations used the normal-score transform of the log-conductivity data. Thus a unit variance and a zero mean were also specified. The inverse-normal-score transform was applied to each realization to obtain a log K realization with the same univariate distribution as the original log K data. The bfLm simulations required approximately 45 minutes each on a 333 MHz Pentium II running the LINUX operating system.

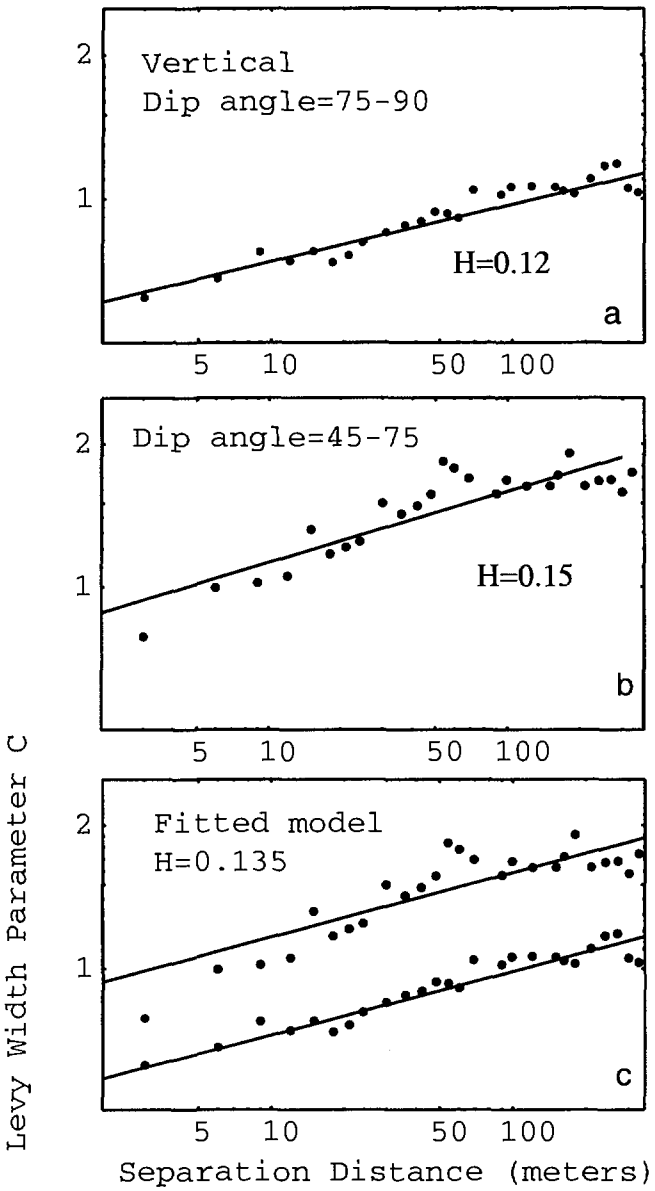

Figure 9: Levy structure functions for Log K fitted by a power-law model. These are analogous to the variograms shown in Figure 5, but are more robust against extreme variations in Log K. The power-law behavior over a wide-range is additional evidence for long-range spatial dependence. The bottom plot shows the two empirical variograms compared with the final fitted model (H=0.1).

