



Fractional Dispersion, Lévy Motion, and the MADE Tracer Tests

DAVID A. BENSON¹, RINA SCHUMER¹, MARK M. MEERSCHAERT² and
STEPHEN W. WHEATCRAFT³

¹*Desert Research Institute, Division of Hydrologic Science, 2215 Raggio Parkway, Reno, NV 89512, U.S.A.*

²*Department of Mathematics, University of Nevada, Reno, NV 89557, U.S.A.*

³*Department of Geologic Sciences, University of Nevada, Reno, NV 89557, U.S.A.*

(Received: 21 January 1999; in final form: 20 December 1999)

Abstract. The macrodispersion experiments (MADE) at the Columbus Air Force Base in Mississippi were conducted in a highly heterogeneous aquifer that violates the basic assumptions of local second-order theories. A governing equation that describes particles that undergo Lévy motion, rather than Brownian motion, readily describes the highly skewed and heavy-tailed plume development at the MADE site. The new governing equation is based on a fractional, rather than integer, order of differentiation. This order (α), based on MADE plume measurements, is approximately 1.1. The hydraulic conductivity (K) increments also follow a power law of order $\alpha = 1.1$. We conjecture that the heavy-tailed K distribution gives rise to a heavy-tailed velocity field that directly implies the fractional-order governing equation derived herein. Simple arguments lead to accurate estimates of the velocity and dispersion constants based only on the aquifer hydraulic properties. This supports the idea that the correct governing equation can be accurately determined before, or after, a contamination event. While the traditional ADE fails to model a conservative tracer in the MADE aquifer, the fractional equation predicts tritium concentration profiles with remarkable accuracy over all spatial and temporal scales.

Key words: fractional derivative, fractional Laplacian, anomalous dispersion, Lévy motion, α -stable, heavy tails, Fokker–Planck equation, MADE site.

1. Introduction

The idea of a fractional derivative is nearly as old as the definition of an integer-order derivative (see the history compiled by Oldham and Spanier, 1974). Fractional derivatives effectively interpolate between successive integer-order derivatives when applied to many kinds of functions. For example, the q th fractional derivative with respect to x (as defined in Appendix A) of e^{ax} is $a^q e^{ax}$. Early in this century, a direct link was made between some stochastic processes and an integer-order differential equation. Einstein (1908) showed that a particle undergoing random Gaussian displacements fulfills the second-order diffusion equation. Kolmogorov extended the link between the second-order diffusion equation and generalized finite-variance processes (see Gnedenko and Kolmogorov, 1954; Feller, 1971). This model is extremely robust, primarily because it is a restatement

of the central limit theorem (CLT). As a result, many different processes behave, in the long-term limit, in accordance with the classical diffusion equation. Einstein formulated his famous treatise on Brownian motion and the diffusion equation prior to Lévy's (1937) discovery of the stable densities, which are a superset of the Gaussian. Had Lévy's discoveries come several decades sooner, Einstein might have chosen a more general approach that would have led to a fractional diffusion equation 90 years earlier. Instead, the definitive link between all Lévy motions and a fractional differential equation has been developed only recently (Sheshadri and West, 1982; Gorenflo and Mainardi, 1998; Meerschaert *et al.*, 1999).

Lévy's stable laws are important because they are the limit distributions of sums of random variables. If a parcel of tracer is characterized as a particle that undergoes independent displacements as it moves through heterogeneous material, then the total displacement tends to converge to a stable variable. The motion is called a Lévy motion; Brownian motion is a subset. A large number of particles, i.e., the plume as a whole, fulfills a large portion of the probability distribution for a single particle. Then the ensemble concentration replaces the probability distribution for a single particle. Benson (1998) showed that the second-order advection–dispersion equation (ADE), even in a nonlinear form with the space-dependent dispersion tensor, cannot generally predict the stable Lévy motions. This is due to the Gaussian-density fundamental solution (Green function) of the ADE. An infinitude of nonlinear superpositions of the Gaussian must still scale proportional to the square root of time (cf., Crank, 1975). Further, combinations of the exponentially decaying tail of the Gaussian cannot model the ‘heavy’ power-law tails that characterize Lévy motion.

Recent work on random motions with power-law excursion probabilities (Fogedby, 1994; Compte, 1996; Saichev and Zaslavsky, 1997; Benson, 1998; Gorenflo and Mainardi, 1998; Meerschaert *et al.*, 1999; Compte and Caceres, 1998; Metzler *et al.*, 1998, 1999; Rocco and West, 1999; Grigolini *et al.*, 1999) has shown that the governing equations are generalizations, using fractional derivatives, of the classical ADE. In this paper, we review the derivation of the governing equation of stable Lévy motions in 1D. We then analyze data from the macrodispersion experiment (MADE) site in *a posteriori* and *a priori* modes to see (1) whether a fractional ADE is an appropriate model; and (2) to see whether the correct order of the fractional ADE, and the values of parameters, can be discerned before any contamination takes place.

We use the MADE site because it defies the traditional approaches used to model contaminant transport based on the classical integer-order ADE. The degree of K variability is very high (Rehfeldt *et al.*, 1992), and the plumes of conservative tracer are highly skewed. Moreover, the relative mass of tracer in the leading edge is very high, or in the vernacular of this study, ‘heavy-tailed.’ The first modeling approach used small-perturbation methods to give the local dispersion tensor within the classical ADE (Adams and Gelhar, 1992). Failure of this method led to more complex approaches, such as finely discretized numerical implementations

of the classical ADE (Zheng and Jiao, 1998). This method still does not adequately reproduce the extreme tailing and skewness of the plume. Some have advocated a model based on diffusion-limited transport at this (C. Zheng, pers. comm., 1998) and other sites (Brusseau, 1992; Haggerty and Gorelick, 1995). The diffusion-limited, or two-site model, is a bimodal method that assumes a certain portion of the plume transfers in and out of a zero-velocity mode, and that Gaussian transport still occurs in each mode. Each of these successive approaches requires an added level of complexity, or required information, in the governing equation. The first needs a different dispersion parameter for each scale. The second adds spatial variability. The third adds transfer coefficients. Berkowitz and Scher (1998) proposed that the underlying transport process at the MADE site is not correctly modeled by a Gaussian process. We agree, and derive a simple (and lowest information-containing) equation that is based on the very large velocity contrasts experienced by the plume.

2. Stable Laws and Motions

Before developing the fractional advection–dispersion equation, it is useful to first investigate the generalized central limit theorem (GCLT) that leads to Lévy’s stable laws. Complete discussions are given by Gnedenko and Kolmogorov (1954) and Feller (1971). In short, if X_1, X_2, \dots, X_n are independent, identically distributed (iid) random variables, then the sum of these converges in distribution to an α -stable random variable (Y). The distribution of Y scales predictably with the number of summands (n) according to

$$\frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma n^{1/\alpha}} \Rightarrow Y, \quad (1)$$

where σ is a positive real constant and \Rightarrow denotes convergence in probability. An example is the sum of finite-variance variables, which converge to (are in the domain of attraction of) a Gaussian random variable. In this case, the scaling rule dictates that $\alpha = 2$, σ is the standard deviation, and (1) is the traditional CLT. All finite-variance random variables are in the Gaussian domain of attraction. Many infinite variance random variables (e.g., Pareto, or power-law) are in the domain of attraction of an α -stable random variable, for which the rate of scaling in (1) follows $0 < \alpha < 2$. A special case of (1) arises when the summands (X) are themselves stable random variables: the sum of any number of iid stable variables is also a stable variable. The GCLT can also be generalized to vectors in any number of dimensions (Meerschaert, 1986).

Liu and Molz (1997a, b), Molz *et al.* (1997), and Painter (1996a, b, 1997) analyze the underlying heavy-tailed distribution of hydraulic conductivity (K) at several sites. In particular, Painter (1996b) shows clear evidence of these α -stable distributions in vertical borehole measurements of K . The stable variables are notably different from the normal or lognormal because the tails of the density func-

tions decay according to a power-law. If Y is a stable variable, then for large arguments

$$P[|Y| > y] \approx W y^{-\alpha}, \quad (2)$$

where $P[\cdot]$ denotes probability and W is a positive constant. This leads to an infinite variance, since $\int y^2 (y^{-\alpha-1}) dy$ diverges for all $\alpha < 2$. The power-law (Pareto) tail allows a higher probability of extreme values than a typical finite-variance distribution (even the lognormal) can predict. Any random variable that has a power-law tail (only the tail is important, virtually any function for small arguments $[y]$ is permissible) belongs to the domain of attraction of an α -stable variable.

Densities of α -stable variables are represented by their Fourier transforms (Feller, 1971; Samorodnitsky and Taquq, 1994). If Y is an α -stable random variable with a probability density function $f(y)$, then the Fourier transform of the density (denoted by the overhat and the change to the Fourier variable k) has the form

$$\hat{f}(k) = \int_{-\infty}^{\infty} e^{-iky} f(y) dy = \exp[-|k|^\alpha \sigma^\alpha \left(1 + i\beta \text{sign}(k) \tan\left(\frac{\pi\alpha}{2}\right)\right) - \mu ik], \quad (3)$$

where the parameters $0 < \alpha \leq 2$, $\sigma > 0$, $-1 \leq \beta \leq 1$, and μ describe the index of stability, the spread, the skewness, and the location of the density, respectively. The $\text{sign}(k)$ function is -1 for $k < 0$ and 1 otherwise. The characteristic function for $\alpha = 1$ (the Cauchy density) is slightly different from (3) and will not be discussed here for the sake of brevity. An important feature of these densities is that they are invariant after shifting by μ and dividing by σ . All versions of the density are scaled and shifted versions of the *standard* density with $\sigma = 1$, $\mu = 0$. We also use an equivalent form

$$\hat{f}(k) = \exp[q\mathcal{C}(-ik)^\alpha + p\mathcal{C}(ik)^\alpha - \mu ik], \quad \alpha \neq 1, \quad (4)$$

where p and q range between 0 and 1 with $p + q = 1$, and \mathcal{C} is a positive constant. The two forms (3) and (4) are related by $\sigma^\alpha = \mathcal{C} |\cos(\pi\alpha/2)|$ and $\beta = p - q$. When the density is *symmetric*, the skewness parameter β is zero (so $p = q = 1/2$), and the symmetric characteristic function is

$$\hat{f}(k) = \exp(-\sigma^\alpha |k|^\alpha - \mu ik). \quad (5)$$

In this form, it is easy to see that the Gaussian (normal) density is α -stable with $\alpha = 2$. It is a simple matter to show that for $\alpha > 1$, the expectation $\mathbf{E}(Y) = \mu$. The mean is undefined for $\alpha \leq 1$. A standard, symmetric, α -stable distribution is characterized by the compact formula $\hat{f}(k) = \exp(-|k|^\alpha)$.

