Hydraulic conductivity, velocity, and the order of the fractional dispersion derivative in a highly heterogeneous system

Matt G. Herrick,[1,2] David A. Benson,[1] Mark M. Meerschaert,[3] and Katherine R. McCall,[4]

Received 4 September 2001; revised 17 April 2002; accepted 17 April 2002; published 12 November 2002.

[1] A one-dimensional, fractional order, advection-dispersion equation accurately models the movement of the core of the tritium plume at the highly heterogeneous MADE site. An a priori estimate of the parameters in that equation, including the order of the fractional dispersion derivative, was based on the assumption that the observed power law (heavy) tail of the hydraulic conductivity (K) field would create a similarly distributed velocity field. Monte Carlo simulations were performed to test this hypothesis. Results from the Monte Carlo analysis show that heavy tailed K fields do give rise to heavy tailed velocity fields; however, the exponent of the power law (the tail parameter) describing these two distributions is not necessarily the same. The tail parameter that characterizes a velocity distribution is not solely dependent on the tail parameter that characterizes the K distribution. The K field must also have long-range dependence so that water may flow through relatively continuous high-K channels. *INDEX TERMS:* 1829 Hydrology: Groundwater hydrology; 1869 Hydrology: Stochastic processes; 3250 Mathematical Geophysics: Fractals and multifractals; *KEYWORDS:* fractional derivative, Lévy motion, tail estimation

Citation: Herrick, M. G., D. A. Benson, M. M. Meerschaert, and K. R. McCall, Hydraulic conductivity, velocity, and the order of the fractional dispersion derivative in a highly heterogeneous system, *Water Resour. Res.*, *38*(11), 1227, doi:10.1029/2001WR000914, 2002.

1. Introduction

[2] A fractional order advection-dispersion equation (ADE) describing the transport of solutes in a porous medium has been applied to several tracer tests [*Benson*, 1998; *Benson et al.*, 2000, 2001; *Meerschaert et al.*, 2001]. In this model, differential advection is described by a dispersion derivative of fractional order $0 < \alpha \leq 2$. The one-dimensional (1-D) form of the fractional ADE is

$$\frac{\partial C}{\partial t} = -V\frac{\partial C}{\partial x} + D\left(p\frac{\partial^{\alpha} C}{\partial x^{\alpha}} + q\frac{\partial^{\alpha} C}{\partial(-x)^{\alpha}}\right), \qquad (1)$$

where C is concentration, t is time, V is the mean drift velocity, x is the spatial domain, $D > 0$ is the dispersion coefficient, and the relative weight of the positive constants p and q, with $p + q = 1$, describes the skewness. *Schumer et al.* [2001] argue that skewness should be maximal, so that either p or q will be zero. For the MADE site plumes, *Benson et al.* [2001] use an equation with $p = 1$ and $q = 0$, leaving a single dispersion term. *Baeumer et al.* [2001] argue that the entire advection-dispersion operator should be of fractional order. For light-tailed (finite variance) random motions like the Gaussian or log normal, $\alpha = 2$ and the classical ADE is recovered. The classical ADE predicts a Gaussian plume with a Fickian growth rate of the centered second moment (or variance) that is proportional to time. To match field data, the Gaussian model requires a dispersion coefficient that grows with scale [*Pickens and Grisak*, 1981; *Dagan*, 1988; *Neuman and Zhang*, 1990; *Rajaram and Gelhar*, 1995]. For heavy tailed (infinite variance) random motions, $\alpha < 2$, and the measured plume variance grows more rapidly (with $t^{2/\alpha}$) [*Benson et al.*, 2001]. Because of the fractional order dispersion derivative, equation (1) predicts super-Fickian growth with a constant dispersion coefficient. In addition, when $p \neq q$, the plume is skewed, a feature Gaussian models do not allow.

[3] Equation (1) is the governing equation of Lévy motion, which is similar to Brownian motion except that incremental motions are heavy tailed (defined in detail below). In order to use equation (1) without prior observation of tracer movement, the parameters must be estimated based on aquifer properties. First, the aquifer hydraulic conductivity (K) field must be heavy tailed with infinite correlation length. Given this situation, one might assume: (1) that the velocity field has a similar distribution. From a heavy-tailed velocity field, one might further assume that: (2) particle "jumps" in small amounts of time are also heavy tailed with similar distribution. These two assumptions allow simple estimation of all parameters in equation (1) [*Benson et al.*, 2001].

[4] The focus of the present research is to test the first assumption and to discern a relationship between the tail parameters that describe the K and velocity distributions and the possible fractional order of the dispersion derivative. *Neuman* [1995] made analytic arguments that the velocity distribution should be Gaussian within a finite-variance K field with infinite correlation lengths. We are not aware of analytic methods for heavy-tailed K fields, so we take a numerical approach. Large-scale flow simulations

[1]Division of Hydrologic Sciences, Desert Research Institute, Reno, Nevada, USA.
[2]Now at Broadbent and Associates, Inc., Reno, Nevada, USA.
[3]Department of Mathematics, University of Nevada, Reno, Nevada, USA.
[4]Department of Physics, University of Nevada, Reno, Nevada, USA.

0043-1397/02/2001WR000914$09.00

in a generic hydrogeologic system with K statistics similar to the MADE site were completed. Due to the variability in K from simulation to simulation, a Monte-Carlo analysis was performed to obtain an ensemble average of many simulations.

2. Stable Distributions and Heavy Tails

[5] Stable distributions are solutions to the fractional ADE for an instantaneous injection of solute, i.e., the Green function. The properties of the stable distributions are most easily investigated in terms of their characteristic function $\varphi(k)$, which is the Fourier transform of the probability density function, i.e., $\varphi(k) = \int_{-\infty}^{\infty} \exp(ikx) f(x) dx$. The characteristic function of a symmetric stable distribution with no shift (zero mean for $\alpha > 1$) is, $\varphi(k) = \exp(-|k|^{\alpha}\sigma^{\alpha})$where the parameter σ describes the spread of the data or plume, and the characteristic exponent α defines the scaling properties and tail thickness of the distributions. Smaller values of α result in heavier tails. When $\alpha < 2$, the distribution has infinite variance and when $\alpha \leq 1$, it has an undefined mean as well. The α-stable densities are characterized by their "heavy" tails since, for $\alpha < 2$, one or both of the tails of the density are asymptotically power laws and plot as straight lines with slope -1-α on a log-log plot [see, e.g., *Benson et al.*, 2001].

[6] We loosely define a "heavy tailed" distribution as one in which either the left or right tail follows this power law relationship. For large values, the heavy tailed random variable Y approximately has the probability $P(|Y| \geq y) \approx Wy^{-\alpha}$, where $P(\cdot)$ denotes probability and W is a positive constant. For heavy tailed distributions, the exponent of the power law (α) is termed the tail parameter. *Lévy*'s [1937] extended central limit theorem dictates that for $\alpha < 2$, these heavy tailed random numbers are in the domain of attraction of (i.e., properly normalized sums converge to) a stable distribution with infinite variance, while for $\alpha \geq 2$ they are in the domain of attraction of a normal distribution [see *Feller*, 1971]. It is this convergence property that allows us to examine a generic heavy tailed K distribution.

3. MADE Site

[7] The Macrodispersion Experiments (MADE) were conducted at the Columbus Air Force Base in northeastern Mississippi in a highly heterogeneous aquifer. A tail parameter estimator developed by *Meerschaert and Scheffler* [1998] was used to estimate the tail thickness of the vertical increments of the K data collected at the site. The flowmeter measurements were collected on approximately 15 cm intervals [*Rehfeldt et al.*, 1992; *Boggs et al.*, 1993]. The increments (ΔK) data yielded a tail parameter of $\alpha = 1.1$ [*Benson et al.*, 2001].