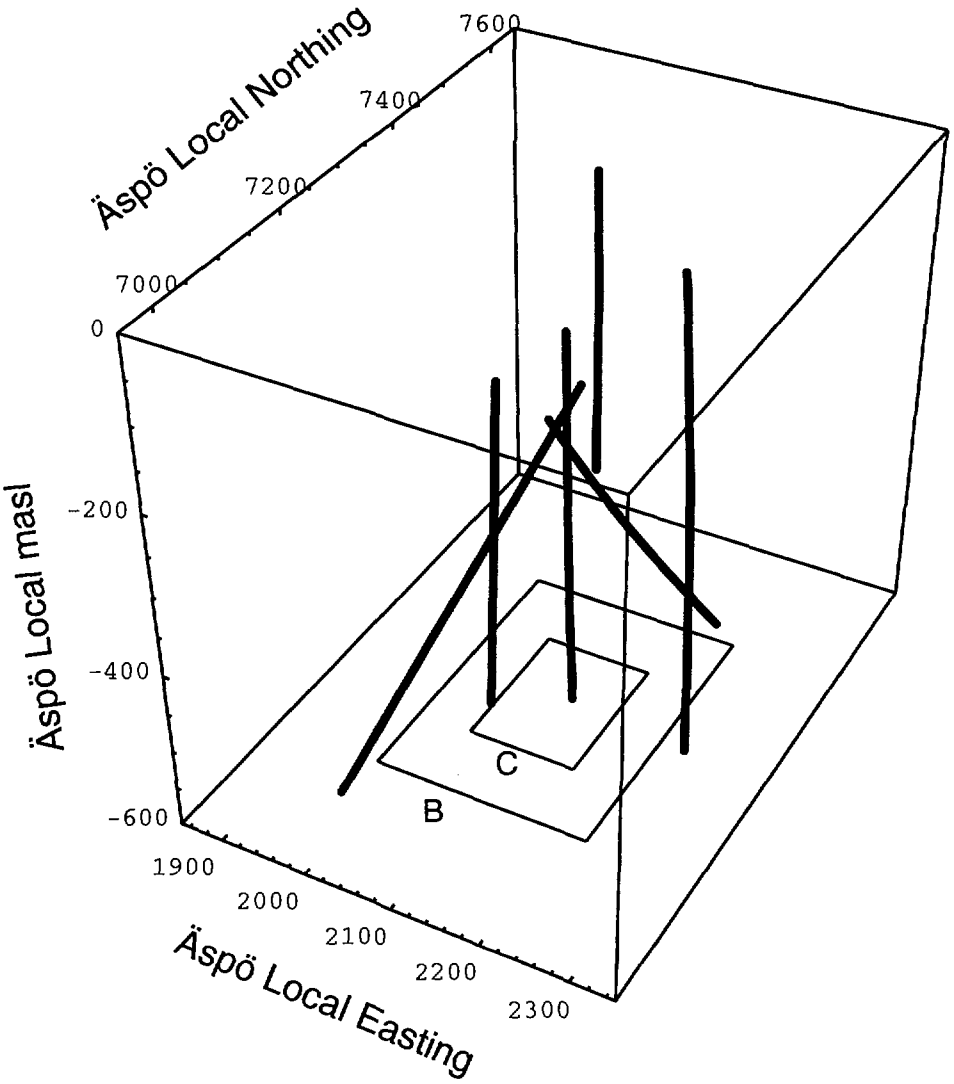

Figure 10: Domain and location of conditioning data for the stochastic simulation. The coordinates are in meters in the local Äspö coordinates. The two rectangles on the lower face mark the release area for release scenarios B and C.

One limitation of the fBm and bfLm formalisms is that, being non-stationary models, they offer no control over the univariate distribution. Further research is needed to address this limitation. For the purposes of the present study, a variant of the post-processing method originally proposed by Journel and Xu (1994) was used to correct the histogram of each stochastic simulation. The histogram correction method improves the reproduction of the univariate distribution while still honoring the conditioning data and without significant alterations to the spatial correlation pattern of the origin realization. The main modification to their original histogram correction method was to apply it recursively to improve the reproduction of the target univariate distribution.

One realization from the bfLm simulations after post-processing is shown in Figure 11. The combination of long-range dependence and high spatial variability associated with the Levy distributions results in significant high conductivity streaks like those associated with fracture zones. These high conductivity streaks can act as conduits for contaminant migration and it is particularly important for the stochastic model to capture these when they are present in the formation. These are missing from the multiGaussian simulation shown for comparison purposes in Figure 12. In this case the disorder is distributed more uniformly throughout the simulation domain.

4 Tracer Transport in the LRD Stochastic Continua

4.1 Flow and transport simulations

The three sets of conditional stochastic simulations based on the bfLm, fBm and multiGaussian models were used in particle tracking transport calculations in order to study the implications of LRD behavior on repository performance. The flow domain used is the same as for the stochastic simulations (Figure 10).

Laplace's equation was first solved for each of the 48 realizations of hydraulic conductivity to obtain hydraulic head under the assumption of single-phase steady state flow. Constant head was specified for the top and bottom faces of the rectangular domain with a head difference of 100 m leading to a macroscopic hydraulic gradient of 1/6. No flow conditions were applied at the four remaining faces of the 3D domain. Standard finite-difference methods were used to discretize Laplace's equation. Simultaneous over-relaxation (SOR) was used to solve for the nodal heads. These were then differenced to obtain the steady-state single phase darcy velocities at each node. A uniform

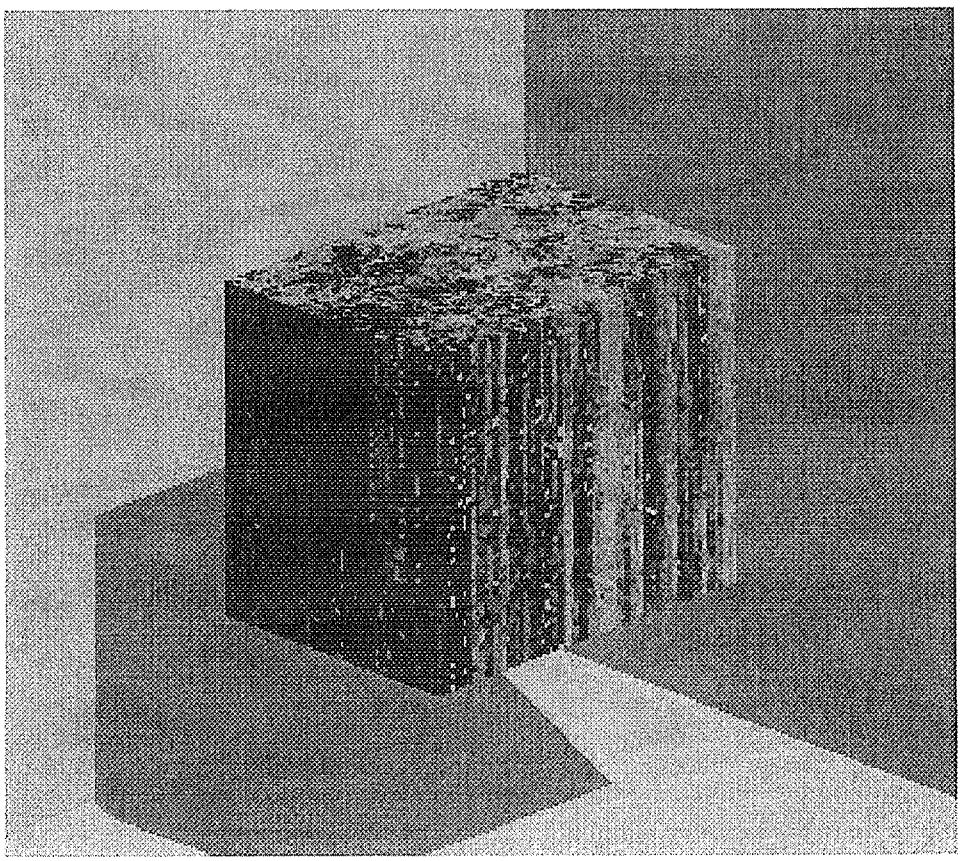

Figure 11: Conditional stochastic simulation generated using the LevySim code and the bfLm model. The long streaks of high and low conductivity are a consequence of the long-range dependence in the model. The grayscale is inverted; dark corresponds to high conductivity.