The most important feature of the α -stable distributions is the characteristic exponent (also called the index of stability) α . The value of α determines how ‘non-Gaussian’ a particular density becomes (Figure 1). As the value of α decreases from a maximum of 2, more of the probability density shifts toward the tails. The densities appear near-Gaussian in untransformed coordinates; the difference lies in

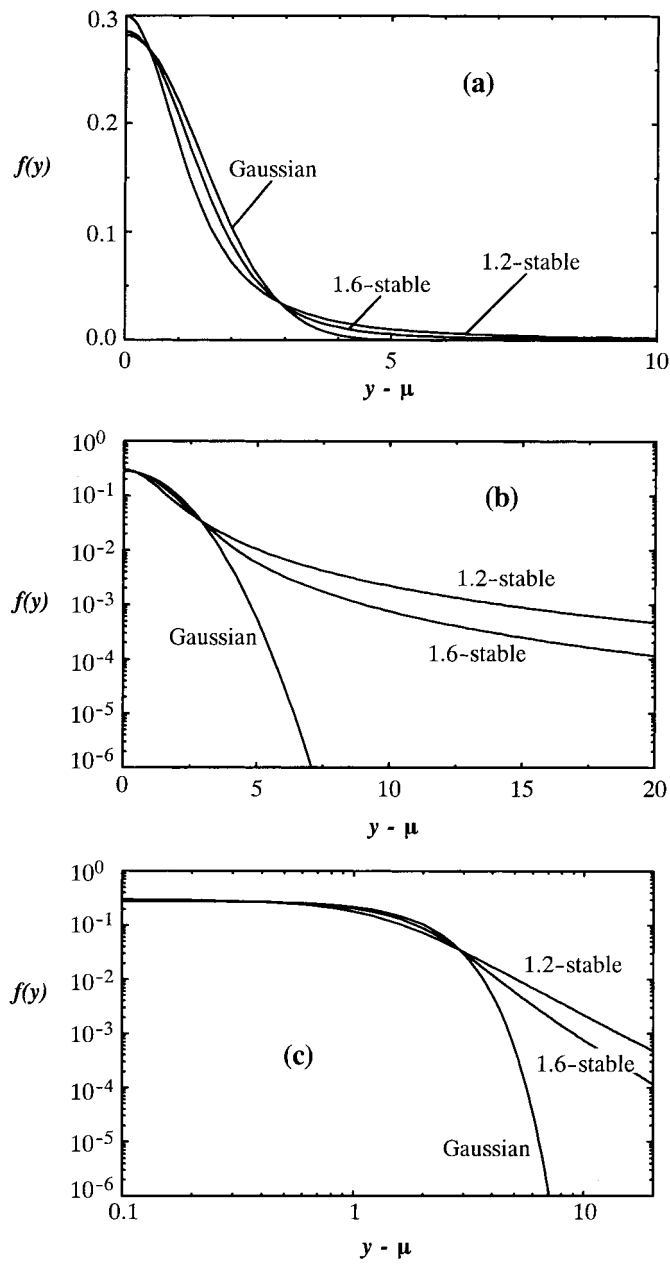


Figure 1. Plots of the positive density functions of several standard symmetric α -stable distributions using (a) linear axes, (b) semilog axes, and (c) log-log axes. For large arguments, the stable densities plot as straight lines with slope $-1 - \alpha$ on log-log axes. For contaminant plumes, $\mu = vt$ is the mean travel distance, and $f(y)$ is the relative concentration (see Section 3).

the relative weight present in the tails. This difference becomes apparent using log axes (Figures 1b and 1c), which show the power-law character of the tails.

3. Governing Equation

Recently, the ADE was generalized to describe all Lévy motion in 3D after a general definition of multidimensional fractional differentiation (Meerschaert *et al.*, 1999):

$$\frac{\partial C}{\partial t} = -v \frac{\partial C}{\partial x} + \mathcal{D} \nabla_M^\alpha C, \quad (6)$$

where v is the drift velocity, \mathcal{D} describes the spread of the process, and the fractional Laplacian operator ∇_M^α specifies the weight of the tails in a moving plume (via α), as well as the propensity for skewed movements in 3D (via the random measure M on the unit sphere). Within this definition, v and \mathcal{D} are constants. The mean advective velocity, or drift of the process, is given by v . If the variance of the particle transition probability is infinite (e.g., power-law tails), then the nonlinear growth of the plume is incorporated within the fractional derivative, not within the leading parameter \mathcal{D} . The derivative is defined so that it correctly captures the nonlinear scaling of the particle transition density. To accurately model particle transport, one need only estimate the order of the fractional derivatives and the constant \mathcal{D} . This equation describes all Lévy motions including the subset of Brownian motion when $\alpha = 2$. In 1D, the random measure M reduces to $p + q = 1$, where the relative weight of p and q describes skewness (Appendix A):

$$\frac{\partial C}{\partial t} = -v \frac{\partial C}{\partial x} + p \mathcal{D} \frac{\partial^\alpha C}{\partial x^\alpha} + q \mathcal{D} \frac{\partial^\alpha C}{\partial (-x)^\alpha}. \quad (7)$$

Solutions to common solute transport boundary value problems (BVPs) are gained through Laplace or Fourier transforms. The 1D fractional-in-space Equation (7) is solved via Fourier transform (Appendix A) to give the fundamental (Green's function) solution:

$$\hat{C}(k, t) = \exp[q \mathcal{D} t (-ik)^\alpha + p \mathcal{D} t (ik)^\alpha - ikvt] \quad (8)$$

which is an α -stable density with an equivalent representation:

$$\hat{C}(k, t) = \exp[-\mathcal{D} t |k|^\alpha |\cos(\pi\alpha/2)| (1 + i\beta \text{sign}(k) \tan(\pi\alpha/2)) - ikvt]. \quad (9)$$

This is the general form of a 1D Lévy motion (Samorodnitsky and Taqqu, 1994). The positive constant $\sigma = (|\cos(\pi\alpha/2)| \mathcal{D} t)^{1/\alpha}$ indicates that the stable density is invariant upon shifting by the mean (vt) and scaling by $t^{1/\alpha}$ (Figure 2). The nonlinear scaling (growth rate) results from the ability to replace the variable k with $t^{1/\alpha} k$ in (9). The distance between two points with constant relative concentration in a plume will grow proportional to $t^{1/\alpha}$, leading to a measured plume variance that is proportional to $t^{2/\alpha}$. Only if $\alpha = 2$ does this reduce to a Gaussian density

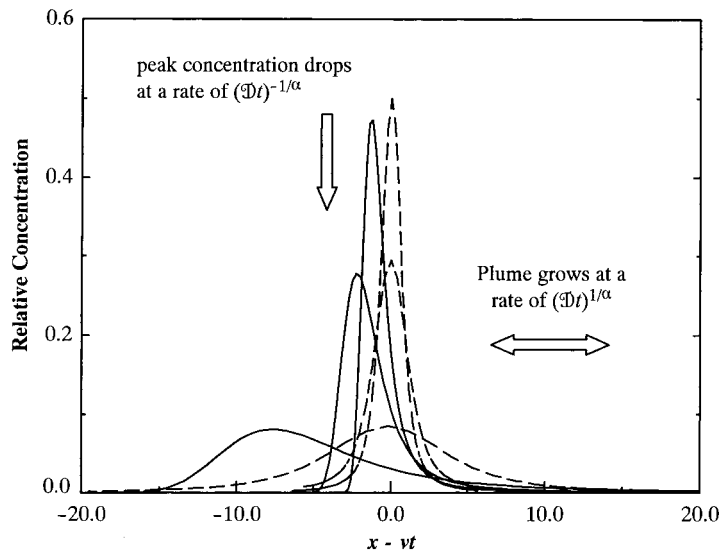


Figure 2. Comparison of the development of spatially symmetric (dashed lines) and positively skewed ($\beta = 1$, solid lines) plumes from a pulse source. Three dimensionless elapsed times (0.5, 1.0, and 5) are shown. All curves use $\alpha = 1.3$ and $\mathcal{D} = 1$. The point $x - vt = 0$ is the plume center of mass for all curves shown. Solutions were gained by numerical integration of (A.15).

with the characteristic Boltzmann scaling (growth) of the plume proportional to \sqrt{t} and a measured variance growing linearly with time.

In addition to growing faster than the Boltzmann scaling predicts, the stable plumes have two significant characteristics. First, the skewness can be positive or negative, depending on whether the bulk of the contaminant is initially placed in lower- or higher-than-average velocity material. This leads to preferential sampling of portions of the velocity distribution. Second, one or both of the tails are 'heavier' than the second-order equation predicts. One or both of the tails asymptotically follow a power law (Figure 1) and can result merely from a heavy-tailed velocity distribution. The peak concentration may also be separated from the mean travel distance (vt) by large distances (Figure 2). The skewness and tailing can arise from the hydraulic characteristics of the aquifer, even in the absence of nonlinear sorption or completely stagnant, diffusion-limited zones.

4. Evidence for Fractional Dispersion in the MADE Tracer Tests

The MADE site is located on the Columbus Air Force Base in northeastern Mississippi. The two large-scale, natural-gradient tracer tests performed there differ from other large-scale tracer experiments such as the Borden site in Ontario, Canada (e.g., Sudicky, 1986) or the Cape Cod site in Massachusetts (e.g., LeBlanc *et al.*, 1991) because of the strong heterogeneity of the aquifer (Rehfeldt *et al.*, 1992;

Boggs *et al.*, 1993). The unconfined, alluvial aquifer consists of generally unconsolidated sands and gravels with smaller clay and silt components. Irregular lenses and horizontal layers were observed in an aquifer exposure near the site (Rehfeldt *et al.*, 1992). Detailed studies characterizing the spatial variability of the aquifer and the spreading of the conservative tracer plume for the experiment conducted between October 1986 and June 1988 (MADE-1) are summarized by Boggs and Adams (1992), Adams and Gelhar (1992), and Rehfeldt *et al.* (1992). A synopsis of the second experiment (MADE-2), conducted between June 1990 and September 1991, is given by Boggs *et al.* (1993).

4.1. PREVIOUS STUDIES

One purpose of the MADE tracer tests was to create a database for validating groundwater flow and solute transport models. The most commonly applied stochastic theories, based on dispersive mixing in heterogeneous aquifers (e.g., Gelhar and Axness, 1983; Dagan, 1984), were tested by Adams and Gelhar (1992). They found that classical Gaussian models were not able to reproduce the highly skewed plume evolution found at the MADE site. They presented no data concerning the heavy plume tails.

Berkowitz and Scher (1998) show how a very general, non-Gaussian, model of particle transport can be applied to fracture networks. Their work is based on particle transitions described by a joint space-time distribution. If one is able to identify this distribution, then a plume can be generated. Based on numerical generation of fracture networks and an analogy to the preferential flowpaths in the highly heterogeneous MADE site, the authors present several figures that show an improved fit over those presented by Adams and Gelhar. It is not clear whether the density of Berkowitz and Scher (1998) reproduces the extremely heavy tails of the plume or scales at the correct rate, since the authors predict that the ratio of the sample mean and standard deviation is a constant. These quantities measured in the MADE tests do not have the same slope on a log-log plot. Their methodology also depends on measuring an empirical joint space-time particle transition density.

Zheng and Jiao (1998) produced 3D numerical simulations for the MADE-2 experiment using K distributions derived from borehole flowmeter data and various geostatistical methods. They found that the simulated plume was highly sensitive to the distribution of the K field. Additionally, they were not able to concurrently reproduce the movement of the plume's center of mass and the significant spreading of the tracer at the plume's leading edge. Finally, they suggest that inconsistencies in modeling may be due to temporal flow field variations and preferential flowpaths not consistent with the assumption of a Gaussian or finite-variance pore velocity distribution. The large-scale preferential flowpaths lead to a higher probability of sustained high velocity events that the normal or lognormal cannot predict. Further, the K increments (Appendix B) measured at the smallest scale (a vertical lag of 15 cm) suggest that heavy-tailed velocity distribution precludes the use of

the traditional ADE even at the smallest scale within finely discretized numerical models.

The fractional ADE can describe both the heavy-tailed transport and the skewed transitions. We develop a relationship between aquifer properties and the parameters within the fractional ADE, and compare the estimates to those gained from sampling a contaminant plume at early time. The former leads to estimation of parameters *a priori* a contamination event, while the latter leads to *a posteriori* parameter estimation.

