Heavy tailed K distributions imply a fractional advection dispersion equation

Kohlbecker, Matthew V.

Wheatcraft, Stephen

Meerschaert, Mark M.

Abstract

Benson (1998) introduced a Fractional Advection Dispersion Equation (FADE) to model contaminant transport in porous media. This equation characterizes contaminant plume dispersion with a fractional derivative of order $1 \le \alpha \le 2$ and has solutions that are heavy-tailed Levy α-stable densities. The FADE assumes that α-stable hydraulic conductivity distributions, characterized by tail parameter α_K, give rise to α-stable velocity distributions, characterized by tail parameter α_v. This work investigates whether this is a valid assumption.

A computer algorithm was written to generate *K* fields. The algorithm uses a modification of the spectral synthesis method for generating fractional Brownian motion (fBm). The USGS finite difference groundwater code MODFLOW was used to calculate the velocity fields, and the *K* and *v* fields were analyzed to determine whether they were consistent with a stable PDF. After determining that the *K* and *v* fields were consistent with a stable PDF, the tail parameters describing each pair of *K* and *v* fields, $\alpha_{\log K}$ and $\alpha_{\log v}$, were measured with the Nolan (1997) Maximum Likelihood Estimator (MLE). The relationship between $\alpha_{\log K}$ and $\alpha_{\log v}$ was then determined.

There are three conclusions of this research. First, Mandelbrot and PP plots confirm that stable distributions of log *K* give rise to stable distributions of log *v*. Second, $\alpha_{\log K} > \alpha_{\log v}$. Finally, there is a positive statistical relation between $\alpha_{\log K}$ and $\alpha_{\log v}$.

1.0 Introduction

1.1 Hydraulic Conductivity, Velocity, and Particle Jumps

Equations of contaminant transport in porous media are based on assumptions about hydraulic conductivity. These assumptions concern the probability density function (PDF) of *K* (Mercado, 1967; Schwartz, 1977; Smith and Schwartz, 1980, 1981a, 1981b), stationarity of *K* (Gelhar et al., 1979; Gelhar and Axness, 1983), nonstationarity of *K* (Dagan, 1984), or fractal nature of *K* (Wheatcraft and Tyler, 1988; Neuman, 1990; Benson, 1998). The reason for this emphasis on *K* is found in Darcy's Law, the equation governing groundwater flow (e.g., Freeze and Cherry, 1979):

$$v = \frac{-K}{\eta}\nabla h \tag{1}$$

where v is average velocity of a 'parcel' of water, ∇h is the hydraulic gradient, η is porosity, and K is hydraulic conductivity. Since field-measured values of K vary over 13 orders of magnitude and η varies over only 1 order of magnitude (Freeze and Cherry, 1979), differences in the velocity field are dominated by differences in the hydraulic conductivity field as per equation (1) (assuming the contribution of ∇h is also negligible[1]). Note that (1) also implies a linear relation between $\log v$ and $\log K$. The differences in the velocity field result in plume spreading at rates faster or slower than the advective groundwater velocity, a macroscopic dispersion often called differential advection.

[1] ∇h is inversely correlated with K and, like η, varies over a relatively small range of magnitudes in a typical field setting. Determining the significance of ∇h in (1) is part of this work.

Equation (1) is directly relevant to an equation of contaminant transport that assumes a heavy-tailed distribution of K: the Fractional Advection Dispersion Equation (FADE) (Benson, 1998, Benson et al., 2000). This paper discusses the effect of a Levy α-stable distribution of K on the resulting velocity field with a focus on the implications for the FADE.

1.2 Other Approaches to Contaminant Transport

Solute particle velocity can be thought of as a jump; a displacement of random magnitude that a particle undergoes in a discrete amount of time. The Central Limit Theorem (CLT) of probability and statistics is used to understand the spatial distribution of these particles after a large number of jumps have been completed (e.g., Bear, 1972):

$$\frac{x_1 + x_2 + x_3 \quad x_n - n\mu}{\sigma n^{1/2}} \Rightarrow Z \tag{2}$$

where $x_1, x_2, x_3, \ldots x_n$ are independent identically distributed (IID) random variables from a finite variance distribution representing particle jumps, μ is the mean jump size, σ is the standard deviation of the jump size, n is the number of jumps, and Z is a Gaussian random variable to which the sums converge. According to equation (2), the sums of IID random variables from a finite variance probability distribution converge to a Gaussian distribution as $n \to \infty$ (e.g., Bear, 1972). If, then, we have several solute particles with jump magnitudes governed by a finite variance PDF, the accumulated particle jumps will converge to a Gaussian distribution by the CLT.

This approach to contaminant transport—assume particle jumps/velocities are governed by the CLT and related to K by Darcy's Law—has been taken, either implicitly

or explicitly, in equations of contaminant transport such as the Advection Dispersion Equation (ADE) (e.g., Bear, 1972):

$$\frac{\partial C}{\partial t} = -v\frac{\partial C}{\partial x} + \frac{\partial}{\partial x}\left(D\frac{\partial C}{\partial x} \right) \tag{3}$$

where C is concentration, t is time, x is distance, v is velocity, and D is the dispersion coefficient.

In order for equation (3) to accurately describe solute transport, equation (2) must describe particle movement, so the PDF of the $x_1, x_2, x_3, \ldots x_n$ in equation (2) must be a finite variance distribution. Therefore, by equation (1), K should follow a finite variance distribution. Freeze (1975) provided evidence that supported approximating K with a finite variance PDF—log normal. Therefore, the finite variance requirement in (2) was satisfied, and equation (3) was thought to be a valid equation for contaminant transport.

The result of the finite variance assumption is that the concentration profile of a solute plume is Gaussian (Bear, 1961) and plume variance will grow like $t^{1/2}$ (Bear, 1979). Several field studies of macroscopic solute dispersion have indicated, however, that plumes are highly non-Gaussian (Freyberg (1986) and Adams and Gelhar (1992)), and plume variance grows faster than $t^{1/2}$ (Mercado (1967), Pickens and Grisak (1981), Benson et al. (2001)), called super-Fickian dispersion.

Painter (1996a, 1996b), Benson et al. (2001), and Liu and Molz (1997) have presented evidence for approximation of K with an infinite variance PDF, which suggests that a theorem other than equation (2) is appropriate for describing solute migration.

1.3 The Fractional Advection Dispersion Equation

1.3.1 Levy's α-stable densities

An understanding of the FADE requires an understanding of Levy's α-stable densities. These densities belong to a class of probability distributions distinguished from one another by a tail parameter, α, where $0 < \alpha \leq 2$. They have the characteristic of 'heavy tails,' where the tail parameter α describes the 'heaviness' of the tails. Heavy-tailed distributions exhibit power law decay in the tails of the distribution (as compared to, for example, exponential decay in the tails of a Gaussian distribution). Letting $f(x)$ indicate the PDF of a Levy distribution, power law decay can be expressed as:

$$\lim_{x \to \infty} f(x) = Cx^{-1-\alpha} \tag{4}$$

where C is a constant and x is a random variable.

Figure 1 shows Levy's α-stable densities on both arithmetic and log-log axes for several values of α. The special case of the Gaussian, α = 2, is also shown. Note that as α approaches 1, the mass in the PDF shifts from the body of the distribution to the tails, signifying a higher probability of extreme events. As α approaches 2, the mass of the PDF shifts from the tails of the distribution to the body, signifying a lower probability of extreme events.

When α < 2, Levy α-stable distributions have infinite variance; when α < 1, the distributions do not have a defined mean. Infinite variance means that the variance of a sample does not converge as the sample size approaches infinity (e.g., Schumer et al., 2001).