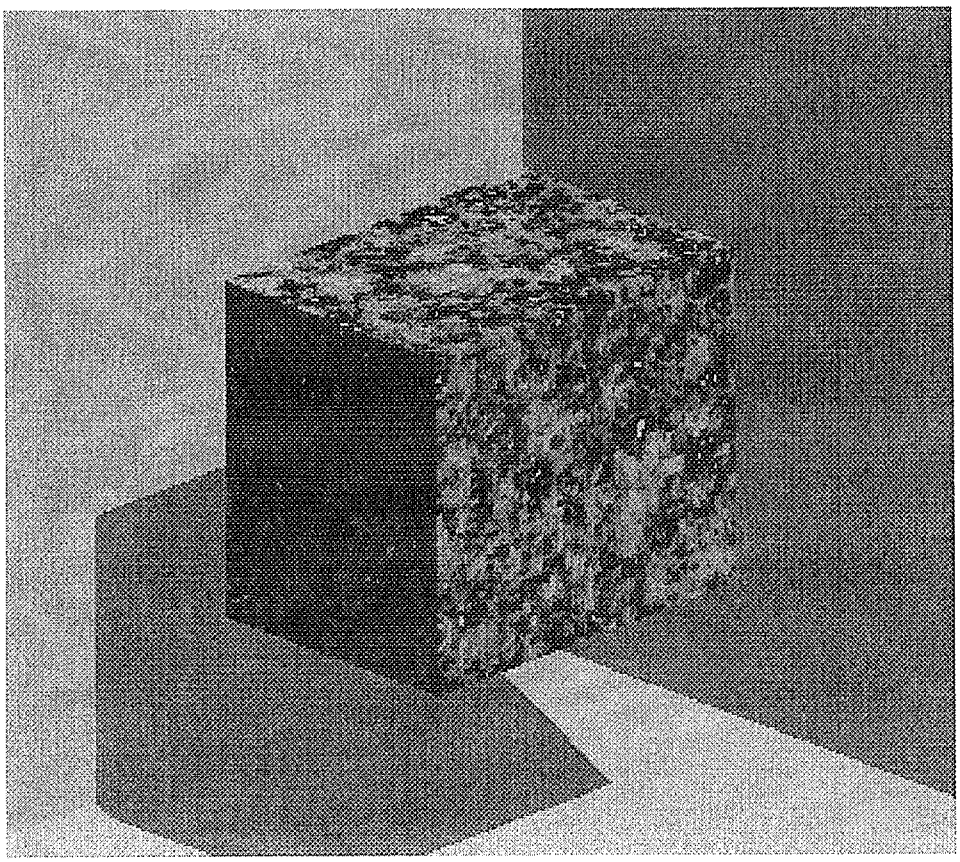

Figure 12: Conditional stochastic simulation generated from an isotropic multiGaussian model. The large regions of high and low conductivity seen in Figure 10 are missing. The grayscale is inverted in this image.

porosity was assumed and set to 1 for convenience.

The velocity field were then used in tracer transport simulations. Particles were launched from the bottom face of the domain and tracked along streamlines defined by the velocity field. Tri-linear interpolation was used to interpolate the velocities between nodes. The time required for each particle to travel across the domain and the x-y position at the top face were recorded for each particle.

Three release scenarios were considered. In scenario A, starting positions were chosen randomly with uniform distribution from the entire bottom face. In scenario B, the particles were launched from 1/4 of the bottom face as shown in Figure 10, while in scenario C, they were launched from the smaller rectangle shown in Figure 10 which covers 1/16th of the bottom face. 1000 particles were used in each realization.

4.2 Flow channeling

One characteristic of flow in fractured rock is that it tends to be highly channeled, i.e. the majority of the flow often occurs through relatively small regions of the overall domain due to heterogeneity. The ability of stochastic continuum models to reproduce this type of behavior is an important first test.

To study the relationship between flow chaneling and the choice of stochastic model, particles were distributed randomly and uniformly across the inlet side of the domain (release scenario A). Each particle was assigned a statistical weight according to the local flow velocity at the release site. This is equivalent to a particle source that is proportional to the local flow velocity. The top (outlet) face was partitioned into 50×70 m^2 panels and the fraction of the statistical weight arriving at each panel was calculated.

Results are shown in Figure 13 for one realization from each of the bfLm, fBm and multiGaussian models (counterclockwise from top). All three cases show some degree of flow channeling. In the multiGaussian case it is due to the high variance (2.99) in Log K for this data set. Channeling in the fBm and bfLm cases is much more pronounced due to the long-range spatial dependence. In the bfLm case the vast majority of the tracer mass arrives in a single panel. This illustrates how the important phenomenon of extreme flow channeling can be captured in a continuum model that combines high variability with long-range dependence.