4.2. A PRIORI ESTIMATION OF PARAMETERS

Recent studies show that permeability variations can be described with stable or other heavy-tailed distributions (e.g., Painter, 1996b). Liu and Molz (1997b) have suggested that the $\ln(K)$ distribution at the MADE site is heavy-tailed with a power-law index that is a function of sample separation. We re-examine the statistical distribution of the K field at the MADE site, since the ability to estimate the order of the fractional dispersion term *a priori* depends on the ability to accurately estimate the Lévy index (α) of aquifer K data. The values of α and \mathcal{D} would follow immediately from a direct measurement of the natural velocity profiles; however, this is a difficult task. The basis of the following estimates is the assumption (as yet unverified) that the head gradient fluctuations within the soil are negligible compared to the magnitude of the K fluctuations. In this case, the small-scale velocity field is dominated by the many order-of-magnitude differences in K , rather than the smaller differences in head gradient. Therefore, if the K distribution is found to have a power-law tail with exponent α , we assume that the velocity distribution would have a similar exponent. Only the scale factor of the distribution (and possibly the low-velocity shape of the distribution) would change according to the mean gradient. An outline of the theory and application of four popular estimators is contained in Appendix B. One method for determining the Lévy index from a sample K or ΔK distribution appears to be most robust and stable; we present it here.

The Pareto is the most simple of the heavy-tailed distribution functions for a random variable X in an α -stable domain of attraction. The distribution function is defined by $P[X < x] = 1 - Wx^{-\alpha}$, where $\alpha > 0$, $W > 0$, and $x \geq W^{1/\alpha}$. Since one or both of the tails of a non-Gaussian stable distribution are asymptotically Pareto, it is also possible to model a stable distribution by evaluating its tails with a shifted Pareto distribution (Anderson and Meerschaert, 1998). Meerschaert and Scheffler (1998) develop a simple method for estimating the thickness of heavy tails for Pareto-like data with $W = 1$. When $0 < \alpha < 2$, $1/\alpha$ can be approximated by

$$\frac{1}{\alpha} \approx \frac{\gamma + \ln \sum_{i=1}^n (X_i - \bar{X})^2}{2(\gamma + \ln n)}, \quad (10)$$

where $\gamma \approx 0.5572$ is Euler's constant, X_1, \dots, X_n are the sample data, and \bar{X} is the sample mean.

Details of borehole flowmeter logging, slug tests, and laboratory permeameter methods for measuring K at the MADE site are described by Rehfeldt *et al.* (1992). Application of Equation (10) to all normalized (each data point divided by the sample median) vertical K increment data yields an estimate of $\alpha = 1.1$. Another popular estimator is given by Hill (1975). Hill's estimator is similar to Meerschaert and Scheffler's except that does not allow the shift of the Pareto. However, Hill's estimator is able to provide the scale (width) parameter W , which is similar to the standard deviation of a Gaussian. It is this scale of the ΔK distribution that ultimately determines the magnitude of the dispersion coefficient.

The Hill's estimator fits the equation $P[\Delta K > x] = Wx^{-\alpha}$. Sums of Pareto data converge to an α -stable with an equality of the scale parameters of $\sigma^\alpha = W\Gamma(1 - \alpha) \cos(\pi\alpha/2)$ (e.g., Feller, 1971; Janicki and Weron, 1994), where $\Gamma(\cdot)$ is the gamma function. From (9) we have $\sigma^\alpha = \mathcal{D}t|\cos(\pi\alpha/2)|$, which has units L^α . The Pareto scale factor W from the ΔK data has units $(\text{cm/s})^\alpha$. Therefore, the ΔK increments must be translated to relative particle separation size over a short period of time. These separation distances are the heavy-tailed random variables in (1) that converge to an α -stable variable governed by (7). The gradient divided by the porosity (J/η) is assumed to be relatively uniform giving the velocity increments $\Delta v = (J/\eta)\Delta K$. The difference in the solute particle travel distances (Δd) is this random variable multiplied by some time interval τ , which integrates the temporal correlation, giving $P[\Delta d > (J\tau/\eta)x] = Wx^{-\alpha}$. With the change of variable $y = (J\tau/\eta)x$, the distribution becomes $P[\Delta d > y] = W(J\tau/\eta)^\alpha y^{-\alpha}$. Now equating the dimensionally correct scale factors of the α -stable solution and this Pareto Δd distribution, $\mathcal{D}\tau = W(J\tau/\eta)^\alpha |\Gamma(1 - \alpha)|$. Therefore, the dispersion coefficient is given by the K statistics, the mean gradient and porosity, and a characteristic transition time by the relationship

$$\mathcal{D} = W(J/\eta)^\alpha \tau^{\alpha-1} |\Gamma(1 - \alpha)|. \quad (11)$$

Note that as α approaches unity, the estimate becomes very insensitive to the choice of transition time (τ). For example, when $\alpha = 1.1$, a ten-fold decrease in τ only changes the estimate of \mathcal{D} by 20%. For any value of $\alpha > 0.5$, the estimate of \mathcal{D} is most strongly sensitive to the value of J/η . The assumed gradient also directly changes the mean velocity, so the estimated values of v and \mathcal{D} must be tracked simultaneously.

Once again assuming that the gradient is relatively constant throughout the soil, Darcy's Law gives $v = \bar{K}(J/\eta)$. The arithmetic mean K value from the flowmeter and core tests is 24 m/d, which compares favorably to a pumping test that indicated a horizontal K of 17 m/d (Boggs and Adams, 1992). Hill's estimator for the K increments (Appendix B) gives $\alpha = 1.1$ and $W = 0.0028 (\text{cm/s})^{1.1}$. Using $\tau = 1$ day in (11) gives $\mathcal{D} \approx 50 \text{ m}^{1.1}/\text{d}(J/\eta)^{1.1}$. Table I summarizes the coefficients v and \mathcal{D} for a range of values (J/η). The gradient experienced by most of the plume

Table I. Relationship of the constants v and \mathcal{D} to the assumed gradient for a range of values reasonable to the MADE site

J/η	v (m/d)	\mathcal{D} (m ^{1.1} /d)
0.005	0.12	0.14
0.010	0.24	0.32
0.015	0.36	0.50

(after a distance from the release point of approximately 10 m) is typically lower than 0.003 (Boggs and Adams, 1992), so we estimate a value of approximately 0.005 for J/η . The validity of this entire procedure can only be addressed through a detailed examination of the relationship between head gradient and the heavy-tailed K distribution, as well as the decay of particle time correlation on the order of 1 day.

4.3. A POSTERIORI ESTIMATION OF PARAMETERS

Two *a posteriori* methods for estimating the Lévy index are presented using information from the MADE database to validate the *a priori* estimate of $\alpha = 1.1$. These methods require far less data collection than extensive K measurements and are more practical methods for real plumes. We use the analyses of the MADE-I test presented by Adams and Gelhar (1992).

Approximately 10^4 L of water containing 2,500 mg/l of bromide were injected into the MADE aquifer for the first test. Over the next 20 months, seven sampling events using an extensive array of multi-level samplers (MLS) were performed. Using the spatial moments of plume growth for the MADE-1 experiment, Adams and Gelhar (1992) compare a normalized distribution of mass with a Gaussian distribution with similar 0th, 1st, and 2nd longitudinal moments. They noted that the longitudinal variance of the plume grew at an accelerated rate. A plot of the apparent plume variance along the travel direction on log–log axes confirms that plume growth is indeed nonlinear (Figure 3a). Because the density of an α -stable plume is scale invariant with $t^{1/\alpha}$, the variance of the discretely-sampled plume should grow proportionally with $t^{2/\alpha}$ (Benson, 1998). Thus, the slope of the graph (Figure 3a) provides a second estimate of the Lévy index ($\alpha = 1.2$), that agrees reasonably closely with the *a priori* estimate. The plume does not appear to follow the early- (slope = 2) or late- (slope = 1) time behavior predicted by local stochastic theories (e.g., Gelhar and Axness, 1983). Though the data show some scatter due to the difficulties of measurement, there is no systematic deviation from the general trend of $\alpha = 1.2$ (Figure 3a).

A simpler method for obtaining parameters for the fractional ADE arises from the fact that the plume is a 3D density that grows at a rate of $(\mathcal{D}t)^{1/\alpha}$ (Meerschaert *et al.*, 1999). The peak concentration should thus follow:

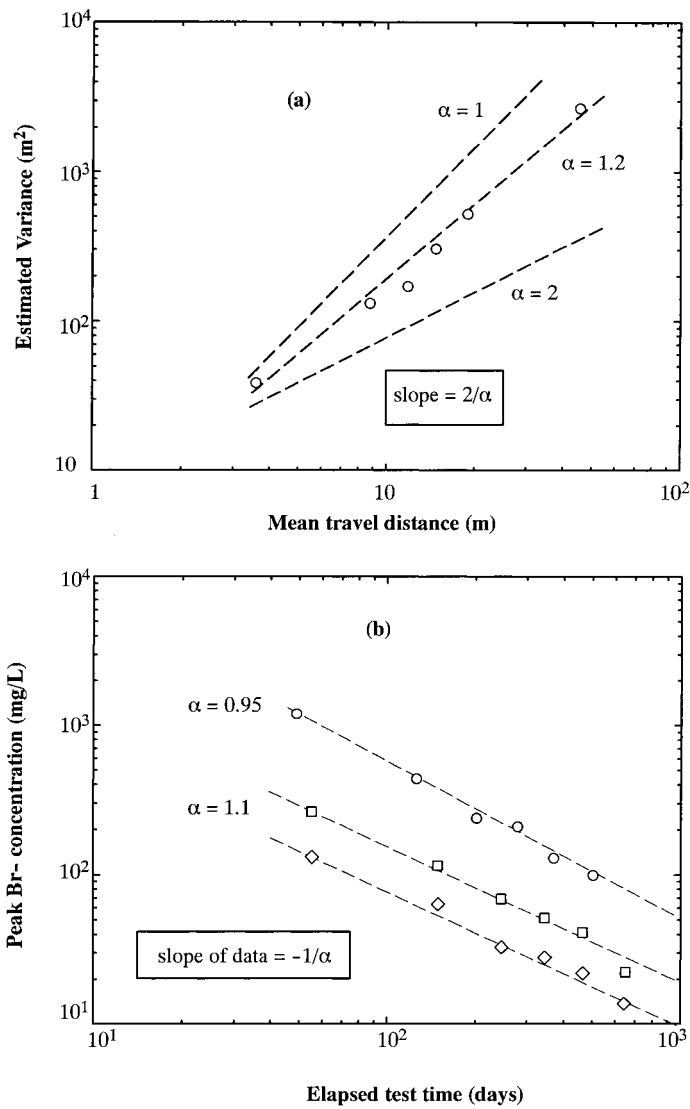


Figure 3. Estimation of Lévy index (α) from plume data: (a) calculated plume variance (Adams and Gelhar, 1992) along the mean direction of travel for the MADE-1 experiment and (b) maximum bromide concentration measured during each snapshot. Circles represent unfiltered data, squares and diamonds are spatially filtered (averaged, see Adams and Gelhar, 1992).

$$C(t) = \frac{C_0}{(\mathcal{D}t)^{1/\alpha}}. \quad (12)$$

The slope of a log-log plot of the peak concentration versus time should be $(-1/\alpha)$. Adams and Gelhar (1992) list the maximum bromide concentrations before and

after spatial averaging (filtering). The unfiltered data indicate that $\alpha = 0.95$, while both sets of filtered data conform to a line with $\alpha = 1.1$ (Figure 3b).