[8] *Benson et al.* [2001] then assumed that the fluctuations in the regional gradient were small compared to the fluctuations in K, and that the two were relatively uncorrelated, so that the velocity (v) increments would be constant multiples of the K increments. This enabled simple a priori estimation of both α and D in equation (1). They argued that this assumption should be best when K fluctuations were very large; however, the validity of this assumption is untested. Figure 1 shows the longitudinal distribution of tritium mass during the MADE-2 tests along the central "core" of the plume for MADE-2 snapshots 3 and 4. The core was obtained by taking the maximum tritium concentration for a given set of samplers some distance from the injection point along the direction of mean trajectory. Also plotted are the α-stable solutions to the fractional ADE (1) and the Gaussian solutions to the traditional ADE. All parameters in equation (1) were estimated a priori using the statistics of the K field and the assumption that the v distribution is similar to that of the K field. The value of the dispersion coefficient is a constant in the fractional ADE for all time periods. The traditional ADE was modeled with a "best fit" dispersion coefficient for each time period, in the same manner as *Adams and Gelhar* [1992]. The fractional ADE is able to model the peak concentration and the highly skewed movement of the core of the plume. Because the stable solutions have heavy tails, the data were also plotted using semilog and log-log axes (Figures 1b and 1c). For brevity, only the last two snapshots are shown in the plots. The later time snapshots show that the transport is dominated by heavy tailed (power law) motions.

[9] The first conjecture used to estimate the parameter α is that heavy-tailed K distributions, $P(K \geq y) \approx Wy^{-\alpha}$, give rise to heavy-tailed velocity distributions with a similar tail parameter. This assumes that fluctuations from the mean hydraulic gradient divided by the porosity are negligible and uncorrelated with the K fluctuations, so that Darcy's law $v = -K(\nabla h/n)$, where ∇h is the regional hydraulic gradient and n is the porosity, is locally valid. This assumption, and the substitution $J = -\nabla h$, leads to the velocity distribution $P(v \geq y) \approx W(ny/J)^{-\alpha}$ with the same tail parameter as the K distribution. The second conjecture is that heavy tailed velocity distributions give rise to heavy tailed particle jumps with a similar tail parameter. This assumes that the particles have a well-defined mean duration of each motion (i.e., the temporal correlation of particle velocity is small), and allows the Markov property to be used in the derivation of equation (1). The focus of this research was to test the first conjecture; further particle tracking simulations can check the second.

4. Monte Carlo Simulations

[10] Two batches of Monte Carlo simulations (100 simulations per batch) were conducted to determine whether the velocity field has a power law tail and to compare the tail parameters that describe the hydraulic conductivity (α_K) and velocity (α_v) fields. The first batch was run with $\alpha_K = 1.1$ (characterizing a highly heterogeneous system) and the second batch with $\alpha_K = 1.8$ (characterizing a mildly heterogeneous system). The $\alpha_K = 1.1$ value was chosen to mimic the K field at the MADE site [*Benson et al.*, 2001]. The $\alpha_K = 1.8$ value was chosen to test the same assumption, but on a system that is not so heterogeneous. The two batches of simulations may also indicate a relationship between the K and velocity field tail parameters.

4.1. Heavy Tailed *K* Fields With Long-Range Dependence

[11] The generation of the K fields described below is a modified version of the p-field simulation approach described by *Goovaerts* [1997] and *Deutsch and Journel*

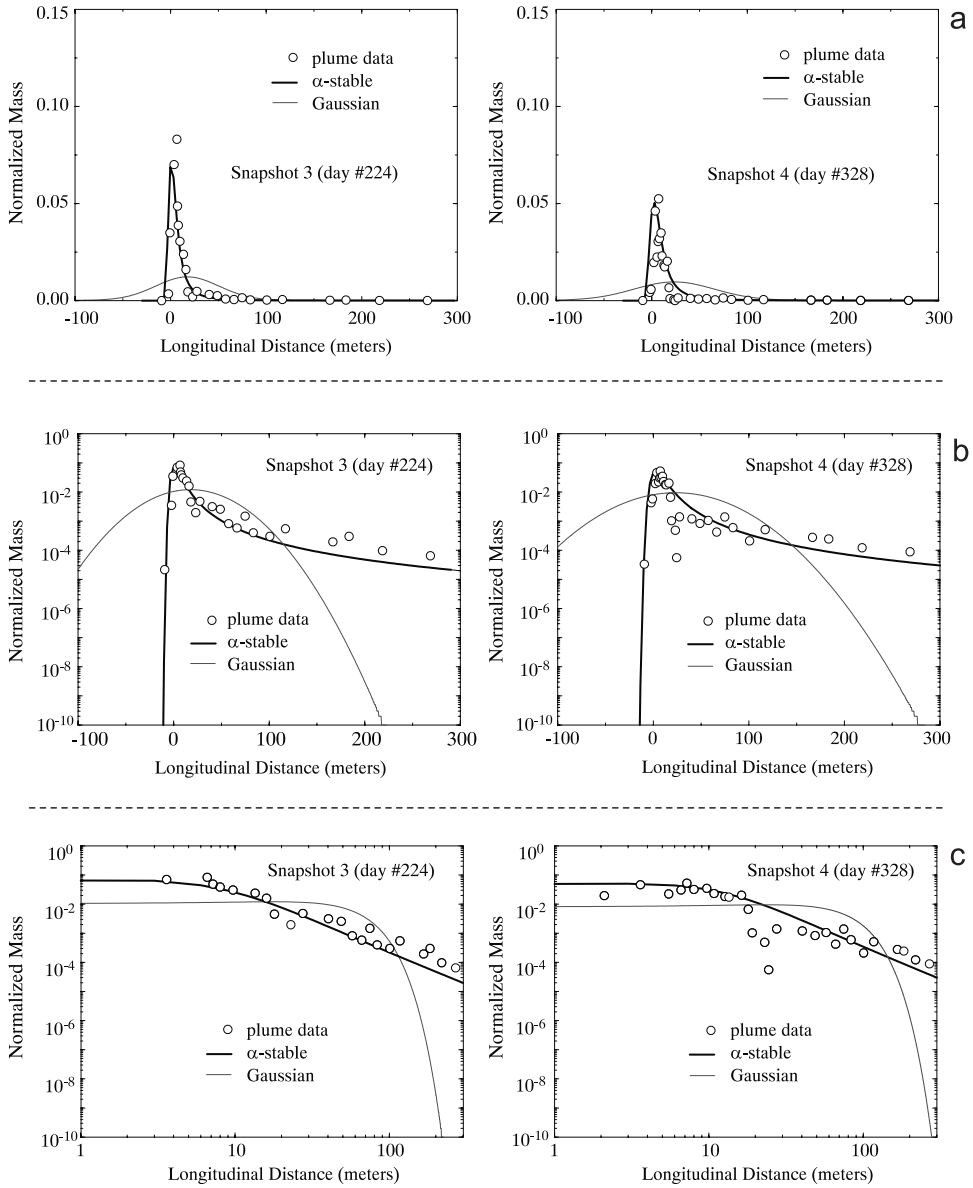

Figure 1. Plots of the MADE-2 normalized longitudinal tritium mass distribution from the final two snapshots (3 and 4) along with analytic solutions of the classical ADE and the fractional ADE using (a) linear axes, (b) semilog axes, and (c) log-log axes [after *Benson et al.*, 2001].

[1998]. With this method, a correlated random field is generated based on a certain distribution function. The values are ranked and replaced by similarly ranked values drawn from a different distribution [*Painter*, 2001]. Since the fractional ADE implies infinite velocity correlation lengths [*Benson*, 1998], we created K fields with both the long-range dependence (LRD) and a heavy tailed distribution. For a detailed description of various random processes with long-range dependence and high variability, see *Taqqu* [1987]. A fractional Brownian motion (fBm) algorithm [*Painter*, 2001] was used to generate the LRD portion of the K fields. A Hurst coefficient of $H = 0.40$ in agreement with *Molz and Liu*'s [1997] study of the vertical variation of K at the MADE site was used in the fBm algorithm. Anisotropy was then introduced by specifying that the correlation extended 10 times farther in the horizontal versus vertical direction.

[12] In general agreement with the ΔK data from the MADE site [*Benson et al.*, 2001], we used the shifted Pareto distribution function to define the underlying distribution of the K fields:

$$F(K) = 1 - W \cdot (K - s)^{-\alpha}, \tag{2}$$

where $\alpha > 0$, W is the Pareto scale factor ($W > 0$), s is the shift ($s < 0$), and $K \geq 0$. A shift of $s = -10^{-2}$ cm/s, resulting in a scale factor of $W = (-s)^{\alpha} = 10^{-2\alpha}$ (cm/s)$^{\alpha}$, was used in the simulations. The shifted Pareto was chosen because the numbers falling within the distribution are positive, heavy tailed, in the domain of attraction of a stable, and fit the ΔK data at the MADE site.