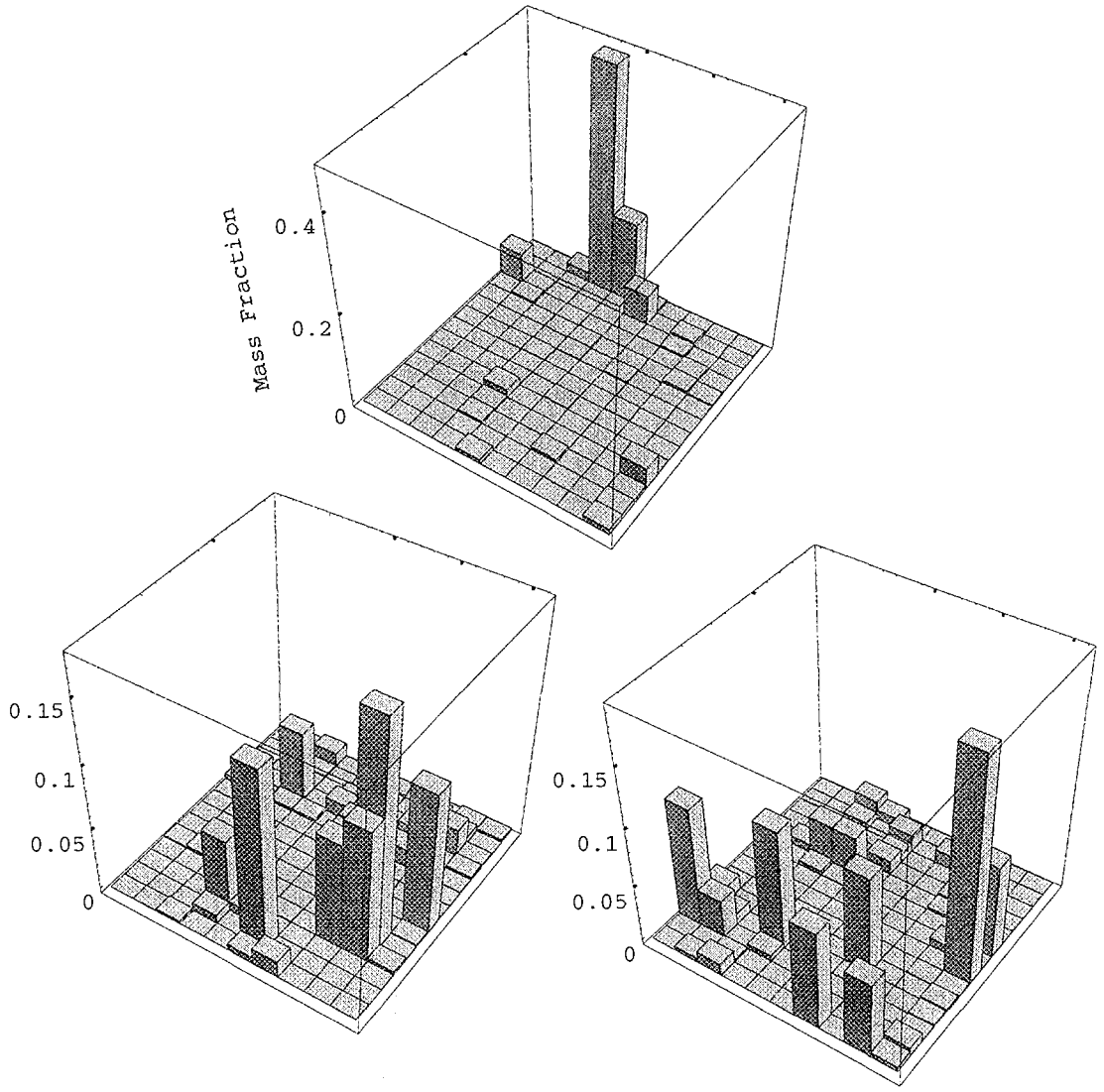

Figure 13: Spatial distribution of mass arrival at the top face of the simulation domain illustrating flow channeling in the stochastic continuum representation. The stochastically generated $\log K$ fields utilized the bfLm, fBm and multiGaussian models (counterclockwise from top). All three show evidence for flow channeling, but it is much stronger in the bfLm case. Particles were released from the entire bottom face of the simulation cube, with the particle source proportional to the local flow rate.

4.3 Sensitivity of predicted breakthrough curves to choice of heterogeneity model

Cumulative breakthrough curves are shown in Figure 14 using the three heterogeneity models and realease scenario A. The particle source was constant across the release area, not weighted by the flow velocity. The vertical axis is the cumulative mass fraction recovered at the top face. The horizontal axis is base-10 logarithm of time in thousands of years. The curves in Figure 14a were obtained using the 16 conductivity realizations generated using the bfLm model. Figures 14b and 14c show the same for the fBm and multiGaussian models, respectively. The aim is to explore how sensitive probabilistic predictions of geosphere performance are to the choice of heterogeneity model.

The two models with long-range dependence result in earlier arrival times compared with the multiGaussian model. If we take as a measure of the bulk arrival time τ_{50}, the time at which 50% of the mass is recovered at the outlet, then the ensemble average bulk arrival time $\langle\tau_{50}\rangle$ is 5.4×10^4 years for the bfLm model, 6.0×10^4 years for the fBm model and 1.8×10^5 years for the multiGaussian model. More important, if the focus is on the arrival of the leading edge of the contaminant plume, then the differences are more dramatic. For example, $\langle\tau_5\rangle = 3.0 \times 10^2$, 3.8×10^2, and 1.3×10^4 for the bfLm, fBm and multiGaussian models, respectively. Here $\langle\tau_5\rangle$ is the time at which 5% of the initial mass is recovered at the outlet. It should be noted that these times are dependent on the choice of macroscopic head gradient and the assumed porosity; they may not represent true times but could be scaled to true times if desired.