Closed-form expressions for Levy's α-stable densities do not exist except in three special cases: α = 2 (Gaussian), α = 1 (Cauchy), and α = ½ (Levy). Stable distributions are therefore described by their characteristic function, $\phi(k)$, which is the Fourier

transform of the PDF. The characteristic function is (using the S(α, β, γ, δ; 0) parameterization of Nolan, 2002):

$$\phi(k) = E\exp(ikX) = \exp\left(-\gamma^{\alpha}|k|^{\alpha}\left[1 + i\beta\left(\tan\frac{\pi\alpha}{2}\right)(\text{sign } k)\left(|\gamma k|^{1-\alpha} - 1\right)\right] + i\delta k \right) \tag{5}$$

where X is a random variable with density $f(x)$, $i = \sqrt{-1}$, and the sign function is:

$$\text{sign } k = \begin{cases} -1 & k < 0 \\ 0 & k = 0 \\ 1 & k > 0 \end{cases} \tag{6}$$

Note that the change of the variable x to k indicates transformation to Fourier space. The characteristic function for the case $\alpha = 1$ is slightly different and is omitted here.

1.3.2 The FADE

The FADE uses Levy's α-stable densities to overcome limitations of the ADE by allowing (1) growth in plume variance faster than $t^{½}$ and (2) plumes with non-Gaussian concentration profiles. In 1D, the FADE is (e.g., Benson, 1998):

$$\frac{\partial C}{\partial t} = -v\frac{\partial C}{\partial x} + D\left(p\frac{\partial^{\alpha} C}{\partial x^{\alpha}} + q\frac{\partial^{\alpha} C}{\partial(-x)^{\alpha}}\right) \tag{7}$$

where C is concentration, t is time, v is advective velocity, x is distance, D is a constant diffusion coefficient, and p and q are related to the skew of the Levy density. Note that (7) collapses into the traditional ADE when $\alpha = 2$ since $p + q = 1$ and $\frac{\partial^2 C}{\partial x^2} = \frac{\partial^2 C}{\partial(-x)^2}$.

The FADE assumes that particle jumps are governed by a PDF with infinite variance. These densities are governed by a generalization of the CLT called the

Generalized Central Limit Theorem (GCLT). We illustrate the GCLT by considering IID random variables, $x_1, x_2, x_3, \ldots x_n$, from a Levy α-stable distribution (e.g., Samorodnitsky and Taqqu, 1994):

$$\frac{x_1 + x_2 + x_3 \quad x_n - n\mu}{\sigma n^{1/\alpha}} = Z \tag{8}$$

where μ is the population mean, σ is the population standard deviation, n is the number of samples in the population, and Z is a Levy α-stable random variable to which the sums converge.

In terms of solute migration, if particle jumps are governed by an infinite variance PDF, the distribution of particles in space after a large number of jumps follows an infinite variance distribution and plumes grow like $t^{1/\alpha}$ (i.e., much faster than ADE growth) (Schumer et al., 2001). The FADE has successfully modeled non Gaussian plumes at the Cape Cod site using $\alpha = 1.8$ (Benson et al., 2000) and the MADE site in Columbus, MS, using $\alpha = 1.1$ (Benson et al., 2001).

1.4 K and the FADE

The FADE, then, (1) allows for super-Fickian dispersion and (2) by approximating contaminant plumes with a Levy α-stable distribution, can approximate plumes observed in the field that are non Gaussian (Benson et al., 2001). The obvious question, however, is what is the mechanism that gives rise to particle jumps that follow a Levy α-stable distribution? Recalling that particle jumps are related to velocity, equation (1) suggests spatial variabililty of hydraulic conductivity as a mechanism assuming that ∇h does not change the PDF of K. This work investigates whether or not spatial

variabililty of K could be a mechanism for heavy-tailed particle jumps by generating K fields with a Levy α-stable distribution of log K and comparing them to the PDF of the resultant velocity fields.

This work sets out to answer two questions:

(1) Does a Levy α-stable probability distribution of log K give rise to a Levy α-stable probability distribution of log velocity as suggested by equation (1)? To answer this question, we examine whether the log K and resulting log v distributions (a) exhibit power law tails and (b) can be fit to a stable distribution.

(2) What is the relationship between the tail parameter describing the velocity distribution, $\alpha_{\log v}$, and that describing the conductivity distribution, $\alpha_{\log K}$? If a given $\alpha_{\log K}$ results in a certain $\alpha_{\log v}$, we could measure $\alpha_{\log K}$ in the field and, pending work on the relationship between $\alpha_{\log v}$ and α, the order of the fractional dispersion derivative in the FADE, use it as an estimate of α. To answer this question, we use box plots and scatter plots to compare $\alpha_{\log K}$ and $\alpha_{\log v}$.

2.0 Previous Work

The application of Levy α-stable densities to hydraulic conductivity distributions has seen some use in the past. Painter (1996a; 1996b) and Liu and Molz (1997) have presented evidence that increments in log K are consistent with a stable distribution. Benson et al. (2001) has shown that increments in raw K are consistent with a shifted Pareto distribution (see equation 9) at the Macrodispersion Analysis site (MADE) in

Columbus, MS. Painter (2001) reexamines some of his earlier data sets and concludes that stable distribution in log K overestimates extreme values of K. Meerschaert et al. (2004) suggests an alternative Laplace or double exponential model for log K. Determination of the PDF of K remains an area of active research.

Herrick et al. (2003) investigated the same questions posed in this research using a shifted Pareto distribution as the PDF of K in conductivity fields. The shifted Pareto is a heavy-tailed distribution where:

$$P(X > x) = (x + s)^{-\alpha} \tag{9}$$

Their study examined K and v statistics over the ensemble of $n = 100$ K fields for two values of α_K: $\alpha_K = 1.1$ and $\alpha_K = 1.8$. They found that (1) Pareto distributions of K give rise to Pareto distributions of velocity (i.e., K is a mechanism for heavy-tailed velocity distributions) and (2) that heavy-tailed K fields give rise to lighter-tailed velocity fields, which they attribute to the inverse correlation between K and ∇h. The latter result is important because it would suggest that ∇h affects the PDF of v in equation (1).

This study seeks to expand on the work of Herrick et al. (2003) by examining the correlation between $\alpha_{\log K}$ and $\alpha_{\log v}$ on an individual, field-by-field basis in light of the complexity found by Herrick et al. (2003). In addition, this study uses a different PDF for K (i.e., log-stable instead of shifted Pareto) and a different algorithm to generate K fields. These differences allow us to evaluate whether the method of K field simulation affects the results.

3.0 Methods of Research

3.1 Introduction

It is not practical to study the relationship between $\alpha_{\log K}$ and $\alpha_{\log v}$ in a field-setting due to the large number of K and velocity fields that must be tested to achieve a statistically significant result. Therefore, we used a Monte Carlo method in which a computer algorithm generates many K field realizations; the corresponding velocity fields are then computed using the USGS finite difference computer code MODFLOW (McDonald and Harbaugh, 1988, 1996, 2000). Statistical analysis was done on the input and output from each Monte Carlo trial (i.e., estimation of $\alpha_{\log K}$ and $\alpha_{\log v}$) and comparisons between input statistics and output statistics were made. These steps are outlined in depth below.

3.2 Algorithm for K field generation

K fields were generated with an algorithm that is a modification of the method for spectral synthesis of fractional Brownian motion fields (fBm) (Saupe, 1988; Brewer and Wheatcraft, 1994). Spectral synthesis is based on the Fourier series, the discrete form of which is (e.g., Weaver, 1983):

$$f(x) = \sum_{k=0}^{\infty} \left[A_k \cos(2\pi\omega_k x) + B_k \sin(2\pi\omega_k x) \right] \tag{10}$$

where A_k and B_k are the pure cosine (real) and sine (imaginary) contents, respectively, ω_k is frequency, and k is the wave number (k = 0, 1, . . . N-1). The spectral synthesis method chooses the real and imaginary parts of the Fourier series, A_k and B_k, in Fourier space from a Gaussian probability distribution. The phase term in (10), $2\pi\omega_k x$, is chosen from a uniform distribution on [0, 2π]. Equation (10) is then multiplied by a filtering function:

$$\frac{1}{k^{2H+1}} \tag{11}$$

where H is the Hurst parameter. The Hurst parameter determines correlation characteristics of the fBm. If $H = 0.5$, increments in the fBm are independent. If $0 \le H < 0.5$, increments are negatively correlated (antipersistence), if $0.5 < H \le 1.0$, increments are positively correlated (persistence) (Samorodnitsky and Taqqu, 1994). The discrete inverse Fourier transform is then taken of the product of (10) and (11) to create a spatially-varying fBm with values that follow a Gaussian distribution. The fBm field is then log-transformed to avoid negative values of K.