[13] The heavy tailed random numbers that defined the K field were truncated at 300 cm/s. We chose this extremely

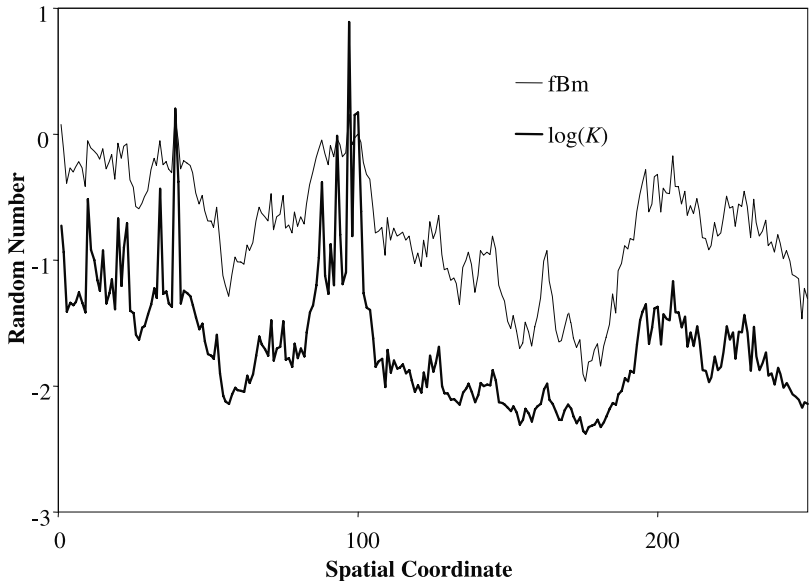

Figure 2. 1-D realization of an fBm field and resulting LRD heavy tailed K field ($H = 0.4$). For a direct comparison, the log(K) field is displayed.

large value for K to maintain the power law for most of the high-K tail. Our objective was to look for similarity in the K and velocity distributions from a theoretical standpoint. It is unknown how reasonable the cutoff is, since over 500,000 K values are generated in these simulations, and real aquifers are never sampled in such detail.

[14] The final step in the p-field algorithm is to map the Pareto-distributed random numbers into the fBm field, i.e., the highest numbers in the fBm are replaced by the highest numbers generated using the shifted Pareto (2) and so on. This maintains the spatial correlations built into the fBm algorithm, but changes the underlying distribution from a Gaussian to a heavy tailed distribution. Figure 2 shows a 1-D realization of a fBm field and resulting LRD heavy tailed K field generated. Although log(K) is shown, it is important to remember that the untransformed K values, not log(K), are heavy-tailed. For a more detailed description of the generation of the K fields, see *Herrick* [2001].

4.2. Flow Model

[15] A 3-D computational domain in the shape of a rectangular box was used. The size of the box was 320 m by 160 m by 10 m, in general agreement with the aquifer studied at the MADE site [*Boggs et al.*, 1993]. The model domain was broken up into 512,000 total cells of size 2 m by 2 m by 0.25 m. Constant head values were maintained on the up-stream and down-stream faces of the domain, with no-flow conditions enforced on the remaining four lateral faces. The constant head values induce an average head gradient of 0.005, also in general agreement with the MADE site. In addition to the anisotropy established in the fBm algorithm, anisotropy in the magnitude of K was introduced by setting the ratio of vertical to horizontal K at 0.1.

[16] The numerical code MODFLOW [*McDonald and Harbaugh*, 1988; *Harbaugh and McDonald*, 1996] was used to compute the steady state groundwater flow equation. The logarithmic mean was used to calculate the inter-block K [*Goode and Appel*, 1992]. Convergence of the PCG2 solver [*Hill*, 1990] was checked by calculating the local mass balance at steady state for each cell in the model domain. Local mass balance errors of less than 0.5% were found in all cells.

5. Estimation of Tail Thickness

[17] The K values were generated with a known α; however, we compared the α that describes the increments of the data due to the question of stationarity of the K fields that are based, in part, on a fBm. The tail parameter that represents the increments of a data set is a good approximation of the tail parameter that represents the actual values for a nonstationary data set [*Davis and Resnick*, 1985]. Performing tail parameter estimates on both the K and velocity data sets also ensures that any bias in a given estimator is applied to both data sets. Thus, a direct comparison can be made between α_K and α_v.

[18] The generation of the K fields was set up so that flow would predominantly be in the horizontal directions. Because the K fields are isotropic in the horizontal direction, only one tail parameter estimate of the vertical increments of hydraulic conductivity (ΔK) was needed. The resulting velocity increments may not be isotropic, so we examined the vertical increments of both the longitudinal (Δv_L) and transverse (Δv_T) velocities. The longitudinal direction of flow was assumed to be perpendicular to the constant head boundary faces, whereas the transverse direction of flow was assumed to be parallel to the constant head boundary faces. This is a valid assumption when averaging the entire model domain, but does not necessarily represent longitudinal and transverse flows for any individual cell. A determination of the exact longitudinal and transverse flows for each individual cell would be numerically intensive.

[19] There are numerous tail estimators, however, a robust estimator that accurately estimates α for all data sets does not exist. Three different numeric estimators and a graphical estimator were employed in this study to examine the K and velocity fields.

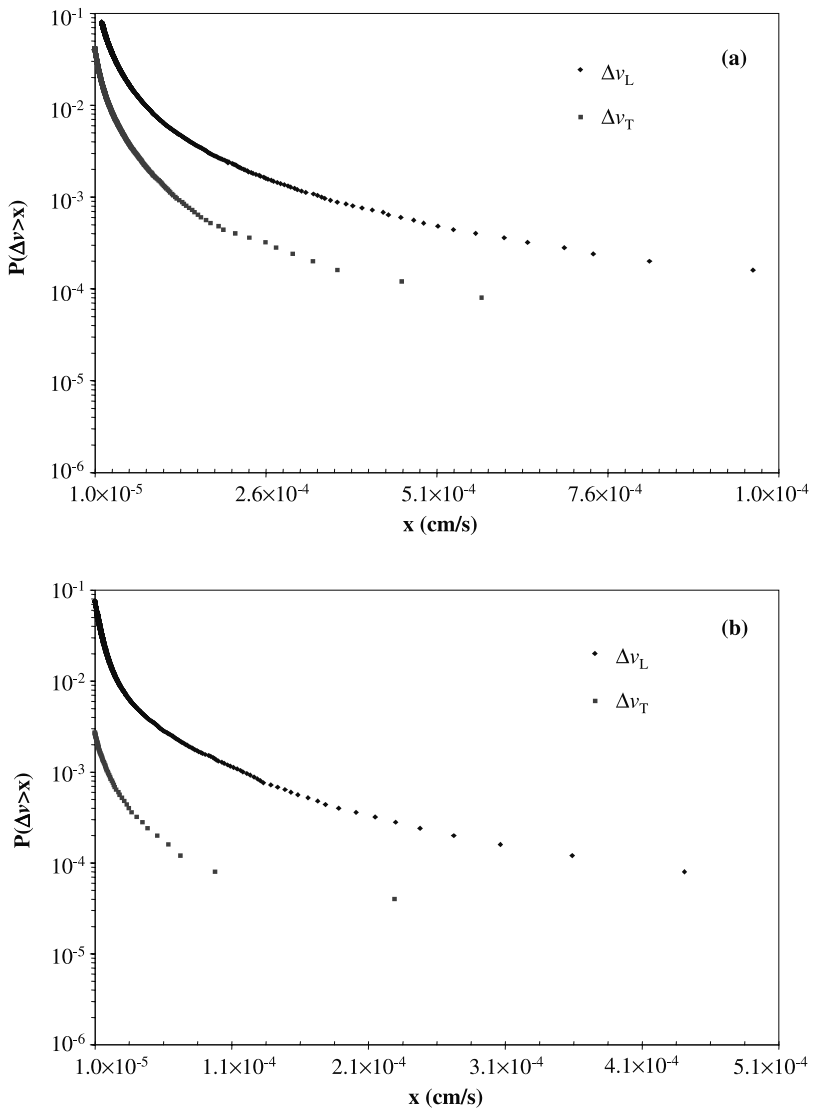

Figure 3. Semilog plots of the empirical distribution function for Δv_L and Δv_T for a single realization from the highly heterogeneous (a) and mildly heterogeneous (b) simulations.