The differences in the predicted $\langle\tau_{50}\rangle$ and $\langle\tau_5\rangle$ have some potentially important implications for repository performance assessment. In this example, the differences due to the choice of heterogeneity model are much greater than the realization-to-realization variability. This result demonstrates the importance of exploring alternative models for heterogeneity in situations where the data do not provide unambiguous support for a single heterogeneity model. Further, the two LRD models predict much earlier arrival times for the leading edge of the contaminant plume. This is due to rapid transport pathways resulting from spatially connected regions of high conductivity. As these are intrinsic features of the LRD models, the result of more rapid breakthrough for the LRD models is expected to be a general one. From the point of view of conservative assessment of repository geosphere performance, it is thus necessary to consider LRD model for heterogeneity unless the data unambiguously rule these out.

The predicted breakthrough curves depend not only on the choice of geo-

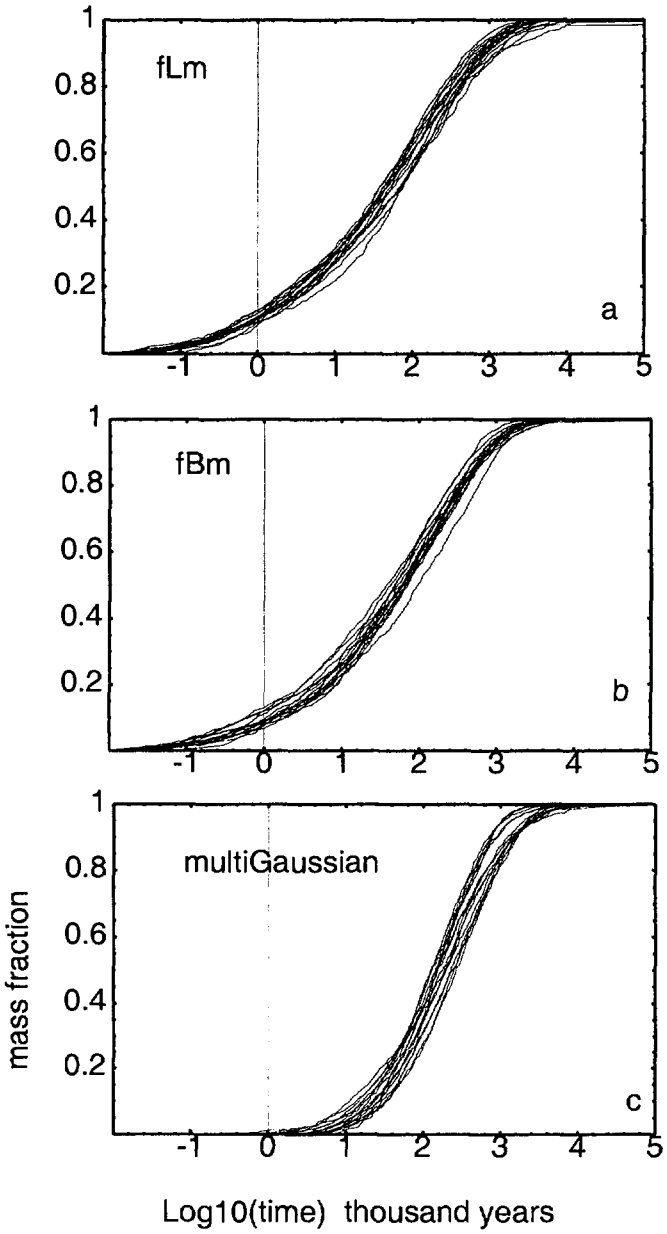

Figure 14: Cumulative breakthrough curves for release scenario A (particles randomly distributed across the bottom face of the simulation domain). The stochastically generated $\log K$ fields employed the bfLm, fBm, and multi-Gaussian models (reading from the top). The two models with long-range dependence predict earlier contaminant arrival.

statistical model, but also on the conditioning data and release scenario. Breakthrough curves calculated by assuming the release scenario B (see Figure 10) are shown in Figure 15. Although the realization-to-realization variability in Figure 15 is significantly larger than in Figure 14 because of the reduced spatial extent of the particle source, there are still significant differences between the three models with the LRD models again predicting significantly earlier arrival times.

4.4 On the relationship between intermediate-scale and large-scale transport properties

Relating geosphere transport properties to intermediate (or block) scale field measurements is an important component to repository performance studies. If the measurement scale is several times the hetorogeneity correlation range, the asymptotic Fickian limit is an appropriate approximation and a dispersion due to heterogeneity can be described by a classical advection/dispersion equation with constant velocity and dispersivity. The conventional method for analyzing block-scale field measurements is to assume that the Fickian limit is reached; dispersivities and velocities obtained from block-scale measurements can then be applied to the larger scales. For media with long-range dependence the Fickian limit is not reached regardless of the measurement scale. Thus care should be taken in extrapolating transport parameters obtained from conventional analysis of intermediate-scale experiments to the larger scale.