The *a posteriori* estimates of the order of the fractional ADE (α) show some variability, from 0.95 to 1.2. It is not known which are the best estimates. The unfiltered (raw data) peak concentrations should have the highest degree of variability. The variance estimations suffer bias, especially at late time, from truncation of distant, low concentrations by the fixed wellfield and finite detection limits. The method based on maximum spatially-averaged concentrations is probably the most useful, since this represents the simplest form of data collection in real-world applications: monitoring wells are typically screened across large vertical distances, providing data similar to Adams and Gelhar's (1992) filtered data. Only the maximum concentration from a vertically averaged (filtered) wellfield at any time is identified and used as the peak concentration. Numerical integration is not needed, and the value of α should be apparent after the first few measurements. This method gives a value equal to the *a priori* estimate of $\alpha = 1.1$. Solutions (e.g., Figure 1) suggest that the recognition of stable transport is more important in predicting concentrations in the leading edge than exact determination of α . A sensitivity analysis is needed to assess the importance of accurate estimation of α and \mathcal{D} .

4.4. ANALYTIC SOLUTIONS

We use the parameters estimated from the aquifer K statistics (supported by Adams and Gelhar's analysis of the MADE-1 bromide plume) to predict the MADE-2 tritium plume. We have only recently received the MADE-1 bromide data and have not yet re-analyzed that test. During the MADE-2 experiment, 9.7 m^3 of tritiated water and four organic compounds were injected into the shallow alluvial aquifer at the test site. Over the next 15 months, five snapshots of the tracer concentration distributions were collected. Several sets of analytic solutions have been generated using the aquifer parameters estimated in previous sections (see Appendix B for a discussion). We compare the solutions of the traditional ADE ($\alpha = 2$) and the fractional ADE ($\alpha = 1.1$):

$$\frac{\partial C}{\partial t} + v \frac{\partial C}{\partial x} - \mathcal{D} \frac{\partial^2 C}{\partial x^2} = C_0 x_0 \delta(t, x), \quad (13)$$

$$\frac{\partial C}{\partial t} + v \frac{\partial C}{\partial x} - \mathcal{D} \frac{\partial^{1.1} C}{\partial x^{1.1}} = C_0 x_0 \delta(t, x), \quad (14)$$

where $C_0 x_0 \delta(t, x)$ is the initial solute concentration (C_0) spread over some injection distance x_0 , which is mathematically concentrated into a Dirac delta function 'spike'. A value of 55,610 pCi per $\text{cm}^3 \times 400 \text{ cm}$ (the approximate length of the injection area) = $2.78 \times 10^7 \text{ pCi/cm}^2$ was used in all solutions for the MADE-2 tritium plume. The values of v and \mathcal{D} in the fractional ADE (14)

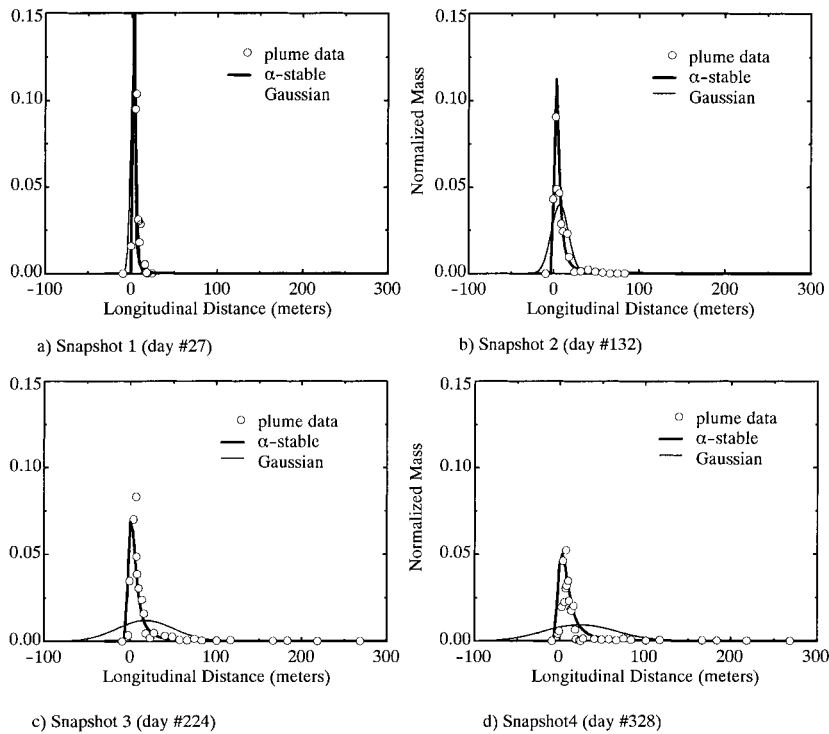


Figure 4. Linear plots of the MADE-2 normalized longitudinal tritium mass distribution at four intervals. Analytic solutions of the ADE and the fractional ADE were gained by numerical integration of (A.15).

are constants for all times of interest (0.12 m/d and $0.14 \text{ m}^{1.1}/\text{d}$, respectively). Similar to Adams and Gelhar (1992), we used ‘best fit’ values of v and \mathcal{D} in the traditional ADE (13) for each time period by calculating the mean and variance of the measured plume.

Figure 4 shows the longitudinal distribution of total mass for MADE-2 snapshots 1–4. The data points represent the maximum concentration measured in vertical slices perpendicular to the direction of plume travel. These maxima were then integrated versus travel distance to get the 1D ‘mass recovery’. The α -stable solutions (14) are superimposed on the normalized mass distributions. Also shown, in the manner of Adams and Gelhar (1992), is a Gaussian distribution with first and second moments equal to those of the normalized sample data. Note that this 1D analysis assumes that the mass lost to transverse dispersion is relatively small. The maximum concentrations were then normalized by the total mass recovered during each snapshot so that each plot has the same 1D mass. We use this method for ease of comparison with the data presented for the MADE-1 test (Adams and Gelhar, 1992). Since the concentration maxima along the plume core represent conditional densities of the entire 3D plume, their evolution will be similar to marginal (integrated) densities (Benson, 1998).

The initial 1D mass was calculated using a constant value of porosity of 0.35 (Boggs *et al.*, 1993). The normalizations represent mass recovery of 145%, 82%, 62%, and 37% for snapshots 1–4, respectively. Mass recovery in excess of 100% in early sampling rounds, followed by decreases in mass recovery to levels below 50% were also observed during the MADE-1 test (Adams and Gelhar, 1992) and the Borden tracer experiment (Freyberg, 1986; Rajaram and Gelhar, 1991). Adams and Gelhar (1992) discuss potential sources for these phenomena. In particular, they show that significant amounts of bromide mass passed beyond the MADE-1 sampling array later in the test. They also show that sampling bias toward higher local K material can seriously underestimate late-time mass. The heavy leading edges predicted by (14) suggest that some of the tritium may be downgradient of the sampling array. No snapshot was bounded in the downstream direction by nondetectable levels of tritium. We have not estimated this mass loss. Radioactive decay of tritium accounts for only about 5% of the lost mass after 328 days. Data for snapshot 5 are not presented because only a small portion of the plume was sampled.

Both the fractional ADE and Gaussian models reproduce the movement of the center of mass in the early time snapshots. In addition, the fractional ADE is able to capture the separation of the positions of the plume mean and the mode, or peak, concentration. The two late-time snapshots clearly show that the fractional ADE is a better model than a local (Gaussian) ADE in predicting the movement of the tritium plume. The fractional ADE was developed with heavy (power-law) tails in mind, so we also plot the data using logarithmic axes for the tritium mass (Figure 5). Here the predictive accuracy of the fractional ADE is dramatically evident, as is the inadequacy of thin-tailed second-order theories at later times. The traditional ADE, even using a best-fit value of \mathcal{D} , underpredicts concentrations throughout most of the downstream portion of the plume by many orders of magnitude. Using the ADE in a predictive mode with smaller values of \mathcal{D} from classical stochastic theories would result in even worse estimates. Adams and Gelhar (1992) and Rehfeldt *et al.* (1992) discuss the fact that classical stochastic theory predicts a dispersion coefficient from 4 to 50 times smaller than the measured plume moments suggest. However, the traditional ADE seems to work very well for the earliest time period. The unsampled portion of the plume beyond 30 m might have measurable concentrations of tritium that would extend the tail of the real plume somewhat. It is also possible that the initial plume placement was far from the high velocity areas, and that some time was needed before the GCLT (1) was reasonably valid.

Since the fractional ADE predicts a leading edge of the plume that is a power function of distance, the decline in normalized mass (concentration) should appear linear on a log–log plot (Figure 1c). This is shown in the last three snapshots (for brevity we show the last two in Figure 6), supporting the notion that Lévy motion is a good model of transport at this site. We mention that unnormalized data concentrations would plot slightly lower (one to two diameters of the circles) on these plots.

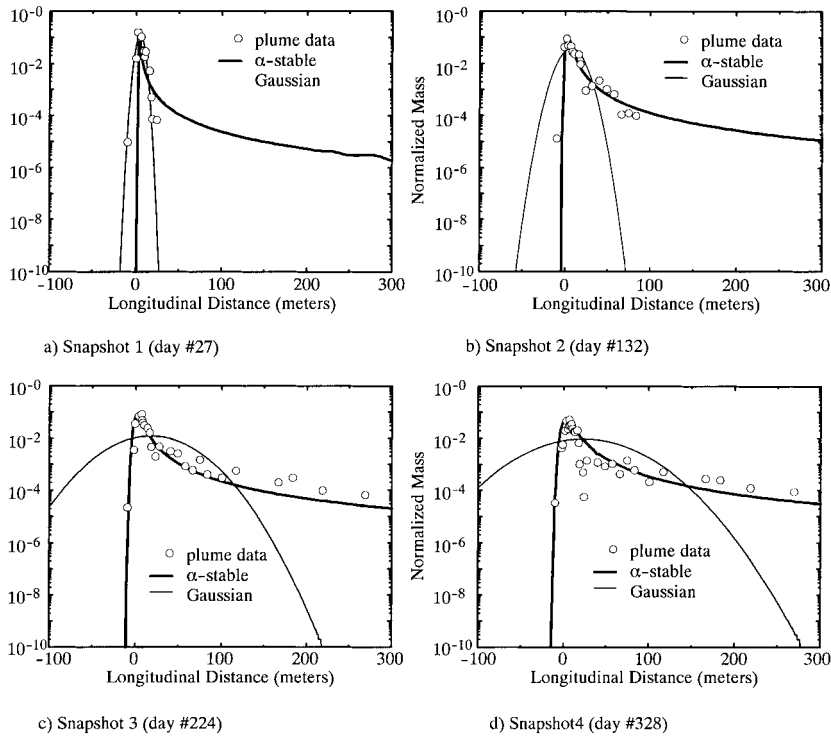


Figure 5. Semi-log plots of the MADE-2 tritium plume. Analytic solutions of the ADE and the fractional ADE were gained by numerical integration of (A.15).