[20] A log-log plot of the empirical distribution function is often used to estimate α from a data set. This graphical method was pioneered by *Mandelbrot* [1963] in his work on heavy tails, and will be called the Mandelbrot plot. If the random data y follow a Pareto distribution $F(y) = P(Y > y) = Wy^{\alpha}$, then $\log F(y) = \log C - \alpha \log y$. By sorting the data in descending order to obtain the order statistics $Y_1 \geq Y_2 \geq \ldots \geq Y_n$, we have approximately that $y = Y_r$ when $F(y) = r/n$. A Mandelbrot plot of $\log(Y_r)$ versus $\log(r/n)$ should be approximately linear with slope $-\alpha$ for heavy tailed data. Simple linear regression can be used to estimate α from a Mandelbrot plot, but this method can be affected by a few extreme values and typically performs poorly [*Benson et al.*, 2001]. A semilog plot of the empirical distribution function of the Δv_L and Δv_T for a single realization from both the highly heterogeneous and mildly heterogeneous simulations are displayed in Figures 3a and 3b. If the data were exponentially distributed, they would plot as a straight line in semilog space. The fact that the data are concave upward for both realizations indicates that the velocity increments are heavier tailed than an exponential distribution. Log-log (Mandelbrot) plots of the same data are shown in Figures 4a and 4b. Straight lines of known slopes are also plotted. From the highly heterogeneous realization, visual inspection indicates that the tail parameters describing Δv fall somewhere between 1.1 and 1.6 depending upon what percentage of the data is examined. Visual inspection from the mildly heterogeneous realization indicates that the tail parameters describing Δv are slightly larger and between 1.8 and 2.0. The magnitude of Δv in the highly heterogeneous realization is a bit larger than Δv in the mildly heterogeneous realization. This is expected since the probability of drawing large K values is greater when α_K is smaller. Mandelbrot plots for both realizations examined also indicate that the slopes defining Δv_L and Δv_T appear to be approximately the same, indicating that α is the same in both directions. Other cases considered exhibit the same pattern from both batches of realizations.

[21] Perhaps the most widely used tail estimator is *Hill*'s [1975] estimator, the conditional maximum likelihood esti-

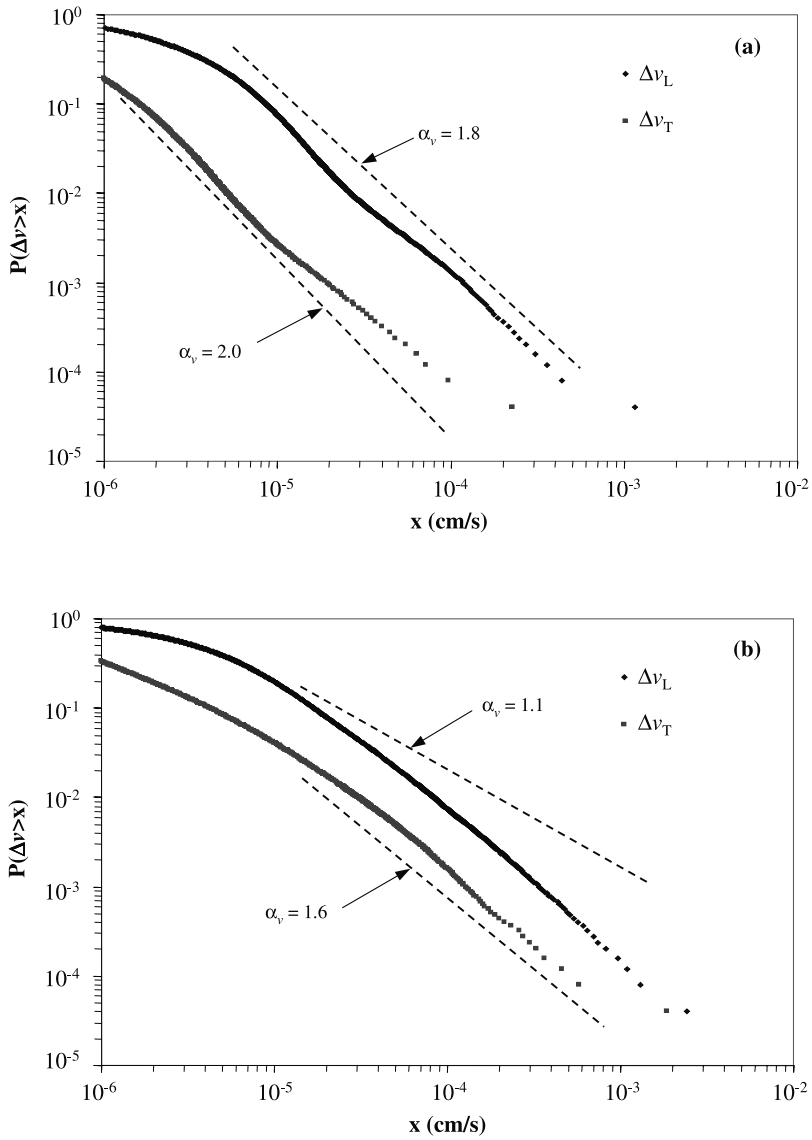

Figure 4. Log-log plots (Mandelbrot plots) of the empirical distribution functions for Δv_L and Δv_T from the highly heterogeneous (a) and mildly heterogeneous (b) simulations examined in Figure 3. Also plotted are dashed lines with known slope α.

mation (MLE) for a Pareto distribution, based on the r largest order statistics. When applying Hill's estimator, one should choose r as large as possible, but small enough so that the largest order statistics lie within the portion of the distributional tail where this approximation is valid. Hill's estimator can give misleading results when applied to stable data [*McCulloch*, 1997]. It has a significant bias for $1.5 < \alpha < 2$ and can yield estimates of α much greater than 2 when the true α is less than 2 [*Meerschaert and Scheffler*, 1998]. Despite its drawbacks, Hill's estimator was used in this study because of its popularity in statistical analysis of heavy tailed data sets.

[22] *Aban and Meerschaert* [2001] recently developed a modification of Hill's estimator. Hill's estimator is scale-invariant, in that a multiplicative factor in the data will not affect the estimate. However, it is not shift-invariant, since an additive factor will distort the estimates. The modified Hill's estimator, called shifted Hill's estimator, is both shift and scale invariant. Shifted Hill's estimator is the conditional MLE for a shifted Pareto distribution (2). As with Hill's estimator, one must choose a certain percentage r of the distribution tail where the approximation is valid. Shifted Hill's estimator exhibits significantly more variability than the traditional Hill's estimator, due to the fact that we are estimating an additional parameter, the shift. The advantage to using shifted Hill's estimator is that, given an optimal shift, shifted Hill's estimator is able to better estimate α when $1.5 < \alpha < 2$ [*Aban and Meerschaert*, 2001].