The Saupe (1988) fBm algorithm applies a Gaussian pdf in Fourier space to generate the Fourier coefficients. The fBm is obtained by taking the inverse Fourier transform. Our plan was to modify this algorithm to produce fractional Levy motion (fLm) by replacing the Gaussian pdf with a Levy-stable pdf. However it's not that simple. It turns out that the Saupe algorithm only works with a Gaussian because the Gaussian is applied in Fourier space. When the inverse Fourier transform is applied, the real space fBm is correct because the Fourier transform of a Gaussian noise is another Gaussian noise.

As a result, our fLm algorithm is considerably different that the original Saupe (1988) and Brewer and Wheatcraft (1994) fBm algorithm. Our modification of the spectral synthesis method generates random numbers from a Levy α-stable distribution in the spatial domain, takes the Fourier transform to find the corresponding A_k and B_k in the Fourier domain, performs the filtering process using the same power law filter, and takes the inverse Fourier transform to convert back to the spatial domain. The resultant field is log-transformed to avoid negative values of K.

The modified algorithm's *K* fields were verified in two ways: first, a Rescaled Range (R/S) analysis (Mandelbrot and Wallis, 1969; Turcotte, 1997) of the resultant fields confirmed that they were Long Range Dependent (LRD) with respect to the Hurst parameter, and second, Mandelbrot plots (Mandelbrot, 1963) confirmed that increments in log *K* exhibit power law decay in the tails. The *K* fields simulated in this study are therefore LRD and exhibit a stable distribution of log *K* characterized by tail parameter $\alpha_{\log K}$.

3.3 K field properties

The *K* fields in this study are 3-dimensional and consisted of 665,640 nodes: 129 nodes in the *x* and *y* directions and 40 nodes in the *z* direction. The *K* fields are anisotropic, with correlation lengths in the *x* direction being 10 times greater than those in the *y* and *z* directions. A value of $H = 0.2$, in the domain of antipersistence, was applied in the *x*, *y*, and *z* directions. This Hurst parameter is consistent with values reported in the literature (e.g., Painter 1996a; 1996b). We generated 100 *K* fields for four values of $\alpha_{\log K}$ ($\alpha_{\log K} = 0.8$, 1.1, 1.4, and 1.7).

3.4 v field properties

A head gradient of $\nabla h = 0.003$, consistent with the gradient at the MADE site (Boggs et al., 1992) was applied to the flow model. The upgradient and downgradient edges of the flow model were set as constant head boundaries while the lateral faces were set as no flow boundaries (Figure 2).

The finite difference groundwater code MODFLOW (McDonald and Harbaugh, 1988; 1996; 2000) solved the steady state groundwater flow equation for the head field (e.g., Bear, 1972):

$$\frac{\partial}{\partial x}\left(-K_x\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(-K_y\frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(-K_z\frac{\partial h}{\partial z}\right)=0 \qquad (12)$$

Since the flow conditions are at steady state, and mass into a grid cell should equal the mass out of a cell, the solver minimizes the mass balance error in and out of the cells. The cell by cell flow terms are then extracted from the MODFLOW solution and converted from mass flux to velocity. Flow in the longitudinal direction (i.e., the component of flow exiting the Front Right Face of the MODFLOW solution) was analyzed in this study.

The global mass balance (i.e., mass error over the entire flow domain) for each Monte Carlo trial was calculated by MODFLOW, and a program was written to calculate the local mass balance (i.e., mass error in flow between cells). The Pre Conditioned Gradient solver (PCG2) gave fast, accurate solutions (defined as a global mass balance of 0 % and local mass balances that did not exceed 0.5%).

3.5 Tail Parameter Estimation and Statistical comparison

Mandelbrot plots and PP plots were used to establish that the log K and log v distributions were consistent with Levy α-stable distributions. The Nolan (1997) Maximum Likelihood Estimator was then used to obtain $\alpha_{\log K}$ and $\alpha_{\log v}$, the tail parameters describing the hydraulic conductivity and velocity distributions. Velocities from nodes within 15 nodes of the upgradient and downgradient constant head

boundaries were excluded from tail parameter estimation in order to avoid interference from constant head boundaries.

Finally, the results were plotted in scatter plots and box plots, and regression analysis was performed to discern a relationship between $\alpha_{\log K}$ and $\alpha_{\log \nu}$.

4.0 Results

4.1 Relationship between the K PDF and the ν PDF

Results of this research are presented as follows: First we outline the results of Mandelbrot plots of log *K* and log *ν* to determine if they exhibit power law decay in their tails, a characteristic of stable distributions (Samorodnitsky and Taqqu, 1994). Second, we outline the results of PP plots of log *K* and log *ν* to determine the goodness of fit to a stable distribution. We then present the results of the Nolan (1997) MLE estimator, used to estimate $\alpha_{\log K}$ and $\alpha_{\log \nu}$. Finally, the tail parameters describing the log *K* and log *ν* distributions are then compared on box plots and scatter plots in order to determine the relationship between the two.

4.1.1 Mandelbrot Plots

Mandelbrot plots establish whether or not the tails of a data set have tails that decay like a power law. Take, for example, a power law distribution:

$$\lim_{x\to\infty} P(X > x) = Cx^{-\alpha} \tag{13}$$

Mandelbrot (1963) points out that by taking logs of both sides of (13) we have:

$$\log P(X > x) = \log(C) - \alpha \log(x) \tag{14}$$

Therefore, a plot of $P(X>x)$ vs. x on log-log axes is a straight line of slope $-\alpha$ if there is some value in the ordered data set, D, above which distribution behaves like a power law.

Mandelbrot plots were made for all 400 K fields. Figure 3 shows Mandelbrot plots of a typical K field from the $\alpha_{\log K} = 1.1$ batch of Monte Carlo trials. Figure 3(a) is a Mandelbrot plot of increments in raw K. Raw K curves concave upward for extreme increments in K. This behavior is consistent with a log-stable distribution of raw K and is to be expected for a stable distribution of log K (Samorodonitsky and Taqqu, 1994). This behavior was evident in all Mandelbrot plots of K fields examined. In Figure 3(b), increments in log K plot in a straight line corresponding to $\alpha_{\log K} = 1.539$, thereby indicating power law behavior in log K. The K fields generated in this study; therefore, exhibit power law decay of log K in their tails.

Mandelbrot plots were also made for all 400 velocity fields. Figure 4 shows Mandelbrot plots of the velocity fields from the same batch of $\alpha_{\log K} = 1.1$ Monte Carlo trials. Figure 4(a) is a Mandelbrot plot of raw v. The extreme values of v curve concave upward from the straight line fit and the distribution is consistent with a log-stable distribution. Figure 4(b) is a Mandelbrot plot of log v and exhibits straight-line power law decay, indicating that log v is characterized by power law decay in the tails. This behavior was evident in all Mandelbrot plots of v fields examined. The v fields generated in this study, therefore, exhibit power law decay of log v in their tails. We remark that the straight line fits in Figures 3(a) and 4(a) for raw K and v might also appear adequate to some scientists. This illustrates the difficulty of tail estimation in general, and perhaps it explains why some investigators believe the raw field data has power law tails, while others believe that the log transformed data has a power law tail.