The relationship between large-scale and intermediate-scale transport properties in LRD media is explored in Figure 16. The 16 bfLm realizations of $\log K$ were used in tracer transport simulations as described previously. In this example, particles were released from the smallest rectangle labeled "C" in Figure 10. The bulk arrival time at 100 m τ_{50}(100 m) and at 600 m τ_{50}(600 m) were recorded. Although the number of realizations is very small, it is clear in the crossplot shown in Figure 16 that there is a strong relationship between the intermediate-scale measurements and the effect of interest. Transport is, however, clearly not Fickian. A simple linear extrapolation from the 100 m scale to the 600 m scale - as would be appropriate in the Fickian limit - does not give accurate results. In particular the advective velocity is clearly increasing with travel distance. Simple extrapolation would thus give non-conservative results in a performance assessment.

There is also considerable scatter in the crossplot in Figure 16, indicating that the two quantities are related only in a statistical sense. This uncertainty should be considered explicitly when upscaling field measurements. Further,

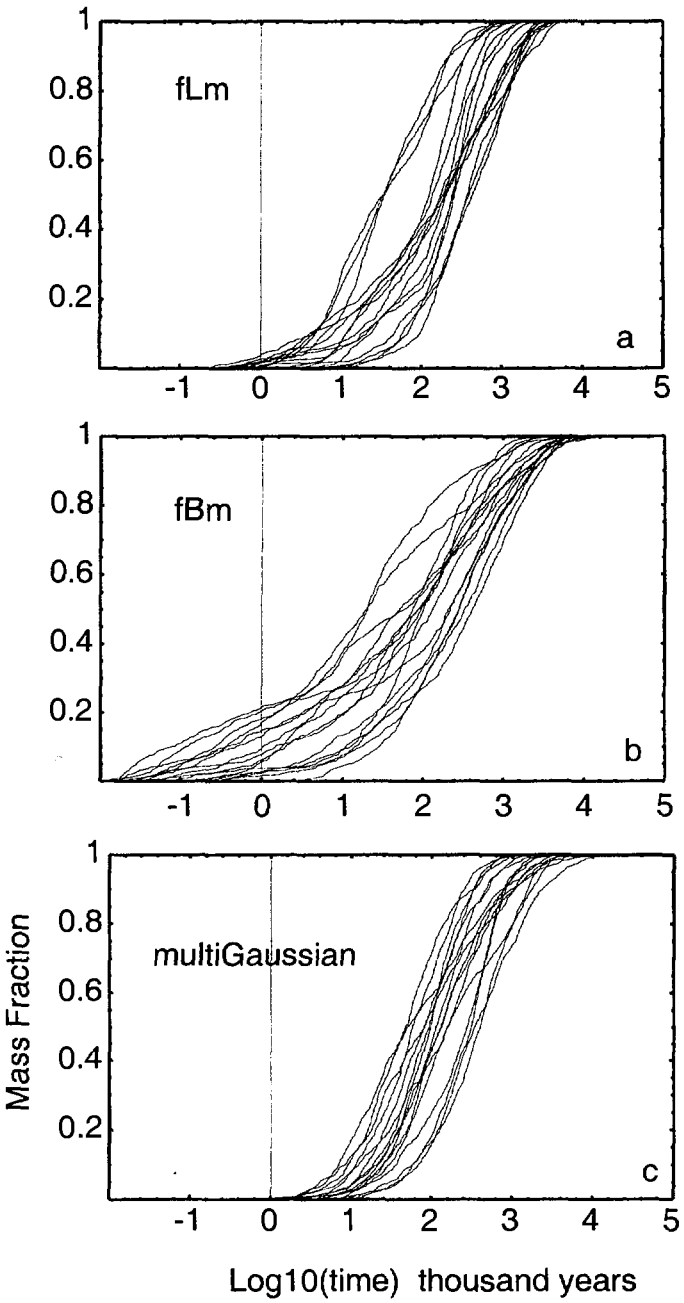

Figure 15: Same as Figure 14 except for the release scenario. In this case the particles were randomly and uniformly distributed across the rectangle B in Figure 10.