[23] Figure 5 shows Hill's and shifted Hill's estimators for Δv_L as a function of the highest order statistics for three realizations from the highly heterogeneous batch of simulations. The three realizations represent approximately the minimum, median, and maximum α estimation for the longitudinal velocity increments. All three realizations are stable over the range representing the largest 2% to 5% of the data using shifted Hill's estimator. The same three realizations are not as stable for Hill's estimator indicating that the data have a nonzero shift. The Δv_T for these realizations as

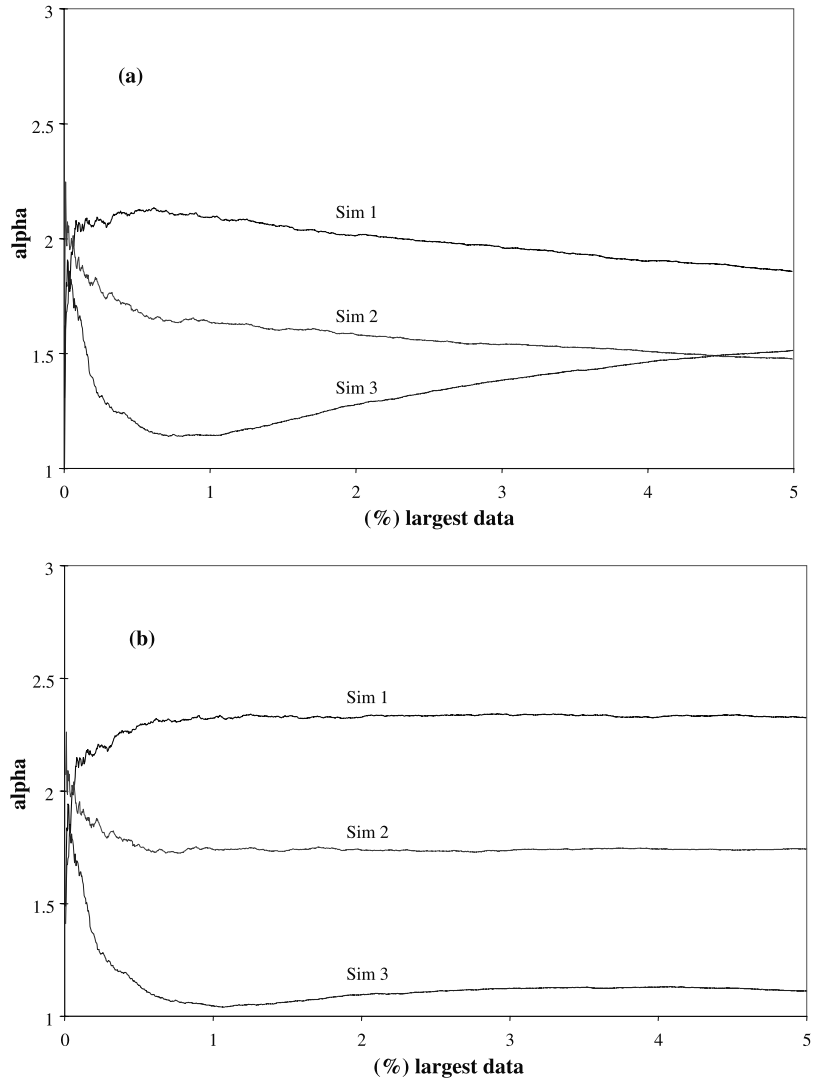

Figure 5. Hill's (a) and shifted Hill's (b) estimation of α as a function of the highest order statistics.

Table 1. Summary Statistics of α Using Meerschaert/Scheffler (M/S), Hill's, and Shifted Hill's (S-Hill's) Estimators Based on 100 Realizations From the Highly Heterogeneous (α_K = 1.1) and Mildly Heterogeneous (α_K = 1.8) Simulations[a]

	Highly Heterogeneous Field (α_K = 1.1)								
	ΔK			Δv_L			Δv_T		
Statistic	M/S	Hill's	S-Hill's	M/S	Hill's	S-Hill's	M/S	Hill's	S-Hill's
Minimum	0.839	0.767	0.730	1.147	1.139	1.114	1.067	1.067	0.950
Lower quartile	0.870	0.842	0.900	1.311	1.427	1.590	1.226	1.303	1.413
Median	0.888	0.890	0.943	1.378	1.587	1.740	1.276	1.442	1.604
Upper quartile	0.898	0.923	0.976	1.440	1.773	1.934	1.345	1.588	1.822
Maximum	0.923	1.104	1.139	1.579	2.678	3.433	1.512	2.228	3.346
	Mildly Heterogeneous Field (α_K = 1.8)								
	ΔK			Δv_L			Δv_T		
Statistic	M/S	Hill's	S-Hill's	M/S	Hill's	S-Hill's	M/S	Hill's	S-Hill's
Minimum	0.815	1.054	0.941	1.305	1.426	1.528	1.239	1.341	1.537
Lower quartile	1.017	1.185	1.328	1.497	1.881	2.176	1.412	1.782	2.094
Median	1.069	1.267	1.423	1.556	2.098	2.451	1.489	1.937	2.428
Upper quartile	1.109	1.328	1.501	1.613	2.325	2.965	1.538	2.129	2.916
Maximum	1.227	1.697	1.923	1.750	3.094	5.719	1.684	2.910	5.473

[a] Results for Hill's and shifted Hill's estimators are for the top 5% of the data.

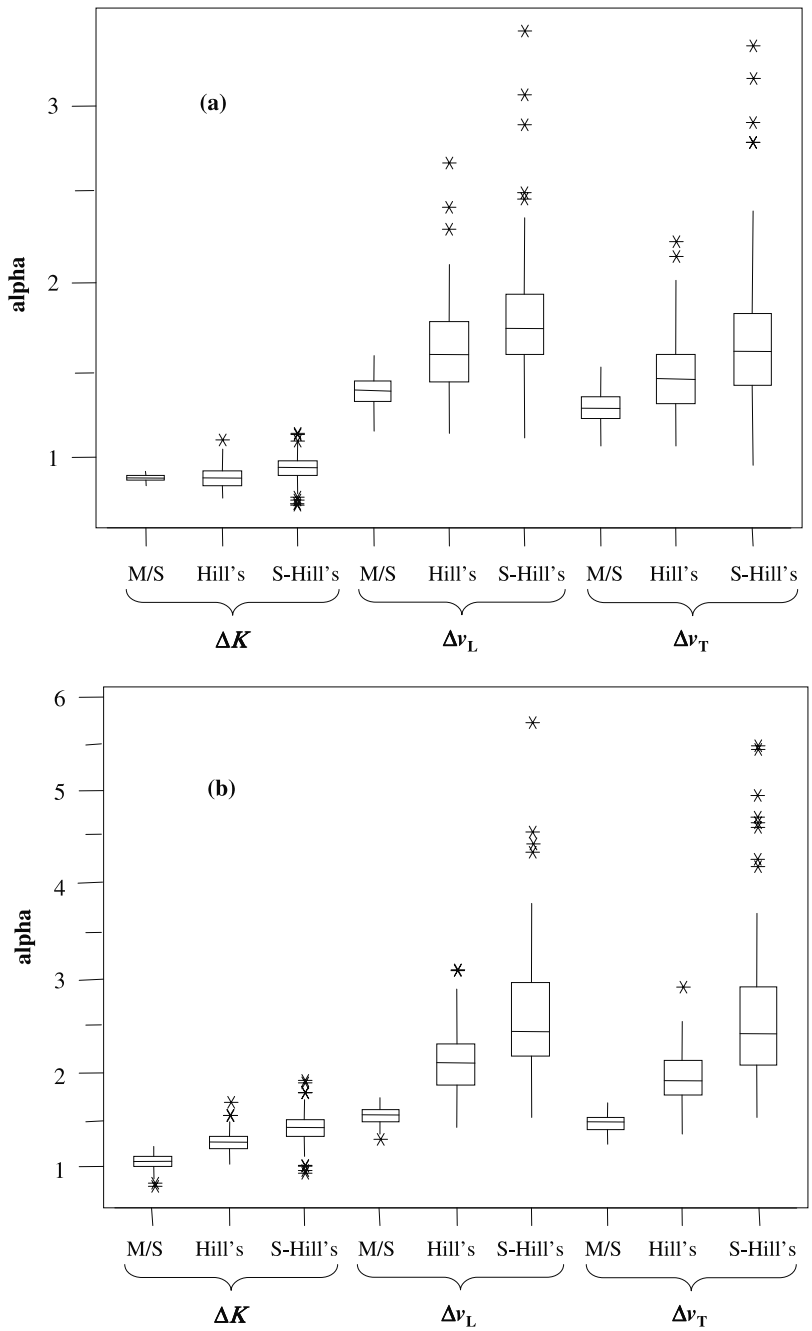

Figure 6. Box plot of α using M/S, Hill's and shifted Hill's estimators, based on 100 realizations from the highly heterogeneous (a) and mildly heterogeneous (b) simulations. Results for Hill's and Shifted Hill's estimators are for the largest 5% of the data. The box for each estimator represents the upper and lower quartiles (Q3 and Q1) and the median (Q2). The fences on either side of the box are of length 1.5(Q3-Q1) and the asterisks represent outliers.

well as realizations from the mildly heterogeneous batch of simulations exhibit similar patterns in regards to stability.