4.1.2 PP Plots

A PP plot graphically measures the goodness of fit of real data to a probability distribution. It consists of two data series: the cumulative probability vs. the expected cumulative probability for the observed data (circles in Figure 5) and the cumulative probability vs. the expected cumulative probability for the best fit probability distribution (solid line in Figure 5). Note that for the best fit probability distribution, cumulative probability equals expected cumulative probability, and the result is a straight line of slope = 1. PP plots were constructed using a program written by Nolan (1997) that (1) calculates the best fit stable parameters for the observed data set (i.e., log K or log v) and (2) constructs a plot of the data showing the fit between the distribution of the observed data and the best fit distribution suggested by the MLE.

Figure 5(a) is a PP plot of log K from a typical K field from the $\alpha_K = 0.8$ batch of Monte Carlo simulations. This plot indicates log K closely follows a stable distribution with $\alpha = 1.1$ and $\beta = 0.006$. Figure 5(b) shows a PP plot of log v from the v field corresponding to the K field in Figure 5(a). Log v is well-approximated by a stable distribution with $\alpha = 1.1$ and $\beta = 0.006$, but in the tails of the distribution the observed cumulative probability is greater than the expected cumulative probability, indicating that the tails of the log v distribution are heavier than the best fit stable law would predict. The extreme tails of the observed data do curve back to the best fit line for the most extreme fractiles, though. PP plots of log K and log v indicate that stable distributions of log K give rise to stable distributions of log v.

4.1.3 Mandelbrot and PP Plot Discussion

Mandelbrot plots show that log K distributions with power law decay in their tails give rise to log v distributions with power law decay in their tails. PP plots show that stable distributions of log K give rise to stable distributions of log v. Taken together, these results suggest that a stable distribution of hydraulic conductivity gives rise to a stable distribution of velocity, and ∇h in equation (1) does not affect the stability of v given a Levy α-stable PDF of K.

This study, combined with the work of Herrick et al. (2003), provides convincing evidence that satisfies a central assumption of the FADE—infinite variance distributions (i.e., stable or heavy-tailed) of conductivity give rise to infinite variance distributions (i.e., stable or heavy-tailed) of velocity. In addition, an infinite variance distribution of v given an infinite variance distribution of K is not contingent on K being approximated as stable or K being approximated as a power law—both types of infinite variance K distributions give rise to an infinite variance velocity distribution.

4.2 Relationship between $\alpha_{\log K}$ and $\alpha_{\log v}$

Now that it has been shown that stable log K fields give rise to stable log v fields, we can determine the relationship, if any, between $\alpha_{\log K}$ to $\alpha_{\log v}$, the tail parameters describing each distribution. Both box plots and scatter plots were used in this comparison.

4.2.1 Box Plots

Box plots compare $\alpha_{\log K}$ to $\alpha_{\log v}$ over the ensemble of Monte Carlo trials. Figure 6 is a box plot that shows the population statistics of $\alpha_{\log K}$ and $\alpha_{\log v}$ as estimated by the Nolan (1997) MLE. The *x* axis indicates the $\alpha_{\log K}$ of the Monte Carlo batch of experiments. The statistics from Figure 6 are presented in Table 1.

The interquartile range of values returned by the Nolan (1997) MLE is notably small, indicating high precision in the estimator. This plot shows that, over the ensemble, heavy-tailed hydraulic conductivity fields give rise to heavier-tailed velocity fields (i.e., $\alpha_{\log K} > \alpha_{\log v}$). This result is opposite of that reported in Herrick et al. (2003). This may be due to the fact that a different simulation methodology was used to generate our K fields, or perhaps because we used a different method for tail estimation. This issue is discussed further in section 5.

4.2.2 Scatter Plots

Scatter plots are used to obtain an individual, trial by trial comparison between $\alpha_{\log K}$ and $\alpha_{\log v}$. Figure 7 shows a scatter plot of $\alpha_{\log K}$ vs. $\alpha_{\log v}$ for the 400 *K* and velocity fields in this study. This plot indicates that $\alpha_{\log K}$ and $\alpha_{\log v}$ are positively correlated and, as indicated by the slope of the regression line, exhibit a near 1:1 linear relationship. The empirical relationship between $\alpha_{\log K}$ and $\alpha_{\log v}$ determined by the least squares regression line is:

$$\alpha_{\log v} = 1.167\left(\alpha_{\log K}\right) - 0.6426 \tag{15}$$

Therefore, if we had a field-measured value of $\alpha_{\log K}$, we could calculate the resulting $\alpha_{\log v}$ by using equation (15). However, a different relation was observed in the work of Herrick et al. (2003).

4.2.3 Conclusions for Relationship

This study finds that heavy-tailed *K* fields give rise to heavier-tailed velocity fields, a different result than that of Herrick et al. (2003). Scatter plots show that $\alpha_{\log K}$ and $\alpha_{\log v}$ are positively correlated and linearly related. A field measured value of $\alpha_{\log K}$ could be used as an estimate of $\alpha_{\log v}$ using equation (15). However, we believe that additional research is necessary to validate the relation between these two parameters. For example, the slope of the linear relation may depend on the Hurst coefficient of the hydraulic conductivity field.

5.0 Discussion

This research has two conclusions. First, a stable log *K* distribution gives rise to a stable log *v* distribution. Second, an empirical relationship between $\alpha_{\log K}$ and $\alpha_{\log v}$, given in equation (15), was found. Therefore, $\alpha_{\log K}$ can be used as an *a priori* estimator of $\alpha_{\log v}$.

The discrepancy between the work of Herrick et al. (2003) and this work concerning the relationship between the tail parameter describing the *K* distribution and that describing the *v* distribution begs the question, what is the true relationship between the two? *A priori*, before examining the work in this study or that of Herrick et al. (2003), it may be tempting to conclude that the tail parameter describing the velocity distribution would be identical to that describing the conductivity distribution (i.e., $\alpha_{\log K} = \alpha_{\log v}$ or $\alpha_K = \alpha_v$). However, this study found that heavy-tailed conductivity fields give

rise to heavier-tailed velocity fields (i.e., $\alpha_{\log K} > \alpha_{\log v}$), which is the opposite the work of Herrick et al. (2003), who found that heavy-tailed conductivity fields give rise to lighter-tailed velocity fields (i.e., $\alpha_K < \alpha_v$).

One possible explanation for this difference is the method used in each study to generate *K* fields. Herrick et al. (2003) simulates raw *K* as a shifted Pareto distribution and this study simulates raw *K* as a log-stable distribution. A significant difference between the two distributions is that a log-stable distribution has more extreme values relative to the population mean than a heavy-tailed distribution (Samorodonitsky and Taqqu, 1994). Therefore, the *K* fields generated in this study are characterized by values of *K* that are much larger in magnitude than those in Herrick et al. (2003) relative to the population mean. Since these *K* values are spatially correlated, they tend to occur in the same location.

The result of highly correlated large *K* values is a field with several 'channels' of high *K*. Figure 8 shows a sample *K* field from the $\alpha_{\log K} = 1.7$ batch of Monte Carlo experiments with several 'channels' of high *K*. The resultant velocity field also exhibits these channels. These channels become preferential flowpaths for water similar to those that would be found in a fluvial geologic setting.

The effect of channeling has been studied in other work. Trefry et al. (2003) simulates *K* fields and studies the velocity fields and contaminant dispersion. Raw *K* is simulated as log normal. As the variance in log *K*, $\sigma^2_{\ln K}$, increases, the velocity fields exhibit exponential to power law tails. Essentially, Trefry et al. (2003) reports an amplification of the velocity field caused by large values of *K* relative to the population mean that are more common as $\sigma^2_{\ln K}$ becomes large. In this study, that amplification

would be manifest as a lower $\alpha_{\log v}$, indicating more extreme velocities relative to the population mean.