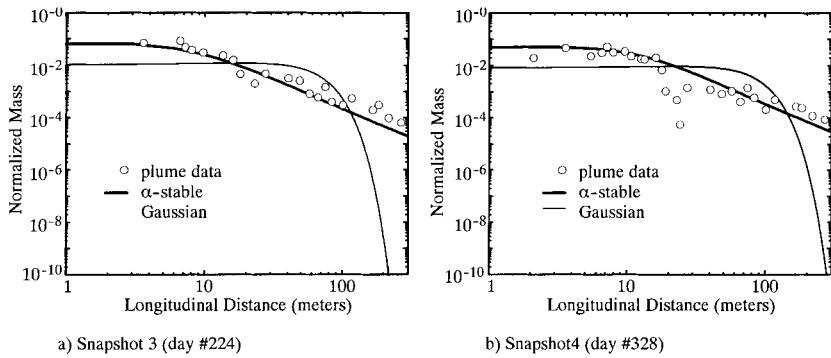


Figure 6. Log-log plots of the MADE-2 tritium plume. Analytic solutions of the ADE and the fractional ADE were gained by numerical integration of (A.15).

4.5. BEHAVIOR OF PLUME PEAK VERSUS MEAN

The rate at which any two portions of a density separate is well known. A density $f(x)$ of a scaled random variable grows by $a^{-1}f(x/a) = f(x)$. The fractional ADE density is invariant according to $\sigma^{-1}f(x/\sigma) = f(x)$, where $\sigma = (|\cos(\pi\alpha/2)|\mathcal{D}t)^{1/\alpha}$. This includes the ADE as well. The distance between

the peak and the shift vt (the mean for $\alpha > 1$) at any time grows like $A(\cos(\pi\alpha/2)\mathcal{D}t)^{1/\alpha}$, where A is some constant that depends on the skewness and order α . If the skewness is zero, so is A . For a positively skewed plume, the peak lies behind the mean, so A is negative, since the distance is always smaller than vt . The distance travelled by the peak (x_p) is merely $x_p = vt + A(\cos(\pi\alpha/2)\mathcal{D}t)^{1/\alpha}$. The velocity of the peak is the time derivative:

$$v_p = v + \frac{A}{\alpha}(\mathcal{D} \cos(\pi\alpha/2))^{1/\alpha} t^{1/\alpha-1}. \quad (15)$$

Combining constants into $B = A\alpha^{-1}(\mathcal{D} \cos(\pi\alpha/2))^{1/\alpha}$ gives $v_p = v + Bt^{1/\alpha-1}$. For $1 < \alpha < 2$, the velocity of the peak (or any concentration percentile) approaches the mean velocity, as observed by Taylor (1953). For α very near unity, the velocity approaches the constant $v \pm B$ for any part of the plume, which is required for piston (wave equation) flow. For $\alpha < 1$, the apparent velocities of all parts of the plume are equal for vanishingly small time and diverge at later time (as does the theoretical plume centroid).

We are not aware of a handy expression for the mode location of an α -stable; however, one can generate a standard density (Appendix A) and get an accurate estimate for the value of A . Numerical estimators and plots of the mode location are also given by Fofack and Nolan (1998) and Nolan (1998). A maximally skewed density for $\alpha = 1.1$ has $A \approx -6$ (Nolan, 1998). The distance between the peak and the measured mean is very sensitive near $\alpha = 1$. For example, $\alpha = 1.14$ has $A \approx -4.5$. Therefore, inaccurate estimation of the parameters v , \mathcal{D} and α could lead to a prediction that the plume peak is moving upstream. A corollary is that such an infeasible scenario would narrow the possible range of values of one of the parameters, given good estimates of the other two.

At the MADE site, the separation of the peak and mean positions is evident for later snapshots (Figure 7). For this plot we adjusted the values of $v = 0.12$ m/d and $\mathcal{D} = 0.14$ m^{1.1}/d predicted from the K distribution and J/η to better match the observed mean and peak locations. Our *a posteriori* estimates of $v = 0.11$ m/d and $\mathcal{D} = 0.12$ m^{1.1}/d based on this plot (Figure 7) are still very close to the initial estimates that were based only on the hydraulic properties of the aquifer. A thorough ‘fitting’ of \mathcal{D} would still depend on an *a posteriori* best-fit value of α . The early time location of the peak concentration is inflated by the large (several meters in the x -direction) size of the initial tritium pulse. Also, the downstream edge of the plume was not adequately sampled during snapshot 5 at 440 days, so the mean was not calculated. The procedure of calculating peak location is complicated for values of $\alpha < 1$, since one must also estimate the plume’s measured mean, which depends heavily on truncation from a finite sampling array. We have not performed this detailed sensitivity analysis. For illustrative purposes, we plot the behavior of the peak for $\alpha = 0.95$ using a best fit value $A(\mathcal{D} \cos(\pi\alpha/2))^{1/\alpha} \approx 0.016$ (Figure 7). A value of $\alpha = 0.95$ also explains the separation of the peak and measured plume mean. This would imply that the distance between readily identifiable portions of

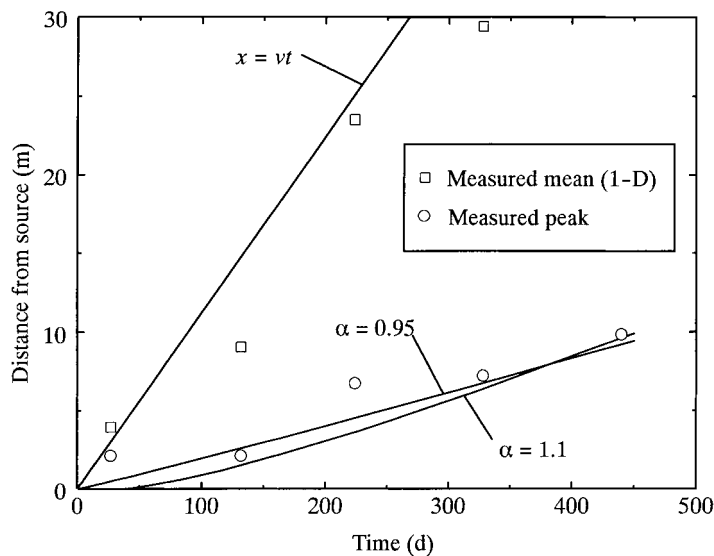


Figure 7. Locations of the center of mass and peak concentration for the positively skewed MADE site tritium plume (symbols) and theoretical estimates for two values of α (0.95 and 1.1). For $\alpha = 1.1$, the plotted mean velocity is $v = 0.11$ m/d and $\mathcal{D} = 0.12$ m^{1.1}/d. Measured center of plume mass calculated by interpolation of 1D values plotted in Figures 4–6.

the plume (i.e., the peak and the measured mean) is diverging with increasing time. The confidence in the slope of the peak location is not high, so we do not consider this a diagnostic tool for estimation of α .

5. Discussion

Skewed and/or heavy-tailed plumes do not necessarily imply a heavy-tailed velocity distribution. Non-stationary velocity fields that arise from variable recharge or boundary conditions such as aquifer thinning can preferentially stretch certain portions of a plume. We have not analyzed the aquifer boundary conditions in detail. Nor have we investigated the applicability of more capable solutions than a local ADE affords. In particular, can the following methods be implemented and produce accurate predictions of the plume? (1) high-variance decomposition methods (Serrano, 1995); (2) nonlocal, second-order equations (e.g., Deng *et al.*, 1993); (3) multi-rate mobile/immobile models (e.g., Haggerty and Gorelick, 1995); or (4) numerical implementations using the most recent non-Gaussian characterizations of the K field (Liu and Molz, 1997a,b).

At the MADE site, the very high degree of K heterogeneity seems to support an assumption of a power-law density and a simple analytic model. The assumption of negligible J/η fluctuations should be most valid when the variability of K is greatest. Therefore, the accuracy of the fractional ADE without any knowledge of plume history should be greatest at highly heterogeneous sites. We make these

assertions without proof that a power-law K distribution gives rise to power-law velocity field with the same exponent. The skewness of the MADE plume may have arisen due to its initial conditions. Since the plume was placed in low-permeability material, the propensity for faster transitions is much higher for quite some time. Had the slug been placed in a high permeability channel, the opposite might be expected, although recent work by Schumer *et al.* (1999) indicates that natural α -stable plumes should be maximally skewed ($\beta = \pm 1$) based on the K distribution alone.

The parameters α and \mathcal{D} come from a 2-parameter K distribution function, the assumption of a characteristic transition time (τ), and (J/η) . The parameter \mathcal{D} in a second-order equation comes from the variance of $\ln(K)$, the functional form of the autocorrelation (either a characteristic length or a simple function of separation), (J/η) , and the local transverse dispersion coefficient. To model a skewed plume with any order equation, one or more additional parameters are needed: (β) for the fractional ADE and a mass transfer rate coefficient and some apportionment of the porosity for mobile/immobile formulations (Haggerty and Gorelick, 1995). Thus, the fractional equation (7) requires the same amount of information or less to produce qualitatively similar predictions as local second-order formulations. The distinguishing characteristic of the fractional solution is the prediction of at least one power-law tail in time and/or space (Benson, 1998).

Tests at mildly heterogeneous sites (Cape Cod) and even within a highly homogeneous sandbox show some of the hallmarks of α -stable plume development (Benson, 1998; Benson *et al.*, 1999b). This is not totally surprising, since random walks by conservative solutes without power-law temporal correlation should converge to some α -stable distribution, based on the GCLT (1). A Gaussian stable plume is a subset of these. The power-law K distribution must have an upper cutoff, hence a finite variance. Since the K distribution reports average velocity in finite volumes, one must assume that actual particle velocity is truncated by a much higher and unknowable upper bound. Mantegna and Stanley (1995) show that truncated α -stable motions stay similarly α -stable for very long times before quickly changing to a Gaussian, particularly when α is much lower than two. Painter (1997) observes this point and gives evidence that a plume might have to travel huge distances before integrating enough of these cutoff transitions. These studies lend support for a fractional ADE formulation for many conservative plumes. The fractional approach has also been shown to work well in unsaturated porous media (Pachepsky, 1998; Pachepsky *et al.*, 1999), where self-generating channels foster similar long-range particle velocity dependence.

6. Conclusions

In some approaches to modeling field-scale transport, such as small perturbation of a stochastic second-order ADE, the dispersive growth process is still fundamentally Gaussian; super-diffusive behavior is forced by scale-dependent macrodispersion

parameters. If the actual growth process is stable but non-Gaussian, then such methods cannot accurately simulate dispersive growth at any scale. This is because the Gaussian fundamental solution cannot reproduce the heavy tails or scale correctly with time. The non-Fickian behavior of the fractional ADE is obtained by matching the order of differentiation of the dispersive derivative to the exponent of the growth process. The fundamental solution to the fractional ADE produces non-Gaussian, heavy-tailed densities (plumes) that simultaneously predict super-Fickian transport and the heavy breakthrough curve tails that are ubiquitously found at field sites and often attributed to channeling, dead-end pores, nonlinear adsorption, and/or zero permeability zones.

There are no restrictions on the fractional-order theory in terms of the variance of $\log(K)$ – it is based on an infinite variance velocity model. The theory is relatively simple to apply: the parameters are constants that can be estimated from early-time plume data or the underlying K data. In contrast, the early plume behavior gives no indication of the late-time macrodispersivity used in a second-order equation. At the MADE site, the fractional ADE gave better predictions with a very low information requirements: a simple equation reproduces very complex behavior.

Using the K data as a surrogate for velocity, the fractional equation parameters were accurately estimated at the MADE site. Our analysis of the K increments at the MADE site using robust tail estimators indicates a power-law of order $\alpha = 1.1$. The power-law tail velocity field imparts instantaneous solute displacements that are power-law tailed as well. These displacements converge to an α -stable plume that is modeled by a fractional dispersion term. This is reflected in the measurements of the MADE-2 tritium plume. Finally, the plume is positively and maximally skewed, which is simply handled by fractional derivatives and the α -stable fundamental solutions.