[24] *Meerschaert and Scheffler* [1998] developed a method for estimating the thickness of heavy tails based on the asymptotics of sums. This robust estimator depends only on α, and not on the exact form of the distribution. The estimator even works for dependent data. *Meerschaert and Scheffler*'s [1998] estimator, like the shifted Hill's estimator, performs quite well when α approaches 2. The primary advantage of the Meerschaert and Scheffler estimator is that it uses all of the data to estimate α, unlike Hill's and shifted Hill's estimators that are based on a percentage of the largest order statistics.

[25] Summary statistics of Meerschaert and Scheffler (M/S), Hill's, and shifted Hill's estimators for 100 realizations from both the highly heterogeneous and mildly heterogeneous simulations are presented in Table 1 and Figures 6a and 6b. The primary result is that the tail

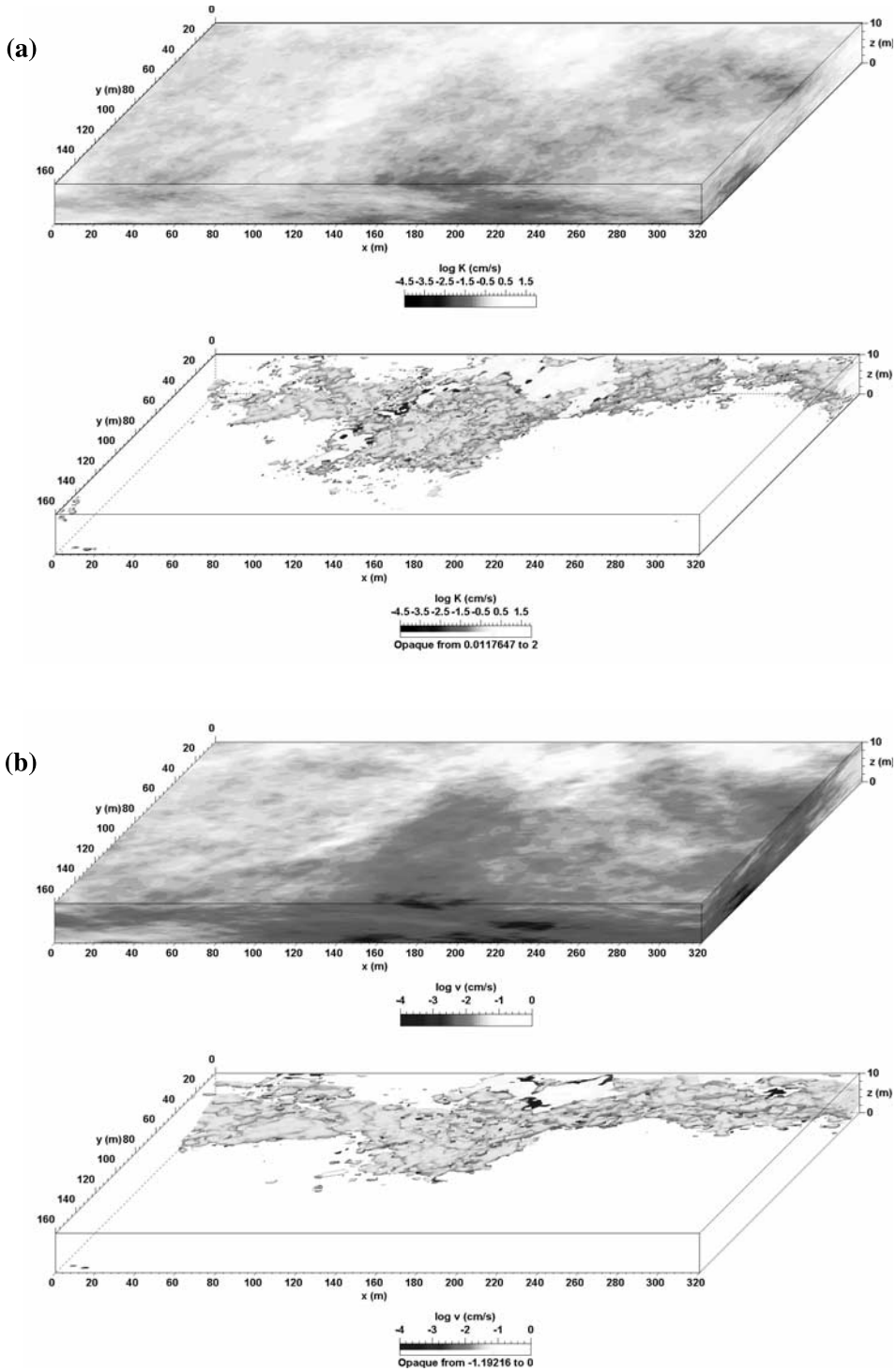

Figure 7. A realization of the K field (a) and resulting velocity field (b) for (Sim 75). This realization consistently gave small α_v values for all estimators.

parameters that describe the vertical velocity increments are statistically significantly larger than the tail parameters that describe the vertical K increments. The statistical significance is shown by the nonoverlap of the box plots. Another notable result is that α is significantly smaller for ΔK than for the K distribution itself. We attribute this to the vertical anticorrelation given to the K field by the fBm with a Hurst coefficient of 0.4. The anticorrelation tends to place small and large K values next to each other in the vertical direction, leading to a higher propensity for large increments.

[26] It is interesting to note that tail parameter estimations for the mildly heterogeneous case are much more broadly distributed than for the highly heterogeneous case. This is probably due to the fact that tail parameter estimation becomes more difficult as α approaches 2. The spread of the sampling distribution for the three estimators is different, however, they all predict the same trend. All estimators