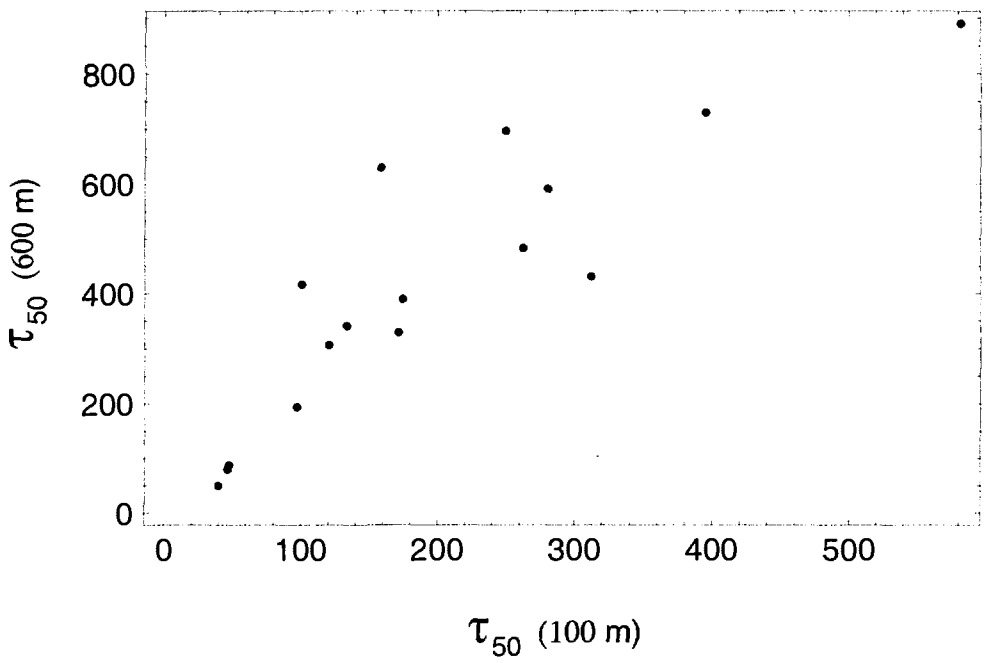

Figure 16: Crossplot of bulk arrival time at 100 m and 600 m obtained from 16 conditional realization of $\log K$ generated using the bfLm model. The Fickian limit is not reached in the bfLm media. There is thus no unique relationship between transport properties measured at different scales. Monte Carlo calculations like those summarized here can by used to upscale the results of block-scale experiments to obtain effective transport properties for the entire domain.

any extrapolation from the intermediate-scale experiments to the large-scale performance should also take into account other data such as borehole measurements of hydraulic conductivity. This problem of merging different types of data in a stochastic framework is one of the most challenging issues in geostatistics. For example, no general methods exist for constraining a stochastic simulation to match both quasi-point measurements and large-scale dynamic data such as the results of tracer tests.

Here a new statistical method for extrapolating intermediate-scale dynamic data to larger scales is proposed. This method accounts for constraints imposed by conditioning data in addition to the intermediate-scale data, and does not require an analytical model for the tracer spreading. The final result is the probability distribution for the effect of interest conditional on the intermediate-scale data and hydraulic conductivity data measured at much smaller scales.

The basic idea is to simulate the intermediate-scale experiment and the large-scale effect of interest using the same stochastically generated K realization. This is repeated a large number of times to build up, in a Monte Carlo sense, the probabilistic relationship between the proxy variable (intermediate-scale measurement) and the target variable (large-scale transport property). The method is easily conditioned to borehole measurements of K by simply making the K generation step a conditional stochastic simulation. Given an actual field measurement of the proxy variable, the probability distribution of the target variable conditional on the proxy measurement is then extracted from the Monte Carlo results.

The issue then is how to take the Monte Carlo results and construct the conditional probability density function for the target variable. The crudest method would be to bin the data into classes and then construct the empirical distribution from the Monte Carlo results in each class. More sophisticated methods that are able to extract more information from the limited number of Monte Carlo results exist for this purpose. In particular, nonparametric methods for fitting conditional quantile functions (Koenker *et al.*, 1994) are well developed. Doing this for several quantiles effectively approximates the entire conditional cumulative distribution function. The method can be made computationally efficient for the present application by using a local linear approximation to the various quantiles in the neighborhood of the given proxy measurement. This approximation should make it feasible to extend this method to the situation of multiple proxy variables. Development and testing of this method will be the subject of future research.

5 Summary

One reason for the interest in discrete-fracture network models is that the alternative, the stochastic continuum models, have traditionally been based on the classical random field models of subsurface hydrology such as the multiGaussian model. These classical models have difficulty modeling abrupt changes in hydraulic conductivity and spatially connected regions of extreme values such as those associated with fracture zones. Given that high conductivity regions associated with fracture zones provide flow and transport pathways through a given volume of rock, it is particularly important to have stochastic models that are capable of realistically modeling these features. In a recent paper (Tsang *et al.*, 1996), a class of geostatistical models based on nonparametric indicator methods was proposed as a framework for stochastic continuum level descriptions of fractured rock. The present report has similar goals, the evaluation of nonclassical stochastic continuum models for fractured rock. However, the focus here is on random field models with long-range spatial dependence. Specifically, a class of fractal scaling models known as bounded fractional Levy motion (bfLm) is evaluated using hydraulic conductivity data from packer tests from boreholes in the vicinity of the Hard Rock Laboratory, Äspö Island, Sweden. The bfLm class of models includes the well-known fractal scaling model fractional Brownian motion (fBm) as a special case, but has increased flexibility to model formations with high spatial variability.

Fractal analysis of the $\log K$ suggests they are consistent with a fractal scaling model. In particular, a bfLm model was fitted to the $\log K$ data and an fBm model to the normal score transform of the $\log K$ data. Both models provided a reasonably accurate fit to the data. However, a conventional variogram model with large nugget and limited spatial correlation range may also provide a similar good fit to the data. It is not possible to distinguish between the competing 3-D models with the limited amount of hydraulic data available.

This inability to distinguish between competing models occurs frequently in practice and is likely to occur also in evaluations of future waste repository sites. Given the need to take a conservative stance in these performance assessment studies, it is particularly important to evaluate alternative models or to at least have an understanding of how the choice of random field model influences the predicted performance. Stochastic simulations conditioned by the Äspö $\log K$ data were coupled with flow and tracer transport calculations to address this relationship between the random field model and the predicted transport properties.

Stochastic simulations based on the two long-range dependent models

contain high-conductivity streaks that mimic fracture zones. Qualitatively, the bfLm model produced better defined high-conductivity streaks as compared to the fBm model, but the differences are small compared to the differences with the classical multiGaussian model. All three models resulted in significant flow channeling, with the bfLm simulations being the strongest followed by the fBm model and the multiGaussian model.