In conclusion, whether one approximates raw *K* as stable or log *K* as stable affects the relationship between the tail parameter describing the *K* distribution and that describing the *v* distribution. An important question, therefore, is which distribution, stable or log stable, is a better approximation of *K* in a typical field setting? Work by Painter (2001) suggests that approximating raw *K* as stable is more appropriate since log stable distributions overestimate extreme values of *K*. The work of Trefry (2003), however, suggests that approximating log *K* as stable is more appropriate to create the preferential flow paths observed in typical field settings. Meerschaert et al. (2004) suggest an alternative Laplace model for log *K*, which is consistent with the assumption that *K* has power law probability tails. Another question concerns the role of the Hurst parameter. Can the Hurst parameter be manipulated to generate *K* fields consistent with the PDF of *K* suggested by Painter (2001) and the channeling of *K* detailed by Trefry et al. (2003)? These questions deserve attention in future research.

Notes

This research was supported by NSF grant DMS-0139927 and a grant from the University of Nevada at Reno Graduate Student Association (GSA).

Works Cited

Adams, E. E. and L. W. Gelhar, 1992, Field study of dispersion in a heterogeneous

aquifer 2: spatial moment analysis, *Water Resources Research*, **28** (12), 3293 – 3307.

Bear, J., 1972, Dynamics of Fluids in Porous Media, Dover, New York, 764 p.

Benson, D. A. *The fractional advection-dispersion equation: development and application*. Unpublished Ph.D. dissertation. University of Nevada Reno, 1998.

Benson, D. A., S. W. Wheatcraft, and M. M. Meerschaert, 2000, Application of a fractional advection dispersion equation, *Water Resources Research*, **36**(6), 1403 – 1412.

Benson, D. A., R. Schumer, M. M. Meerschaert, and S. W. Wheatcraft, 2001, Fractional dispersion, Levy motion, and the MADE tracer tests, *Transport in Porous Media*, **42**, 211 – 240.

Boggs, J. M., S. C. Young, L. M. Beard, L. W. Gelhar, K. R. Rehfeldt, and E. E. Adams, 1992, Field study of dispersion in a heterogeneous aquifer 1: overview and site description, *Water Resources Research*, **28** (12), 3281 – 3291.

Brewer, K., and S.W. Wheatcraft, 1994, Multi-scale reconstruction of hydraulic conductivity distributions using wavelet transforms. in:*Wavelets in Geophysics*, ed.: E. Foufoula-Georgiou and P. Kumar, Academic Press, New York, 1994, p. 213 – 248.

Dagan, G., 1984, Solute transport in heterogeneous porous formations, *Journal of Fluid Mechanics*, **145**, 151 – 177.

Freeze, R. A. and J. Cherry, *Groundwater*, Prentice Hall, Englewood Cliffs, NJ: 1979.

Freyberg, D. L., 1986, A natural gradient experiment on solute transport in a sand aquifer 2: spatial moments and the advection and dispersion of nonreactive tracers, *Water Resources Research*, **22** (13), 2017 – 2030.

Gelhar, L. W., A. L. Gutjahr, and R. L. Naff, 1979, Stochastic analysis of

macrodispersion in a stratified aquifer. *Water Resources Research*. **15** (6), 1387 – 1397.

Gelhar, L. W. and C. L. Axness, 1983, Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resources Research*, **19** (1), 161 – 180.

Herrick (2003)

Hewett, T. A., 1986, Fractal distributions of reservoir heterogeneity and their influence on fluid transport, *Proceedings of the 61st Annual Technical Conference of the Society of Petroleum Engineers*, Report No. 15386.

Liu, H. H. and F. J. Molz, 1997, Comment on 'Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations' by Scott Painter, *Water Resources Research*, **33**(4), 907 - 908.

Mandelbrot, B., The variation of certain speculative prices, 1963, *Journal of Buisness*. **36**, 394 – 419.

Mandelbrot, B. and J. R. Wallis, 1969, Computer experiments with fractional Gaussian noises, Part 2: Rescaled Range and spectra, *Water Resources Research*, **5** (1), 242 – 259.

McDonald, M. G. and A. W. Harbaugh, 1988, A modular three-dimensional finite-difference ground-water flow model, *U.S. Geological Survey Techniques of Water Resources Investigations*, book 6, chapter A1, 586 p.

Meerschaert, M. M., T. J. Kozubowski, F. J. Molz and S. Lu, 2004, Fractional Laplace Model for Hydraulic Conductivity, *Geophysical Research Letters*, **31(**8), 4 pp., doi:10.1029/2003GL019320

Mercado, 1967, The spreading pattern of injected water in a permeability-stratified

aquifer, *Proceedings of the International Association of Science Hydrology Symposium*, Haifa Publication 72, Israel.

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

Neuman, S. P., 1990, Universal scaling of hydraulic conductivities and dispersivities in Geologic media, *Water Resources Research*, **26** (8), 1749 – 1758.

Nolan, J., Numerical calculation of stable densities and distribution functions, *Communications of Statistical Stochastic Models*, **13**(4), pp.759 – 774, 1997.

Nolan, J. P., 2002, *Stable Distributions: Models for Heavy Tailed Data*, Birkhauser, Boston, 352 p.

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

Painter, S., 1996b, Stochastic interpolation of aquifer properties using fractional Levy motion, *Water Resources Research*, **32** (5), 1323 - 1332.

Painter, S., 2001, Flexible scaling model for use in random field simulation of hydraulic conductivity, *Water Resources Research*, **37** (5), 1155 – 1163.

Pickens and Grisak

Samorodonitsky, G. and M. S. Taqqu, *Stable non-Gaussian Random Processes: stochastic models with infinite variance (stochastic modeling).* Chapman and Hall, New York, 1994.

Saupe, D., 1988, Algorithms for random fractals, in: The Science of Fractal Images, eds.: Peitgen, H. O. and Saupe, D., 71 – 136. Springer Verlag, New York, 312 pp.

Schwartz, F. W., 1977, Macroscopic dispersion in porous media: the controlling factors,

Water Resources Research, **13** (4), 743 – 752.

Schumer, R., D. A. Benson, S. W. Wheatcraft, M. M. Meerschaert, 2001, Eulerian derivation of the fractional advection dispersion equation, *Journal of Contaminant Hydrology*, **48** (1/2), 69 – 88.

Smith, L., and F. W. Schwartz, 1980, Mass transport 1: a stochastic analysis of macroscopic dispersion, *Water Resources Research*, **16** (2), 303 – 313.

Smith, L., and F. W. Schwartz, 1981a, Mass transport 2: analysis of uncertainty in prediction, *Water Resources Research*, **17** (2), 351 – 369.

Smith, L., and F. W. Schwartz, 1981b, Mass transport 3: role of hydraulic conductivity data in prediction, *Water Resources Research*, **17** (5), 1463 – 1479.

Trefry, M. G., F. P. Ruan, and D. McLaughlin, 2003, Numerical simulations of preasymptotic transport in heterogeneous porous media: Departures from the Gaussian limit, *Water Resources Research*, **39** (3), 10-1 – 10-13.

Turcotte, D. L. *Fractals and Chaos in Geology and Geophysics*, 2nd ed., Cambridge University Press, Cambridge, 1997.

Wheatcraft, S. W. and S. W. Tyler, 1988, An explanation of scale dependent dispersivity in heterogeneous aquifers using concepts of fractal geometry, *Water Resources Research*, **24** (4), 566 - 578.