Acknowledgements

D.B. received support from the U.S. Dept. of Energy – Basic Energy Sciences grant #DE-FG03-98ER14885. R.S. was supported by the Sulo and Aileen Maki Fellowship from the Desert Research Institute. We thank Chunmiao Zheng at the University of Alabama and the Electric Power Research Institute (Palo Alto) for providing the MADE data.

Appendix A. Derivation of the 1D Fractional ADE

Derivation of a Fokker–Planck equation (FPE) starts with a simple mathematical statement of how a random measure changes state from one moment to the next, after some event has occurred. In this case, we are interested in the probability that a particle has moved from location x_1 to x_3 in the time t_1 to t_3 , or

$p(x_3 - x_1, t_3 - t_1)$. This probability is referred to as the transition density, which is conditional on the initial position (x_1) and the time interval. The particle must move through an intermediate location x_2 , so this probability can be found by summing over all possible intermediate points x_2 . The probability of making both transitions (x_1 to x_2 to x_3) is the product of the single-transition probabilities, giving the Chapman–Kolmogorov equation (Sheshadri and West, 1982; Zaslavsky, 1994):

$$p(x_3 - x_1, t_3 - t_1) = \int p(x_3 - x_2, t_3 - t_2)p(x_2 - x_1, t_2 - t_1) dx_2. \quad (\text{A.1})$$

The particle's initial condition of being at position $x_1 = 0$ at time $t_1 = 0$ is placed into the density to give the probability of a particle's position at any time after moving for time (t):

$$p(x - x_0, t) = \int p(x - x_2, t - t_2)p(x_2 - x_0, t_2) dx_2. \quad (\text{A.2})$$

Note that, in general, the wandering particle is described by a joint space and time density (e.g., Shlesinger *et al.*, 1982; Klafter *et al.*, 1987; Berkowitz and Scher, 1995, 1998) that may be impossible to measure in a field situation. One can simplify by using the conditional space density for a specific time interval (Sheshadri and West, 1982; Zaslavsky, 1994; Benson, 1998; Meerschaert *et al.*, 1999) if one assumes that spatial dependence is much heavier-tailed than temporal correlation. This leads to the Markov assumption that the movement of a particle (a possibly very distant movement) is independent of past movements. We choose an equation that tracks the location of a particle at regular intervals of Δt , so the transition density $p(x - x_2 | \Delta t)$ is now specific to (conditioned on) regular intervals, since taking the limit as Δt goes to zero leads to a differential equation, while the former, in general, does not (Benson, 1998). The 'propagator', or Green function, is the density describing the position of a particle at any time (t) having started from $x_0 = 0$ and is given the shorter notation $p(x - x_0 | t - t_0) = P(x, t)$. This leads to the density of the particle position based on the present position and the transition density:

$$P(x, t + \Delta t) = \int p(x - \zeta | \Delta t)P(\zeta, t) d\zeta. \quad (\text{A.3})$$

Naturally, this is a convolution, since the density of a sum of two random variables (i.e., the distance traveled by a random walker from time zero to t and then again from time t to $t + \Delta t$) is the convolution of the densities of the summands (e.g., Ross, 1988). A requirement on the relative size of the total time (t) versus the transition time (Δt) is sometimes needed. For any general transition density, Δt must be much smaller than the total time, and Δt is called a *mixing time*. This argument is required so that an appropriate limit theorem such as the GCLT (1) can be invoked as a good approximation of the integration of a large number of transitions into a limit distribution. If the transitions are iid *stable* variables, then no

restrictions are placed on the transition time and the convolution is always satisfied. For example, iid Gaussian transitions result immediately in a Gaussian propagator.

By taking infinitesimal values of Δt in (A.3), we will know the change in $P(x, t)$ over a very short time period, leading to a differential equation. One should expect that a particle that travels along fractal paths and/or requires power-law times to complete individual trajectories (e.g., Compte, 1996) will have different limiting behavior than a typical Gaussian process. For example, Hughes *et al.* (1981) show that fractal motions on a lattice have a heavy-tailed transition density. The integer time derivative of $P(x, t)$ is represented:

$$\frac{\partial P(x, t)}{\partial t} = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left(\int p(x - \zeta | \Delta t) d\zeta - P(x, t) \right). \quad (\text{A.4})$$

The convolution is simplified by taking Fourier transforms:

$$\frac{d}{dt} \hat{P}(k, t) = \lim_{\Delta t \rightarrow 0} \frac{\hat{P}(k, t)}{\Delta t} (\hat{p}(k | \Delta t) - 1), \quad (\text{A.5})$$

where $\hat{p}(k | \Delta t)$ is the Fourier transform of the instantaneous particle transition density. Merely identifying a functional form for the instantaneous transition density and placing it in this equation gives an expression for the probability density of the particle at any place and time. Included in this functional form is the relationship of the density to time. A linear relationship is typically assumed, implying that a particle will travel twice as far in twice the time. Other relationships require a more complicated treatment with respect to time (e.g., fractional time derivatives, see Zaslavsky, 1994; Saichev and Zaslavsky, 1997; Benson, 1998). For nonreactive tracers, we shall assume this linear scaling. To illustrate, select a finite-variance transition density based on the cumulant expansion:

$$\hat{p}(k | \Delta t) = 1 + A_1(ik) + \frac{1}{2}A_2(ik)^2 + A_n(ik)^n + \dots \quad (\text{A.6})$$

where $A_1 = -v\Delta t + \sigma(\Delta t)$, $A_2 = \sigma^2\Delta t + \sigma(\Delta t)$, and $A_n = \sigma(\Delta t)$ for $n > 2$. The terms $\sigma(\Delta t)$ contain all trailing terms with powers of Δt greater than unity. The term σ^2 must have units of L^2/T so that the standard deviation of the jump size is L^2 . This requirement provides the linear link between the travel variance (of a particle or plume) and time. These coefficients define the requirements for the transition density that generates the classical ADE – it must have a finite-variance and higher-order moments that essentially disappear as Δt becomes very small. The third moment describes skewness, so highly skewed transitions (finite-variance or not) may not be well modeled by the ADE until time becomes very large and the transition time Δt become proportionally small. Placing (A.6) into (A.5) and taking the limit, we have

$$\frac{d}{dt} \hat{P}(k, t) = -v(ik) \hat{P}(k, t) + (ik)^2 \frac{\sigma^2}{2} \hat{P}(k, t) \quad (\text{A.7})$$

which is easily inverted using the property of Fourier transforms $\mathcal{F}^{-1}[(ik)^n \hat{g}(k)] = d^n/dx^n g(x)$:

$$\frac{\partial}{\partial t} P(x, t) = -v \frac{\partial}{\partial x} P(x, t) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} P(x, t) \quad (\text{A.8})$$

One can develop a parallel theory for the heavy-tailed (infinite variance) case, just as a GCLT (1) arose from the classic CLT. The Fourier transform of the Lévy density can be written:

$$\hat{f}(k) = \exp[q\mathcal{C}(-ik)^\alpha + p\mathcal{C}(ik)^\alpha - \mu ik], \quad \alpha \neq 1 \quad (\text{A.9})$$

where the relative weight of coefficients $p + q = 1$ describes the skewness, and \mathcal{C} is the scale (the spread similar to the variance in the Gaussian). The transition density scales linearly with the transition interval Δt for $\Delta t \rightarrow 0$. This stable transition density has the Fourier transform:

$$\hat{p}(k|\Delta t) = \exp[q\mathcal{D}\Delta t(-ik)^\alpha + p\mathcal{D}\Delta t(ik)^\alpha - v\Delta t(ik)], \quad \alpha \neq 1 \quad (\text{A.10})$$

with an expansion of $\hat{p}(k|\Delta t) = 1 - v\Delta t(ik) + q\mathcal{D}\Delta t(-ik)^\alpha + p\mathcal{D}\Delta t(ik)^\alpha + \sigma(\Delta t)$. This units of \mathcal{D} are $L^\alpha T^{-1}$. This density has an infinite variance for $\alpha < 2$. Placing the expansion into (A.5) and taking the limit:

$$\frac{d}{dt} \hat{P}(k, t) = \hat{P}(k, t)[-v(ik) + q\mathcal{D}(-ik)^\alpha + p\mathcal{D}(ik)^\alpha]. \quad (\text{A.11})$$

One can easily solve this eigenvalue ODE and find that it is the equation of a 1D Lévy motion (see Samorodnitsky and Taqqu, 1994). But this equation admits a remarkable form, based on the Fourier transform of a fractional derivative. The fractional derivative can be defined (Samko *et al.*, 1993; Debnath, 1995; Benson *et al.*, 1999a) by extending the well-known action of Fourier transform on integer derivatives $\mathcal{F}^{-1}[(ik)^n \hat{f}(k)] = (d^n/dx^n) f(x)$ to rational order: $\mathcal{F}^{-1}[(ik)^\alpha \hat{f}(k)] = (d^\alpha/dx^\alpha) f(x)$. In the forward sense,

$$\mathcal{F} \left[\frac{d^\alpha}{dx^\alpha} f(x) \right] = (ik)^\alpha \hat{f}(k).$$

If the derivative is taken with respect to $-x$, we substitute $-k$ for k :

$$\mathcal{F} \left[\frac{d^\alpha}{d(-x)^\alpha} f(x) \right] = (-ik)^\alpha \hat{f}(k).$$

By inverse transform (Samko *et al.*, 1993; Benson *et al.*, 1999a), one finds that the fractional derivative has the 1D representation (note the change of sign inside the integral):

$$\frac{d^q f(x)}{d(\pm x)^q} = \frac{(\pm 1)^n}{\Gamma(n - q)} \frac{d^n}{dx^n} \int_0^\infty \xi^{n-q-1} f(x \mp \xi) d\xi. \quad (\text{A.12})$$

The real-space inverse transform for the propagator (A.11) is now

$$\begin{aligned} \frac{\partial}{\partial t} P(x, t) = & -v \frac{\partial}{\partial x} P(x, t) + p \mathcal{D} \frac{\partial^\alpha}{\partial x^\alpha} P(x, t) + \\ & + q \mathcal{D} \frac{\partial^\alpha}{\partial (-x)^\alpha} P(x, t), \end{aligned} \tag{A.13}$$

or equivalently (Benson, 1998):

$$\frac{\partial P}{\partial t} = -v \frac{\partial P}{\partial x} + (1 + \beta) \frac{\mathcal{D}}{2} \frac{\partial^\alpha P}{\partial x^\alpha} + (1 - \beta) \frac{\mathcal{D}}{2} \frac{\partial^\alpha P}{\partial (-x)^\alpha}, \tag{A.14}$$

where the dimensions of \mathcal{D} (the dispersion coefficient) are $L^\alpha T^{-1}$. This is valid for light- or heavy-tailed particle excursions in 1D. We have introduced the description ‘fractional ADE’ with the understanding that only the dispersion term is correctly described by a fractional derivative. For a large number of independent solute ‘particles’, the probability propagator is replaced by ensemble concentration $C(x, t)$ (cf., Fürth, 1956; Bhattacharya and Gupta, 1990), and the general governing equation for the expected concentration solute movement (the 1D fractional ADE) simplifies to (7). For Gaussian or other light-tailed random motions, $\alpha = 2$ and the classical ADE is recovered.