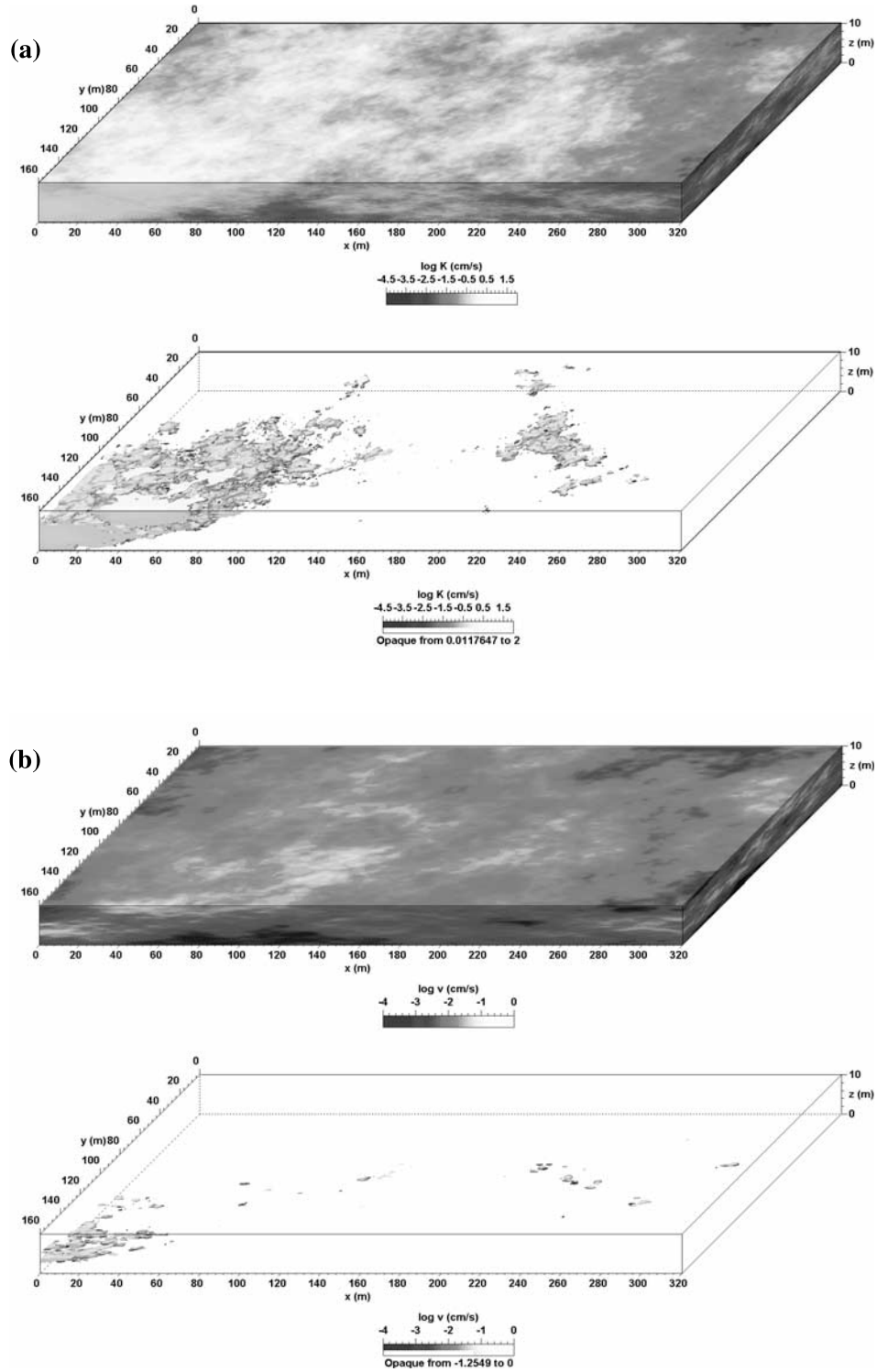

Figure 8. A realization of the K field (a) and resulting velocity field (b) for (Sim 80). This realization consistently gave large α_v values for all estimators.

show that α is slightly smaller but not statistically different in the transverse direction than the longitudinal direction.

6. Correlation of K and Velocity Fields

[27] We now turn our attention to the possible factors that contribute to the increase in α for v versus K. The correlation structure of the K fields was based on the LRD built into the fBm algorithm. LRD simply means that the correlation between K values decreases as a power law of the distance between sample locations. This gives rise to infinite correlation lengths [*Taqqu*, 1987]. In this section, the correlation structure of two realizations from the highly heterogeneous K field ($\alpha_K = 1.1$) is visually examined. The two realizations chosen represent end-members of the spectrum for tail parameter estimation, even though they were created with

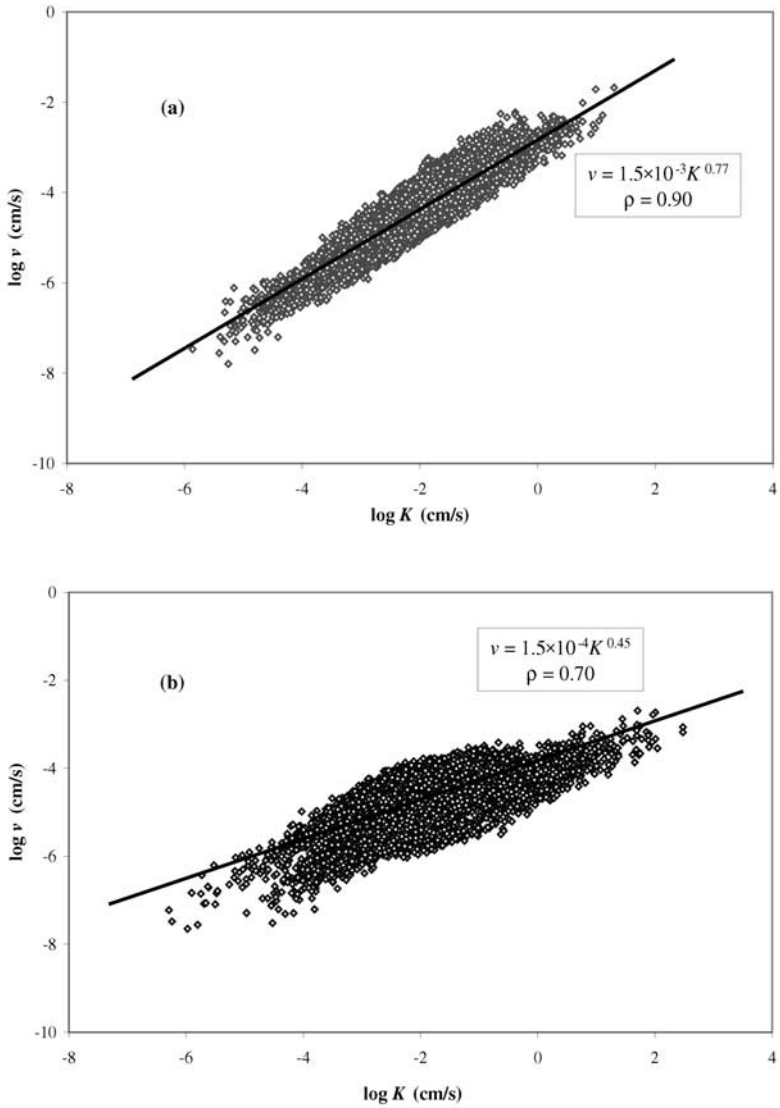

Figure 9. Scatterplot of the K versus total velocity for Sim 75 (a) and Sim 80 (b). The regression analysis and data displayed are for every twentieth data value from the two data sets.

the same K statistics. Hill's estimator for the first realization chosen (Sim 75) gave values of α for Δv_L of 1.14 and for Δv_T of 1.07. Hill's estimator for the second realization chosen (Sim 80) gave values of α for Δv_L of 2.30 and for Δv_T of 2.00.

[28] For visualization purposes a logarithmic transformation of the data was performed. Figure 7 shows the K and total velocity magnitude fields for Sim 75. A plot of the same data with only the largest values visible is also shown. A path of high hydraulic conductivities runs from one constant head boundary to the other. This allows a continuous path of large velocities.

[29] Visualizations of Sim 80 are displayed in Figure 8. Here, there are two separate regions in the model domain where the hydraulic conductivities are large. The resulting velocity field for this K field shows a much smaller percentage of large velocities than for Sim 75. This can be directly attributed to the discontinuous nature of the high K regions for this realization.

[30] The cross-correlation between K and velocity magnitude at zero lag for Sim 75 is shown in Figure 9a. A regression analysis was performed on the log-transformed K and velocity data. This transformation assumes, among other things, that the two parameters are related by a power law and that the residuals are log normally distributed. The regression analysis performed on Sim 75 appears to be valid with a power law relationship $v = 0.0015K^{0.77}$ and strong positive correlation coefficient $\rho = 0.90$ between the logs of the two parameters. The same analysis was also performed on Sim 80 (Figure 9b). A power law relationship $v = 0.00015K^{0.45}$ and weaker correlation coefficient of $\rho = 0.70$ was determined. However, the assumptions of a constant error variance and linearity inherent in a regression analysis are questionable for Sim 80.

7. Discussion and Conclusions

[31] The results for both sets of Monte Carlo simulations show that heavy tailed K fields do give rise to heavy tailed velocity fields. However, the velocity distributions have statistically significantly larger tail parameters than the K

distributions. This indicates that the hydraulic gradient is not uniform and that a weak negative correlation exists between the K fluctuations and the hydraulic gradient. If the tail parameter is considered to be a measure of the heterogeneity of the system, i.e., smaller $\alpha \rightarrow$ heavier tail thickness $\rightarrow$ more heterogeneity [*Painter and Patterson*, 1994], then our results indicate that heterogeneous K fields give rise to less heterogeneous velocity fields.

[32] Flow modeling indicates that the local random variables K and v can be related by a power law:

$$v = CK^m, \tag{3}$$

where the power m is dependent on the strength of the point cross-correlation between the local K and velocity magnitudes. This relationship could also be interpreted as an effective nonlinear Darcy law with the constant C representing some effective regional hydraulic gradient. Substituting the power law K distribution with tail parameter α into equation (3) gives rise to a velocity distribution with probability $P(v \geq y) = Dy^{-\alpha/m}$. From the two realizations examined, $m < 1$, indicating that the power for the velocity distribution is larger than the power for the K distribution.

[33] Continuous connected paths of large hydraulic conductivities are necessary for heavy tailed K distributions to give rise to heavy tailed velocity distributions with a similar tail parameter. Two end-members illustrate this point: flow perpendicular and parallel to perfectly layered soil. For flow perpendicular to layering, conservation of mass dictates that the flow is constant and the equivalent hydraulic conductivity is given by the harmonic mean. Since the velocity is constant, the exponent $m = 0$ in equation (3). On the other hand, for flow parallel to layering, the velocity in each layer is $v = -K(\nabla h/n)$. With a constant $\nabla h/n$, the velocities are a constant multiple of the random K values. Therefore the exponent m in equation (7) is unity, and the tail parameters α_K and α_v are equal.