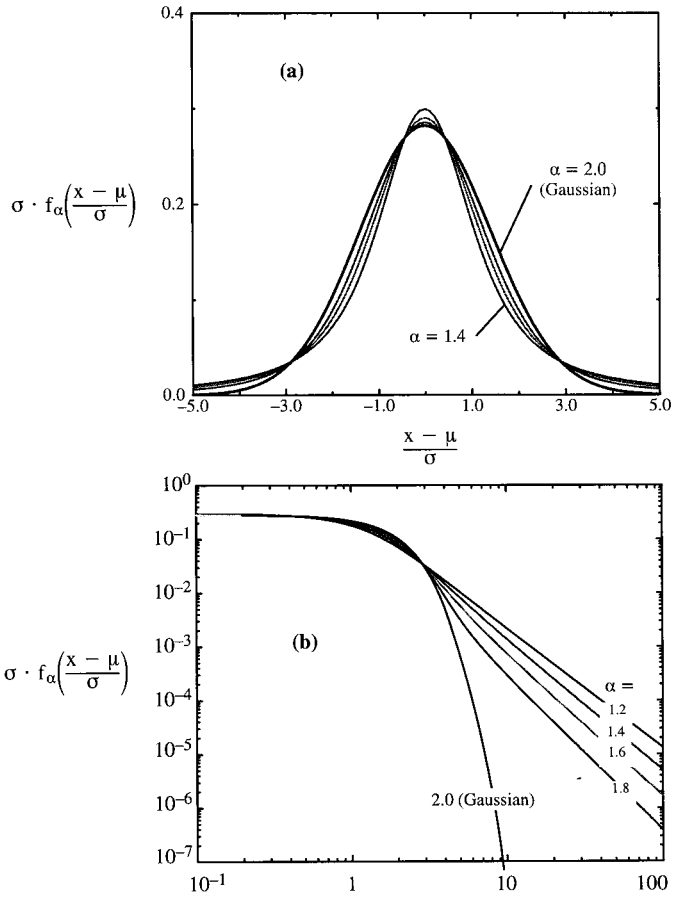


FIGURE 1: Plots of symmetric stable densities. Shown above are plots of the probability density function (PDF) on arithmetic axes. Below are plots on log-log axes showing power law decay of the tails. Note that the case of $\alpha = 2$ is the Gaussian. When $\alpha < 2$, the distribution 'allows' for more extreme events. Taken from Benson (1998).

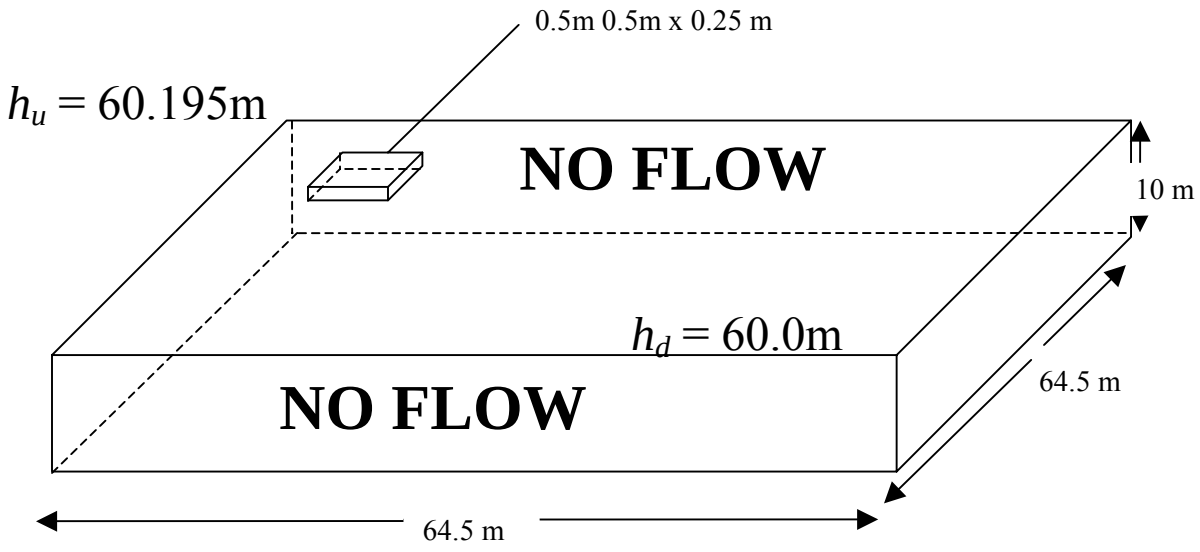
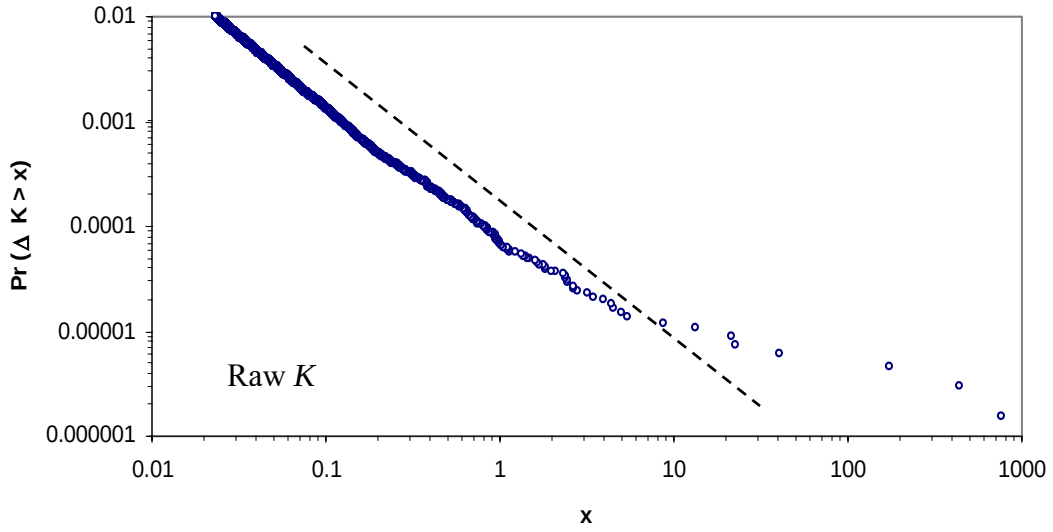


FIGURE 2. Conceptual rendering of the flow model used in this study. The domain is 64.5m x 64.5m x 10m consisting of 665,640 nodes of dimension 0.5m x 0.5m x 0.25m. An upgradient head, h_u, of 60.195 m and a downgradient head, h_d, of 60.0 m are constant head boundaries. The lateral and bottom face of the flow domain are no-flow boundaries.

(a)



(b)