A number of methods can be used to generate the α -stable density solutions (9). Denote the density $f_{\alpha\beta\sigma\delta}(x)$ where α is the index of stability, β is the skewness, $\sigma = [|\cos(\pi\alpha/2)|\mathcal{D}t]^{1/\alpha}$ is the scale, and $\delta = vt$ is the shift. A standard density has $\sigma = 1$ and $\delta = 0$. A general density is related to the standard by $f_{\alpha\beta\sigma\delta}(x) = \sigma^{-1} f_{\alpha\beta 10}((x - \delta)/\sigma)$. Given v and \mathcal{D} , one can specify any x and t , and find the standard density for the argument $(x - vt)/[|\cos(\pi\alpha/2)|\mathcal{D}t]^{1/\alpha}$. The canonical forms of the density assume many different definitions of a skewness parameter. Our choice of p is related to Samorodnitsky and Taqqu’s (1994) common skewness parameter (β) by $\beta = 2p - 1$. A standard density can be found by integrating the expression (Benson, 1998):

$$f_{\alpha\beta 10}(x) = \frac{|cx|^{1/(\alpha-1)} \alpha c}{2|1-\alpha|} \int_{-\theta}^1 U_\alpha(\phi, \theta) \exp(-|cx|^{\alpha/(\alpha-1)} U(\phi, \theta)) d\phi, \tag{A.15}$$

where

$$\begin{aligned} c = [1 + (\beta \tan(\pi\alpha/2))^2]^{-1/2\alpha}, \quad \theta = \frac{2}{\pi\alpha} \tan^{-1}(\beta \tan(\pi\alpha/2)), \\ U_\alpha(\phi, \theta) = \left(\frac{\sin \frac{1}{2}\pi\alpha(\phi + \theta)}{\cos \frac{1}{2}\pi\phi} \right)^{\alpha/(1-\alpha)}, \end{aligned}$$

or by Feller’s (1971) series for $1 < \alpha \leq 2$:

$$f_{\alpha\gamma 10}(x) = \frac{1}{\pi x} \sum_{k=0}^{\infty} \frac{\Gamma(k\alpha^{-1} + 1)}{k!} (-x)^k \sin \frac{\pi k}{2\alpha} (\gamma - \alpha), \tag{A.16}$$

where Feller's skewness parameter γ is obtained from β by Samorodnitsky and Taqqu (1994):

$$\gamma = \frac{2}{\pi} \arctan[\beta \tan(\pi(\alpha - 2)/2)]. \quad (\text{A.17})$$

For symmetric densities, setting $\beta = \gamma = 0$ in (A.16) yields a formula that converges rapidly even for large arguments. The symmetric standard density for $1 < \alpha \leq 2$ is given by the series:

$$f_{\alpha 010}(x) = \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \Gamma\left(\frac{2k+1}{\alpha} + 1\right) x^{2k}. \quad (\text{A.18})$$

Appendix B. Estimation of Power-Law Content

Several numerical and graphical routines are typically used to estimate the 'power-law' content, if any, in data sets. Here we discuss the basis for each estimator and the pitfalls encountered when using each on different types of data. Each estimator is applied to the MADE K data in raw and incremental form, which has serious effect on the stability of some estimators.

The family of stable distributions is parameterized by the tail thickness $0 < \alpha \leq 2$, the skewness $-1 \leq \beta \leq 1$, the location $-\infty < \mu < \infty$, and the scale $\sigma > 0$ (Equation (3)). There are several established methods for stable parameter estimation, which can be used to fit a stable probability distribution to a data set. McCulloch (1986) tabulates quantiles of a stable distribution as a function of the parameters, and uses these tables to interpolate the best-fitting parameter values. These rough estimates can also be used to initialize the optimization for maximum likelihood estimation (MLE). Although the stable density cannot be written in closed form, Nolan (1997) has developed fast numerical approximations that allow practical computation of the stable MLE on a personal computer. A computer program to compute both the McCulloch quantile estimators and the MLE, including diagnostics to check for distributional fit, is available at <http://www.cas.american.edu/~jpnolan/>.

Anderson and Meerschaert (1998) review parameter estimation methods for heavy tails and apply them to a river flow time series. For data sets that are known to be stable, the MLE is the best parameter estimation method. Robust estimators are more appropriate when the exact probability distribution of the data is unknown. The most popular robust estimator was introduced by Hill (1975). Hill's estimator is the conditional MLE for a Pareto distribution $P[X > x] = Wx^{-\alpha}$ based on the largest order statistics. This model is appropriate because the upper tail of a stable distribution is approximately Pareto with the same index (α) (e.g., Samorodnitsky and Taqqu, 1994). A recent extension of Hill's estimator by Aban and Meerschaert (1999) fits a shifted Pareto $P[X > x] = W(x-s)^{-\alpha}$ using the same method. Since Hill's original estimator is not shift-invariant, this new method is more reliable.

Table II. Estimates of K and ΔK tail parameter α using four estimators. Due to the robust methodology of the Meerschaert and Scheffler (M&S) method, a value of $\alpha = 1.1$ is thought to be reliable.

Estimator (type)	McCulloch (quantile stable)	Nolan (MLE stable)	Hill (MLE Pareto)	M&S (asymptotics)
Raw K	0.8	0.4	3.0	1.1
K increments	0.7	0.8	1.1	1.1

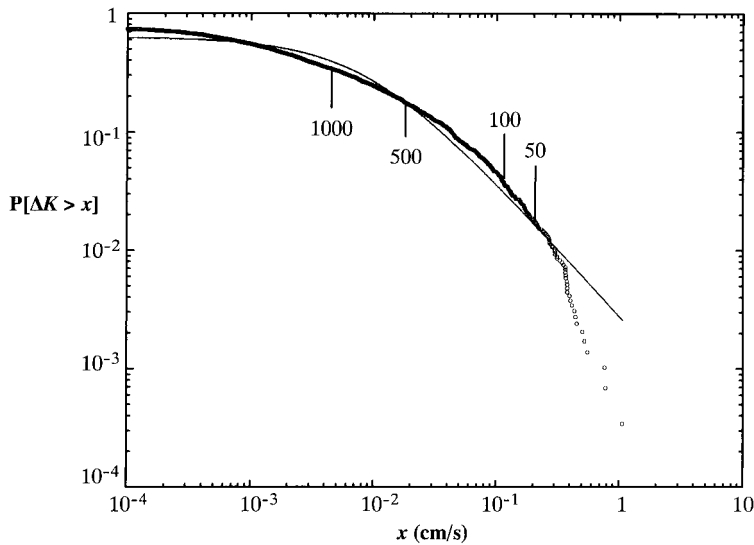


Figure 8. Mandelbrot plot of measured K increments at 15 cm lag ($|\Delta K|$) and shifted Pareto $P[|\Delta K| > x] = 0.0028(x + 0.01)^{-1.1}$ (solid line). Numbers on plot indicate rank of the datum.

Hosking and Wallis (1987) suggest using the MLE for the generalized Pareto distribution to fit hydrologic data. This two parameter model is equivalent to a Pareto with a specific shift that makes X range over the positive reals. McCulloch (1997) shows that Hill's estimator can yield misleading results when applied to stable data. Meerschaert and Scheffler (1998) suggest another robust estimator based on the asymptotics of the sum. This estimator performs about as well as Hill's estimator in most practical situations, and it performs best in exactly those situations (stable or nearly stable data with alpha near two) in which Hill's estimator is most likely to fail.

For the untransformed MADE K data the McCulloch quantile estimator gives $\alpha = 0.8$ and $\beta = 1$ (totally skewed to the right). Nolan's MLE gives $\alpha = 0.4$ and $\beta = 1.0$, indicating very heavy tails. A probability plot indicates significant deviation from the best fitting stable model, indicating that the data are probably not drawn from a stable distribution. This is a situation in which robust estimators

are more appropriate. The generalized Pareto MLE yields $\alpha = 0.5$ but once again a probability plot indicates that there is not a good fit to this particular distribution. Hill's estimator uses the largest r observations out of $n = 2915$ for this data set. The resulting parameter estimates vary with r , for example we get $\alpha = 2.3$ at $r = 50$ and $\alpha = 1.0$ at $r = 500$, and in general, α steadily decreases with r , indicating a possible distributional misspecification. For the Pareto model $y = P[X > x] = Wx^{-\alpha}$ leads to $\ln(y) = \ln(W) - \alpha \ln(x)$. If we sort the data $X_1 \dots X_n$ in descending order to obtain the order statistics $X(1) \geq \dots \geq X(n)$, then we can estimate $y = i/n$ and $x = X(i)$. A log-log plot of i/n versus $X(i)$ should show a straight line with slope $-\alpha$ for Pareto data. This method was pioneered by Mandelbrot (1963) in his early work on heavy tails, and we will call this the Mandelbrot plot. Simple linear regression can be used to estimate (α) from the Mandelbrot plot, but this method typically performs poorly. More importantly, this graphical method can be used to test for fit to a Pareto model. The choice of r for Hill's estimator, based on the Pareto distribution, should be small enough that the Mandelbrot plot of the r largest order statistics appears linear. Since Hill's estimator is not shift-invariant, it is also reasonable to shift the data until the plot appears linear. This amounts to fitting a shifted Pareto. Visual inspection indicates that a shift of $s = -0.2$ cm/s (add 0.2 cm/s to each K observation) gives a Mandelbrot plot that appears linear, and Hill's estimator for the largest $r = 375$ observations yields $\alpha = 3$. The more exact shifted Hill's estimator (conditional MLE) yields $\alpha = 3.019$ at the optimal shift of $s = -0.1998$, so in this case, visual inspection is adequate. Overall, there is a wide discrepancy in α estimates (Table II), due to mismatches of estimation technique with the data distribution. Estimators like the MLE, which are tailored to a specific distributional form, are not robust with respect to deviations from that probability distribution. After rescaling the data (divide by the median), the robust estimator of Meerschaert and Scheffler yields $\alpha = 1.1$ and in this case we consider this our best parameter estimate.

To test the fits to a specific distribution, we also analyzed the variations in K . It is typical in econometrics to difference the data in this way (subtract one observation from the next) to unmask the effects of heavy tails. Theoretically speaking, the differences should exhibit the same tail thickness parameter (α) as the original data set (e.g., Davis and Resnick, 1985). For our data set, the McCulloch quantile method gives $\alpha = 0.7$ and $\beta = 1$ and Nolan's MLE gives $\alpha = 0.8$ and $\beta = 1$ for the differenced data, but once again, diagnostic plots indicate a significant deviation from a stable distribution. Meerschaert and Scheffler's robust estimator (again we rescale by the median) yields $\alpha = 1.1$, the same as for the raw data (Table I). Hill's estimator with $r = 150$ also gives $\alpha = 1.1$ and the Mandelbrot plot indicates that the fit to a shifted Pareto is good (Figure 8). In addition, Hill's estimator gives a scale coefficient $W = 0.0028$ (cm/s)^{1.1}. Since the conditional MLE of Aban and Meerschaert indicates that the optimal shift is near zero, the usual Hill's estimator should give good results. The generalized Pareto MLE is not reported here because

we could not obtain convergence. Overall, it seems safe to say that the K data fit a heavy-tail model with $\alpha = 1.1$ indicating an infinite variance but a finite mean.