[34] The K fields we created were heterogeneous in all directions. If the K correlation decayed quickly, then flow would have to cross heterogeneous units. However, if the spatial correlation of the K fields is such that continuous connected paths of large hydraulic conductivities exist from one constant head boundary to another, then flow in parallel is a more appropriate approximation. All random K fields must lie somewhere between the end-members discussed above, so the connectivity of high K layers is important in estimating the exponent m in the effective Darcy's law (3), and the eventual order of the fractional derivative in equation (1).

[35] This also has ramifications for conducting numerical simulations when LRD is present. Many sites under study do not have physically based no-flow and constant head boundaries. The K field will largely direct water flow and create flow in parallel. The effective no-flow and constant head boundaries become aligned with the connected flow paths themselves. However, when a numerical model of the site is built, the boundaries are typically imposed without regard to the alignment of the long-range, connected, high-K fraction. Arbitrarily positioning the boundaries so that flow is forced to go perpendicular to the prevailing layering or connectivity may not accurately represent a real system. In short, generation of a random K field with LRD may not fully specify the flow field, since different realizations of the same model can significantly change the velocity statistics, particularly for the highest velocities.

[36] Additional work needs to be performed to investigate the validity of a fractional ADE. A natural extension of the work performed here would be to track particles through the velocity fields examined in this study. An examination of the temporal correlation, as well as the spatial moments as a function of the mean travel distance would give another estimate of the order of the fractional dispersion derivative. The results would also test the second assumption stated above: that heavy tailed velocity fields give rise to heavy tailed particle jumps and a convergence to Lévy motion modeled by a fractional ADE.

[37] **Acknowledgments.** This material is based upon work supported by the National Science Foundation under grants EAR-9980489, EAR-9980484, DMS-0139927, and DMS-0139943, and by the U.S. Department of Energy, Basic Energy Sciences grant DE-FG03-98ER14885. M. Herrick was also partially supported by the G.B. Maxey Fellowship from the Desert Research Institute.

References

Aban, I. B., and M. M. Meerschaert, Shifted Hill's estimator for heavy tails, *Commun. in Stat. Simul. Comput.*, *30*(4), 949–962, 2001.

Adams, E. E., and L. W. Gelhar, Field study of dispersion in a heterogeneous aquifer, 2, Spatial moments analysis, *Water Resour. Res.*, *28*(12), 3293–3397, 1992.

Baeumer, B., D. A. Benson, M. M. Meerschaert, and S. W. Wheatcraft, Subordinated advection-dispersion equation for contaminant transport, *Water Resour. Res.*, *37*(6), 1543–1550, 2001.

Benson, D. A., The fractional advection-dispersion equation: Development and application, Ph.D. dissertation, Univ. of Nev., Reno, Nev., 1998.

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

Benson, D. A., R. Schumer, M. M. Meerschaert, and S. W. Wheatcraft, Fractional dispersion, Lévy motion, and the MADE tracer tests, *Transp. Porous Media*, *42*(1/2), 211–240, 2001.

Boggs, J. M., L. M. Beard, S. E. Long, and M. P. McGee, Database for the second macrodispersion experiment (MADE-2), *EPRI Topical Rep. TR-10,2072*, Electr. Power Res. Inst., Palo Alto, Ca., 1993.

Dagan, G., Time-dependent macrodispersion for solute transport in anisotropic heterogeneous aquifers, *Water Resour. Res.*, *24*(9), 1491–1500, 1988.

Davis, R., and S. Resnick, Limit theory for moving averages of random variables with regularly varying tail probabilities, *Ann. Probab.*, *13*, 179–195, 1985.

Deutsch, C. V., and A. G. Journel, *Geostatistical Software Library and User's Guide* (*GSLIB*), Oxford Univ. Press, New York, 1998.

Feller, W., *An Introduction to Probability Theory and Its Applications*, vol. II, 2nd ed., John Wiley, New York, 1971.

Goode, D. J., and C. E. Appel, Finite-difference interblock transmissivity for unconfined aquifers and for aquifers having smoothly varying transmissivity, *U. S. Geol. Surv. Water Resour. Invest. Rep. 92-4124*, 79 pp., Reston, Va., 1992.

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

Harbaugh, A. W., and M. G. McDonald, User's documentation for MODFLOW-96: An update to the U. S. Geological Survey modular finite-difference ground-water flow model, *U. S. Geol. Surv. Open File Rep.96-485*, 56 pp., Reston, Va., 1996.

Herrick, M. G., Hydraulic conductivity, velocity, and the order of the fractional dispersion derivative in a highly heterogeneous system, M.S. thesis, Univ. of Nev., Reno, Nev., 2001.

Hill, B., A simple general approach to inference about the tail of a distribution, *Ann. Statist.*, *3*(5), 1163–1173, 1975.

Hill, M. C., Preconditioned conjugate-gradient 2 (PCG2): A computer program for solving ground-water flow equations, *U. S. Geol. Surv. Water Resour. Invest. Rep. 90-4048*, 43 pp., Reston, Va., 1990.

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

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

McCulloch, J., Measuring tail thickness to estimate the stable index α: A critique, *J. Business Econ. Statist.*, *15*, 74–81, 1997.

McDonald, M. G., and A. W. Harbaugh, A modular three-dimensional finite-difference ground-water flow model, *U. S. Geol. Surv. Tech. of Water Resour. Invest., Book 6*, chap. A1, 586 pp., Reston, Va., 1988.

Meerschaert, M. M., and H.-P. Scheffler, A simple robust estimator for the thickness of heavy tails, *J. Stat. Plann. Inference*, *71*(1–2), 19–34, 1998.

Meerschaert, M. M., D. A. Benson, and B. Baeumer, Operator Lévy motion and multiscaling anomalous diffusion, *Phys. Rev. E*, *63*(2), #021112, 2001.

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

Neuman, S. P., On advective transport in fractal permeability and velocity fields, *Water Resour. Res.*, *31*(6), 1455–1460, 1995.

Neuman, S. P., and Y. K. Zhang, A quasi-linear theory on non-Fickian and Fickian subsurface dispersion, 1, Theoretical analysis and application to isotropic media, *Water Resour. Res.*, *26*(5), 887–902, 1990.

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

Painter, S., and L. Patterson, Fractional Lévy motion as a model for spatial variability in sedimentary rock, *Geophys. Res. Lett.*, *21*(25), 2857–2860, 1994.

Pickens, J. F., and G. E. Grisak, Scale-dependent dispersion in a stratified granular aquifer, *Water Resour. Res.*, *17*(4), 1191–1211, 1981.

Rajaram, H., and L. W. Gelhar, Plume-scale dependent dispersion in aquifers with a wide range of scales of heterogeneity, *Water Resour. Res.*, *31*(10), 2469–2482, 1995.

Rehfeldt, K. R., J. M. Boggs, and L. W. Gelhar, Field study of dispersion in a heterogeneous aquifer, 3, Geostatistical analysis of hydraulic conductivity, *Water Resour. Res.*, *28*(12), 3309–3324, 1992.

Schumer, R., D. A. Benson, M. M. Meerschaert, and S. W. Wheatcraft, Eulerian derivation of the fractional advection-dispersion equation, *J. Contam. Hydrol.*, *48*(1/2), 69–88, 2001.

Taqqu, M. S., Random processes with long-range dependence and high variability, *J. Geophys. Res.*, *92*, 9683–9686, 1987.

D. A. Benson, Desert Research Institute, Division of Hydrologic Sciences, 2215 Raggio Parkway, Reno, NV 89512, USA. (dbenson@dri.edu)

M. G. Herrick, Broadbent and Associates, Inc., 2000 Kirman Ave., Reno, NV 89502, USA. (mherrick@broadbentinc.com)

K. R. McCall, Department of Physics, University of Nevada, Reno, NV 89557, USA. (mccall@physics.unr.edu)

M. M. Meerschaert, Department of Mathematics, University of Nevada, Reno, NV 89557, USA. (mcubed@math.unr.edu)