Nonreacting tracer transport simulations utilizing stochastically generated K fields show different breakthrough curves depending on the choice of the random field model. In general the random field models with long-range dependence predict earlier arrival time for the mass and significantly more dispersion in a single-realization sense. This has some important implications for performance assessment. In order to be conservative in the prediction of geosphere performance, these LRD models need to be considered unless the data specifically rule them out.

Finally, long-range dependence has some implications for the extrapolation of intermediate-scale experiments to larger scales. Since the Fickian limit is never reached for transport in fractal media, the conventional practice of fitting an effective dispersivity to intermediate-scale tracer experiments may result in an inaccurate prediction of the larger-scale transport behavior. In particular, it may introduce a systematic bias and be non-conservative from a performance assessment perspective. It also ignores completely the issue of prediction uncertainty. A new Monte Carlo method for up-scaling the results of intermediate-scale experiments was outlined. This method incorporates small-scale hydraulic conductivity measurements in addition to the intermediate-scale dynamic data, and treats explicitly the issue of prediction uncertainty. Implementation and testing of the new method will be the focus of future research.

References

Boggs, J. M., Young, S. C., Beard, L. M., Gelhar, L. W., Rehfeldt, K. R., & Adams, E. E. 1992. A field study of dispersion in a heterogeneous aquifer, 1: Overview and site description. *Water Resources Research*, **28**, 3281–3292.

Crane, S. D., & Tubman, K. M. 1995. Reservoir Variability and Modeling with Fractals. *Chap. 13 of:* Barton, C. C., & Pointe, P. R. La (eds), *Fractals in Petroleum Geology and Earth Processes.* New York City: Plenum Press.

Feder, J. 1988. *Fractals.* New York: Plenum.

Feller, W. 1971. *An Introduction to Probability Theory and its Applications, Vol.2.* New York: Wiley.

Hewett, T. 1986. Fractal distributions of reservoir heterogeneity and their influence on fluid transport. *In: Proceedings of Society of Petroleum Engineers Annual Technical Meeting.*

Journel, A. G., & Huijbregts, C. J. 1978. *Mining Geostatistics.* New York City: Academic Press.

Journel, A. G., & Xu, W. 1994. Posterior identification of histograms conditional to local data. *Water Resources Research*, **26**(3), 323–359.

Koenker, Roger, Ng, Pin, & Portnoy, Stephen. 1994. Quantile Smoothing Splines. *Biometrika*, **81**(4), 673–680.

Kumar, P., & Foufoula-Georgiou, E. 1993. A multicomponent decomposition of spatial rainfall fields 2. Self-similarity in fluctuations. *Water Resources Research*, **29**, 2533–2544.

Levy, P. 1937. *Théorie de l'Addition des Variables Aléatoires.* Paris: Gauthier-Villars.

Liu, H. H., & Molz, Fred J. 1997. Comment on 'Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations'. *Water Resources Research*, **33**(4), x–908.

Lovejoy, S., & Schertzer, D. 1985. Generalized Scale Invariance in the Atmosphere and Fractal Models of Rain. *Water Resources Research*, **21**(8), 1233–1250.

Mandelbrot, Benoit, & Ness, J. W. Van. 1968. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, **10**, 422–437.

Molz, Fred J., & Boman, G. K. 1993. A fractal-based stochastic interpolation scheme in subsurface hydrology. *Water Resources Research*, **29**, 3769–3774.

Molz, F. J., & Boman, G. K. 1995. Further evidence of fractal structure in hydraulic conductivity distributions. *Water Resources Research*, **22**(18), 2545–2548.

Molz, F. J., Liu, H. H., & Szulga, J. 1997. Fractional Brownian motion and fractional Gaussian noise in subsurface hydrology: A review, presentation of fundamental properties, and extensions. *Water Resources Research*, **33**(10), 2273–2286.

Neuman, S. P. 1990. Universal Scaling of Hydraulic Conductivities and Dispersivities in Geologic Media. *Water Resources Research*, **26**(8), 1749–1758.

Painter, S. 1995. Random fractal models of heterogeneity: the Levy-stable approach. *Mathematical Geology*, **27**, 813–830.

Painter, S. 1996. Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations. *Water Resources Research*, **32**(5), 1183–1195.

Painter, S. 1998. Numerical method for conditional simulation of Levy random fields. *Mathematical Geology*, **30**(2), 163–179.

Painter, S, & Paterson, L. 1994. Fractional Levy motion as a model for spatial variability in sedimentary rock. *Geophysical Research Letters*, **21**, 2857–2860.

Painter, S., Beresford, G., & Paterson, L. 1995. On the distribution of seismic amplitudes and seismic reflection coefficients. *Geophysics*, **60**, 1187–1194.

Rhen, I., Gustafson, G., Stanfors, R. & Wikberg, P. 1997. Geoscientific evaluations 1997/5. Models based on site characterization 1986-1995. *SKB TR 97-06.*

Taqqu, M. S. 1987. Random processes with long-range dependence and high variability. *Journal of Geophysical Research*, **92**, 9683–9686.

Todoeschuck, J. P., Jensen, O. G., & Labonte, S. 1990. Scaling Geology: evidence from well logs. *Geophysics*, **55**, 480–484.

Tsang, Y. W., Tsang, C. F., Hale, F. V., & Dverstorp, B. 1996. Tracer transport in a stochastic continuum model of fractured media. *Water Resources Research*, **32**(10), 3077–3092.

Zolotarev, V. M. 1986. *One-dimensional Stable Distributions.* Providence Rhode Island: American Mathematical Society.