References

- Aban, I. and Meerschaert, M.: 1999, Shifted Hill's estimator for heavy tails, Preprint.
- Adams, E. E. and Gelhar, L. W.: 1992, Field study of dispersion in a heterogeneous aquifer, 2, Spatial moments analysis, *Water Resour. Res.* **28**(12), 3293–3307.
- Anderson, P. and Meerschaert, M. M.: 1998, Modeling river flows with heavy tails, *Water Resour. Res.* **34**(9), 2271–2280.
- Benson, D. A.: 1998, The fractional advection–dispersion equation: Development and application, Unpublished Ph.D. thesis, Univ. of Nevada, Reno, <http://www.hydro.unr.edu/homepages/benson>.
- Benson, D. A., Wheatcraft, S. W. and Meerschaert, M. M.: 1999a, The fractional-order governing equation of Lévy motion, Preprint, <http://www.hydro.unr.edu/homepages/benson>.
- Benson, D. A., Wheatcraft, S. W. and Meerschaert, M. M.: 1999b, Application of a fractional advection–dispersion equation, Preprint, <http://www.hydro.unr.edu/homepages/benson>.
- Berkowitz, B. and Scher, H.: 1995, On characterization of anomalous dispersion in porous and fractured media, *Water Resour. Res.* **31**(6), 1461–1466.
- Berkowitz, B. and Scher, H.: 1998, Theory of anomalous chemical transport in random fracture networks, *Phys. Rev. E* **57**(5), 5858–5869.
- Bhattacharya, R. and Gupta, V. K.: 1990, Application of central limit theorems to solute transport in saturated porous media: from kinetic to field scales, Chapter IV, in: J. H. Cushman (ed.), *Dynamics of Fluids in Hierarchical Porous Media*, Academic Press.
- Boggs, J. M. and Adams, E. E.: 1992, Field study of dispersion in a heterogeneous aquifer, 4; Investigation of adsorption and sampling bias, *Water Resour. Res.* **28**(12), 3325–3336.
- Boggs, J. M., Beard, L. M., Long, S. E. and McGee, M. P.: 1993, Database for the second macrodispersion experiment (MADE-2), EPRI report TR-102072, Electric Power Res. Inst., Palo Alto, CA.
- Brusseu, M.: 1992, Transport of rate-limited sorbing solutes in heterogeneous porous media: Application of a one-dimensional multifactor nonideality model to field data, *Water Resour. Res.* **28**(9), 2485–2497.
- Compte, A.: 1996, Stochastic foundations of fractional dynamics, *Phys. Rev. E* **53**(4), 4191–4193.
- Compte, A. and Caceres, M. O.: 1998, Fractional dynamics in random velocity fields, *Phys. Rev. Lett.* **81**, 3140–3143.
- Crank, J.: 1975, *The Mathematics of Diffusion*, Oxford University Press, Oxford, Great Britain.
- Dagan, G.: 1984, Solute transport in heterogeneous porous formations, *J. Fluid Mech.* **145**, 151–177.
- Davis, R. and Resnick, S.: 1985, Limit theory for moving averages of random variables with regularly varying tail probabilities, *Ann. Probab.* **13**, 179–195.
- Debnath, L.: 1995, *Integral Transforms and Their Applications*, CRC Press, New York.
- Deng, F.-W., Cushman, J. H. and Delleur, J. W.: 1993, A fast Fourier transform stochastic analysis of the contaminant transport problem, *Water Resour. Res.* **29**(9), 3241–3247.
- Einstein, A.: 1908, *Investigations on the Theory of the Brownian Movement*, translation by Dover Publications in 1956 of the original manuscript.
- Feller, W.: 1971, *An Introduction to Probability Theory and Its Applications*, Volume II, 2nd ed., Wiley, New York.
- Fofack, H. and Nolan, J.: 1998, Tail behavior, modes and other characteristics of stable distributions, Preprint, <http://www.cas.american.edu/~jpnolan/>.
- Fogedby, H. C.: 1994, Lévy flights in random environments, *Phys. Rev. Lett.* **73**(19), 2517–2520.

- Freyberg, D. L.: 1986, A natural gradient experiment on solute transport in a sandy aquifer, 2, Spatial moments and the advection and dispersion of nonreactive tracers, *Water Resour. Res.* **22**(13), 2031–2046.
- Fürth, R.: 1956, Notes in: Einstein, A. E., *Investigations on the Theory of the Brownian Movement*, translation by Dover Publications.
- Gelhar, L. W. and Axness, C. L.: 1983, Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.* **19**(1), 161–180.
- Gnedenko, B. V. and Kolmogorov, A. N.: 1954, *Limit Distributions for Sums of Random Variables*, Addison-Wesley, Reading, Mass.
- Gorenflo, R. and Mainardi, F.: 1998, Fractional calculus and stable probability distributions, *Arch. Mech.* **50**(3), 377–388.
- Grigolini, P., Rocco, A. and West, B. J.: 1999, Fractional calculus as a macroscopic manifestation of randomness, *Phys. Rev. E* **59**, 2603.
- Haggerty, R. and Gorelick, S. M.: 1995, Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity, *Water Resour. Res.* **31**(10), 2383–2400.
- Hill, B.: 1975, A simple general approach to inference about the tail of a distribution, *Ann. Statist.* **3**, 1163–1173.
- Hosking, J. and Wallis, J.: 1987, Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics* **29**, 339–349.
- Hughes, B. D., Shlesinger, M. F. and Montroll, E. W.: 1981, Random walks with self-similar clusters, *Proc. Natl. Acad. Sci. USA* **78**(6), 3287–3291.
- Janicki, A. and Weron, A.: 1994, Can one see α -stable variables and processes?, *Stat. Sci.* **9**(1), 109–126.
- Klafter, J., Blumen, A. and Shlesinger, M. F.: 1987, Stochastic pathway to anomalous diffusion, *Phys. Rev. A* **35**(7), 3081–3085.
- LeBlanc, D. R., Garabedian, S. P., Hess, K. M., Gelhar, L. W., Quadri, R. D., Stollenwerk, K. G. and Wood, W. W.: 1991, Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 1, Experimental design and observed tracer movement, *Water Resour. Res.* **27**(5), 895–910.
- Lévy, P.: 1937, *Théorie de L'addition des Variables Aléatoires*, Gauthier-Villars, Paris.
- Liu, H. H. and Molz, F. J.: 1997a, Comment on 'Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations' by Scott Painter, *Water Resour. Res.* **33**(4) 907–908.
- Liu, H. H. and Molz, F. J.: 1997b, Multifractal analyses of hydraulic conductivity distributions, *Water Resour. Res.* **33**(11) 2483–2488.
- Mandelbrot, B.: 1963, The variation of certain speculative prices, *J. Business* **36**, 394–419.
- Mantegna, R. N. and Stanley, H. E.: 1995, Ultra-slow convergence to a Gaussian: the truncated Lévy flight, in: M. F. Shlesinger, G. M. Zaslavsky and U. Frisch (eds), *Lévy Flights and Related Topics in Physics*, Springer-Verlag, pp. 301–312.
- McCulloch, J. H.: 1986, Simple consistent estimators of stable distribution parameters, *Comm. Statist. Simul. Comput.* **15**, 1109–1136.
- McCulloch, J. H.: 1997, Measuring tail thickness to estimate the stable index alpha: A critique, *J. Business Econ. Statist.* **15**, 74–81.
- Meerschaert, M.: 1986, Regular variation and domains of attraction in \mathbb{R}^k , *Stat. Prob. Lett.* **4**, 43–45.
- Meerschaert, M. and Scheffler, H.-P.: 1998, A simple robust estimator for the thickness of heavy tails, *J. Stat. Plann. Inference* **71**(1–2), 19–34.
- Meerschaert, M. M., Benson, D. A. and Bäumer, B.: 1999, Multidimensional advection and fractional dispersion, *Phys. Rev. E* **59**(5) 5026–5028.
- Metzler, R., Klafter, J. and Sokolov, I. M.: 1998, Anomalous transport in external fields: Continuous time random walks and fractional diffusion equations extended, *Phys. Rev. E* **58**, 1621–1633.
- Metzler, R., Barkai, E. and Klafter, J.: 1999, Deriving fractional Fokker–Planck equations from a generalized master equation, *Europhys. Lett.* **46**, 431–436.

- Molz, F. J., Liu, H. H. and Szulga, J.: 1997, 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.
- Nolan, J.: 1997, Numerical calculation of stable densities and distribution functions: Heavy tails and highly volatile phenomena, *Comm. Statist. Stoch. Models* **13**, 759–774.
- Nolan, J.: 1998, Parameterizations and modes of stable distributions, *Statist. Probab. Lett.* **38**(2), 187–195.
- Oldham, K. B. and Spanier, J.: 1974, *The Fractional Calculus*, Academic Press, New York.
- Pachevsky, Y. A.: 1998, Transport of water and chemicals in soils as in fractal media, *Agronomy Abstracts*, p. 202.
- Pachevsky, Y. A., Benson, D. A. and Rawls, W.: 1999, Simulating scale-dependent solute transport in soils with the fractional advective-dispersive equation, Preprint.
- Painter, S.: 1996a, Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations, *Water Resour. Res.* **32**(5), 1183–1195.
- Painter, S.: 1996b, Stochastic interpolation of aquifer properties using fractional Lévy motion, *Water Resour. Res.* **32**(5), 1323–1332.
- Painter, S.: 1997, Reply to comment on ‘Evidence for non-Gaussian scaling behavior in heterogeneous sedimentary formations’ by H. H. Liu and F. J. Molz, *Water Resour. Res.* **33**(4) 909–910.
- Rajaram, H. and Gelhar, L. W.: 1991, Three-dimensional spatial moments analysis of the Borden tracer test, *Water Resour. Res.* **27**(6), 1239–1251.
- Rehfeldt, K. R., Boggs, J. M. and Gelhar, L. W.: 1992, Field study of dispersion in a heterogeneous aquifer. 3: Geostatistical analysis of hydraulic conductivity, *Water Resour. Res.* **28**(12), 3309–3324.
- Rocco, A. and West, B. J.: 1999, *Physica A* **265**, 535.
- Ross, S.: 1988, *A First Course in Probability*, 5th ed., Prentice Hall, NY.
- Saichev, A. I. and Zaslavsky, G. M.: 1997, Fractional kinetic equations: solutions and applications, *Chaos* **7**(4), 753–764.
- Samko, S. G., Kilbas, A. A. and Marichev, O. I.: 1993, *Fractional Integrals and Derivatives: Theory and Applications*, Gordon and Breach.
- Samorodnitsky, G. and Taqqu, M. S.: 1994, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York.
- Schumer, R., Benson, D., Meerschaert, M. and Wheatcraft, S.: 1999, A physical derivation of the fractional advection–dispersion equation, Preprint, <http://www.hydro.unr.edu/homepages/benson>.
- Serrano, S. E.: 1995, Forecasting scale-dependent dispersion from spills in heterogeneous aquifers, *J. Hyd.* **169**, 151–169.
- Sheshadri, V. and West, B. J.: 1982, Fractal dimensionality of Levy processes, *Proc. Natl. Acad. Sci.* **79**, 4501–4505.
- Shlesinger, M. F., Klafter, J. and Wong, Y. M.: 1982, Random walks with infinite spatial and temporal moments, *J. Stat. Phys.* **27**(3), 499–512.
- Sudicky, E. A.: 1986, A natural gradient experiment on solute transport in a sandy aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process, *Water Resour. Res.* **22**(13), 2069–2082.
- Taylor, Sir, G. I.: 1953, Dispersion of soluble matter in solvent flowing slowly through a tube, *Proc. R. Soc. A., London* **219**, 186–203.
- Zaslavsky, G. M.: 1994, Renormalization group theory of anomalous transport in systems with Hamiltonian chaos, *Chaos* **4**(1), 25–33.
- Zheng, C. and Jiao, J. J.: 1998, Numerical simulation of tracer tests in heterogeneous aquifer, *J. Environ. Eng.* **124**(6), 510–516.