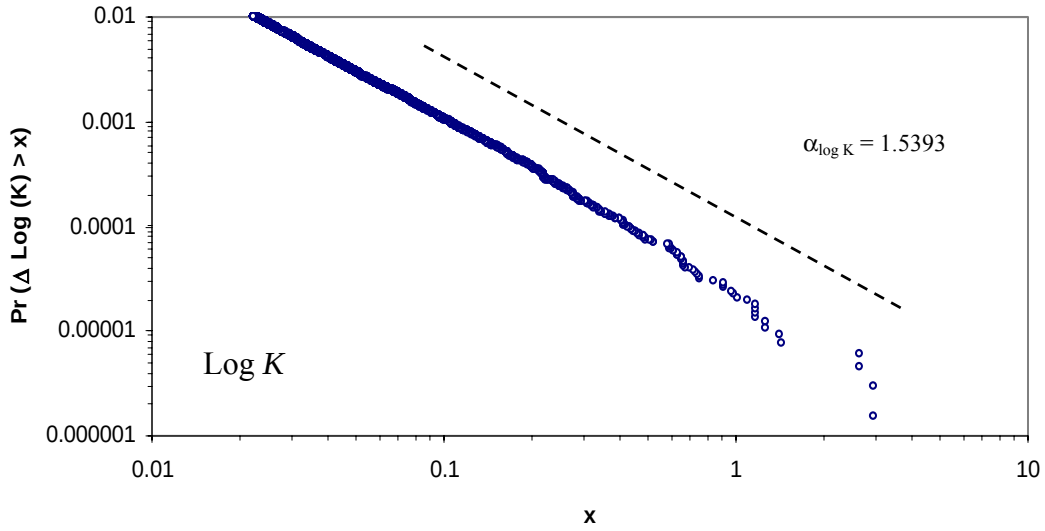
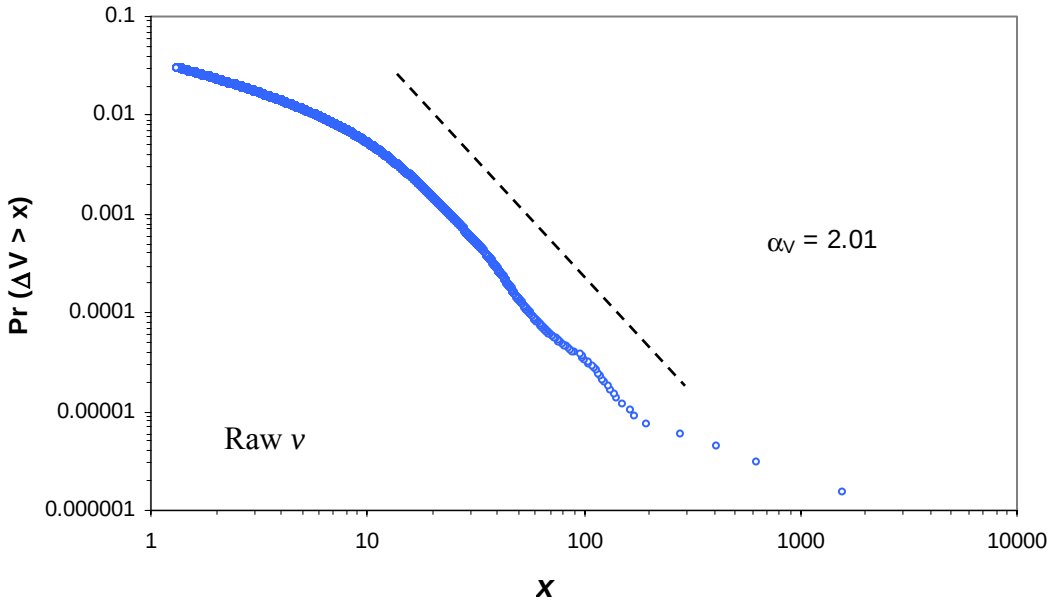


FIGURE 3. (a) Mandelbrot plot of raw K for a realization with $\alpha_{\log K} = 1.1$, $H = 0.2$, and anisotropy $K_x = 10$, $K_y = K_z = 1$. The concave upward curvature of the tail indicates that raw K is consistent with a log stable distribution. (b) Mandelbrot plot of the same realization where log K is shown. The power law decay of the upper tail is consistent with a stable distribution of log K. These plots confirm that raw K is consistent with a log stable distribution and log K is consistent with a stable distribution, as is expected.

(a)



(b)

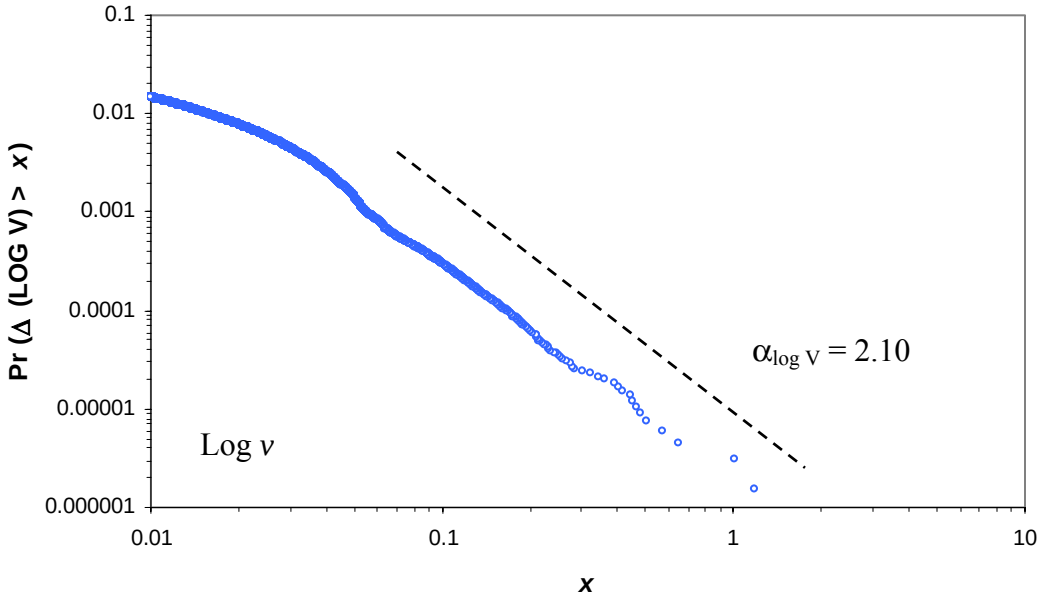


FIGURE 4. (a) Mandelbrot plot of raw *v* from the $\alpha_{\log K} = 1.1$ Monte Carlo simulations. The concave upward curvature of the tail indicates that raw *v* is is consistent with a log stable distribution.. (b) Mandelbrot plot of log *v* from the $\alpha_{\log K} = 1.1$ Monte Carlo simulations. The power law decay of the upper tail is consistent with a stable distribution of log *v*.

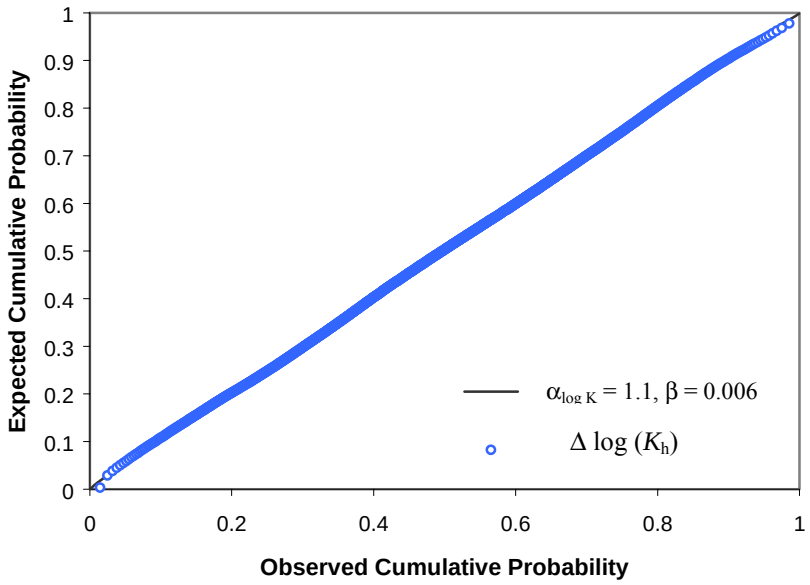


FIGURE 5a. PP plot of increments in log *K* (circles) against a best fit stable distribution with $\alpha = 1.1$ and $\beta = 0.006$ (solid black line) suggested by the Nolan (1997) MLE. Log *K* is well-fit by a stable distribution. This data is from a Monte Carlo trial with $\alpha_{\log K} = 0.8$. N = 1000 data points.

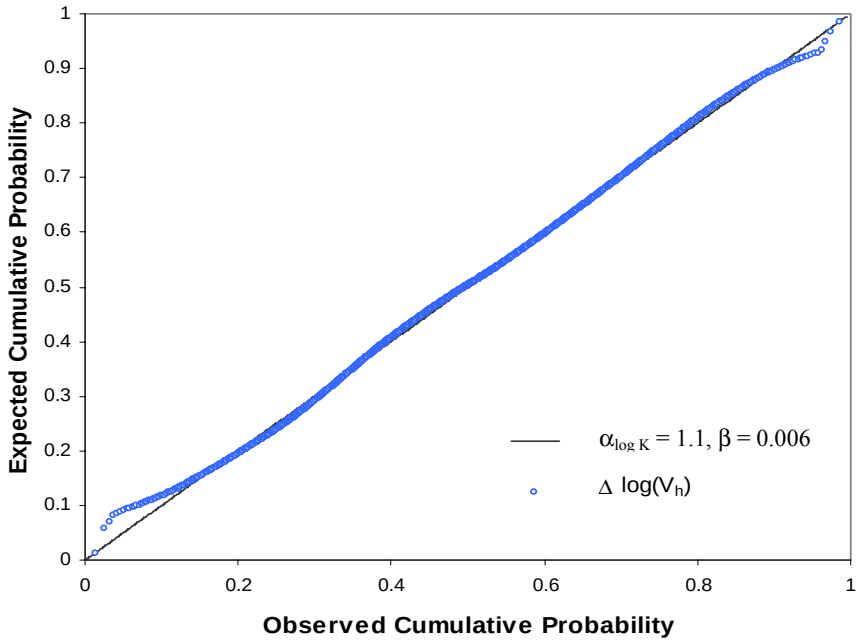


FIGURE 5b. PP plot of increments in log *v* (circles) against a best fit stable distribution with $\alpha = 0.97$ and $\beta = -0.079$ (solid black line) suggested by the Nolan (1997) MLE. Log *v* is well-approximated by a stable distribution. This data is from a Monte Carlo trial with $\alpha_{\log K} = 0.8$. N = 1000 data points.

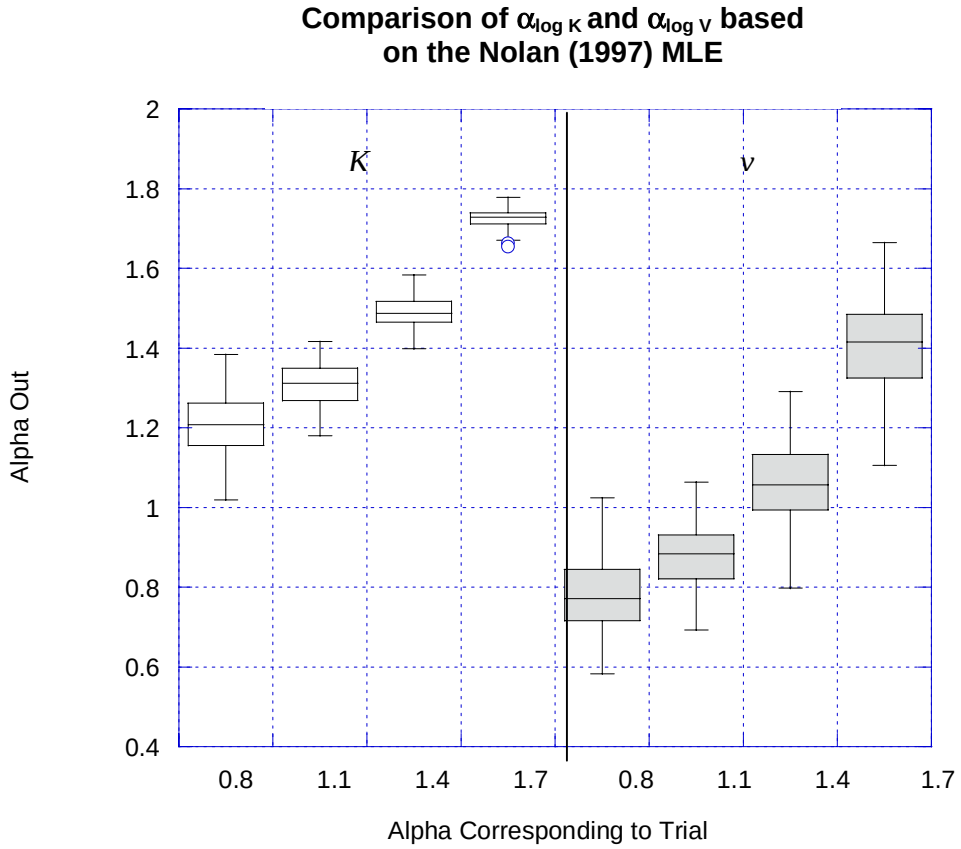


Figure 6. Box plot showing results of the Nolan (1997) MLE. This estimator predicts that heavy-tailed hydraulic conductivity fields characterized by tail parameter $\alpha_{log\ K}$ give rise to heavier-tailed velocity fields with a lower tail parameter $\alpha_{log\ V}$.

TABLE 1. Results of Nolan (1997) Maximum Likelihood Estimator

	K_h				V_h			
	$\alpha_{log\ K}$ = 0.8	$\alpha_{log\ K}$ = 1.1	$\alpha_{log\ K}$ = 1.4	$\alpha_{log\ K}$ = 1.7	$\alpha_{log\ K}$ = 0.8	$\alpha_{log\ K}$ = 1.1	$\alpha_{log\ K}$ = 1.4	$\alpha_{log\ K}$ = 1.7
Maximum	1.38	1.42	1.58	1.78	1.02	1.06	1.29	1.67
Upper 25th	1.26	1.35	1.52	1.74	0.84	0.93	1.13	1.48
Median	1.21	1.31	1.49	1.73	0.77	0.89	1.06	1.42
Lower 25th	1.16	1.27	1.47	1.71	0.72	0.82	1.00	1.33
Minimum	1.02	1.18	1.40	1.66	0.58	0.69	0.80	1.11
Int. Q. Range	.011	0.08	0.05	0.03	0.13	0.11	0.14	0.16

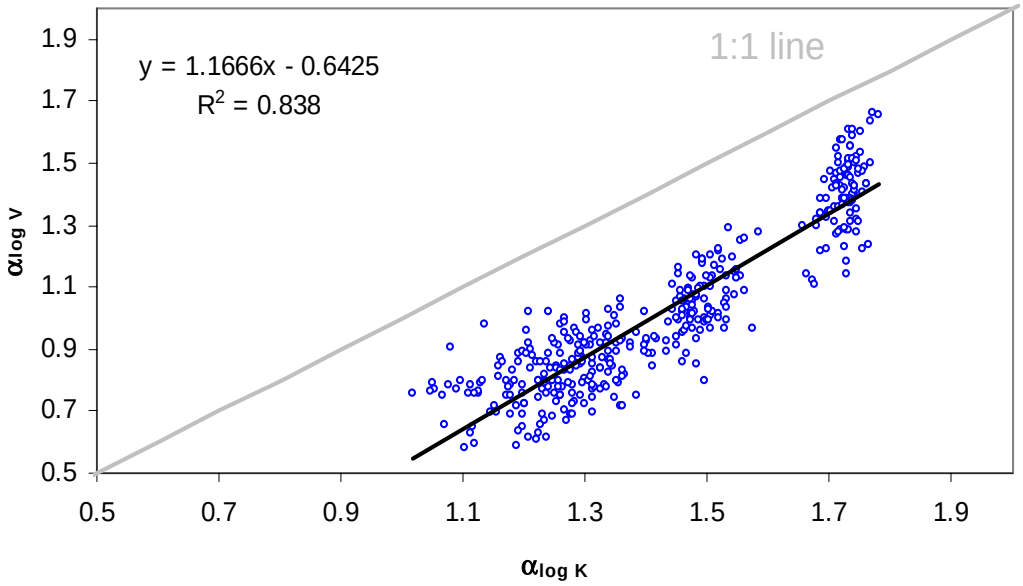


FIGURE 7. Scatter plot of α_K vs. α_V based on the results of the Nolan (1997) MLE. The equation for the regression line (black) is given in the upper left hand corner of the plot, and the 1:1 fit is shown in gray. This estimator predicts that the relationship between $\alpha_{\log K}$ and $\alpha_{\log V}$ is almost 1:1. This plot is based on N = 400 Monte Carlo realizations of *K*.