Fractional advection-dispersion equations for modeling transport at the Earth surface

Rina Schumer

Desert Research Institute, Reno, NV, USA

Mark M. Meerschaert

Michigan State University, East Lansing, MI, USA

Boris Baeumer

University of Otago, Dunedin, NZ

Rina Schumer, Division of Hydrologic Sciences, Desert Research Institute, Reno, NV, 89512, USA.(rina@dri.edu)

Mark M. Meerschaert, Department of Statistics and Probability, Michigan State University, East Lansing, MI, 48824, USA. (mcubed@stt.msu.edu)

Boris Baeumer, Department of Mathematics and Statistics, University of Otago, Dunedin, NZ(bbaeumer@maths.otago.ac.nz)

Abstract. Characterizing the collective behavior of particle transport on the Earth surface is a key ingredient in describing landscape evolution. We seek equations that capture essential features of transport of an ensemble of particles on hillslopes, valleys, river channels, or river networks, such as mass conservation, super-diffusive spreading in flow fields with large velocity variation, or retardation due to particle trapping. Development of stochastic partial differential equations such as the advection-dispersion equation (ADE) begin with assumptions about the random behavior of a single particle: possible velocities it may experience in a flow field and the length of time it may be immobilized. When assumptions underlying the ADE are relaxed, a fractional ADE can arise, with a non-integer order derivative on time or space terms. Fractional ADEs are non-local; they describe transport affected by hydraulic conditions at a distance. Space fractional ADEs arise when velocity variations are heavy-tailed and describe particle motion that accounts for variation in the flow field over the entire system. Time fractional ADEs arise as a result of power law particle residence time distributions and describe particle motion with memory in time. Here we present a phenomenological discussion of how particle transport behavior may be parsimoniously described by a fADE, consistent with evidence of super-diffusive and sub-diffusive behavior in natural and experimental systems.

1. Introduction

An important class of problems in the Earth-surface sciences involves describing the collective behavior of particles in transport. Familiar examples include transport of solute and contaminant particles in surface and subsurface water flows, the behavior of soil particles and associated soil constituents undergoing biomechanical transport and mixing by bioturbation, and the transport of sediment particles and sediment-borne substances in turbulent shear flows, whether involving shallow flows over soils, deeper river flows, or ocean currents. These and many other examples share three essential features. First, is the behavior of a well defined ensemble of particles. These particles may be considered tracers whose total mass is conserved or otherwise accounted for if radioactive decay, physical transformations or chemical reactions are involved. Second, these tracers typically alternate between states of motion and rest over many time scales, and indeed, most tracers of interest in Earth-surface systems are at rest much of the time. Third, when in transport, some tracers move faster, and some move slower, than the average motion due to spatiotemporal variations in the mechanisms inducing their motion. Tracer motions thus may be considered as consisting of quasi-random walks with rest periods.

During motion, tracers may experience characteristically different 'hop' or 'jump' behaviors. Consider releasing at some instant an ensemble of tracers within a system, say, marked particles within a river. Assuming all tracers undergo and remain in motion, then due to turbulence fluctuations in the case of suspended particles, or to effects of near-bed turbulence excursions together with tracer-bed interactions in the case of bedload particles, during a small interval of time Δt, most of the tracers move short distances

⁴⁷ downstream whereas a few move longer distances. The displacement (or hop) distance

⁴⁸ Δx generally may be considered a random variable. If the probability density function

⁴⁹ (pdf) describing hop length has a form that declines at least as fast as an exponential

⁵⁰ distribution, then it possesses finite mean and variance. Moreover, after finite time t,

⁵¹ the center of mass of the tracers is centered at the position $x = x_0 + vt$, where x_0 is

⁵² the starting position and v is the mean tracer speed. The spread or 'dispersion' around

⁵³ average tracer position, described by standard deviation, grows at a constant rate with

⁵⁴ time $\sigma = (Dt)^{1/2}$ where D is the dispersion coefficient. As described below, this behavior

⁵⁵ know as Fickian or Boltzmann scaling may be characterized as being 'local', in that dur-

⁵⁶ ing a small interval, tracers in motion mostly move from location x to nearby positions.

⁵⁷ Conversely, during the same interval, tracers arriving at position x mostly originate from

⁵⁸ nearby positions. Letting $C(x, t)$ denote the local tracer concentration, then in this local

⁵⁹ case the behavior of C may be adequately described by the classic advection-dispersion

⁶⁰ equation (ADE).

⁶¹ Consider, in contrast, the possibility that the hop length pdf declines with distance in

⁶² such a way that the variance is undefined, for example, a power-law distribution. In this

⁶³ case, long displacement distances, albeit relatively rare, are nonetheless more numerous

⁶⁴ than with, say, an exponential distribution - so that hop length density possesses a so-

⁶⁵ called 'heavy tail'. For example, during downstream transport, sediment tracers dispersed

⁶⁶ by turbulent eddies and by momentary excursions into parts of the mean flow that are

⁶⁷ either faster or slower than the average may occasionally experience super dispersive events

⁶⁸ during excursions into sites with unusually high flow velocities [*Bradley et al.*, 2009]. In

⁶⁹ these situations where the hop length density is heavy-tailed, the scaling of dispersion

is non-Fickian, that is, $\sigma = (Dt)^{1/\alpha}$, where $1 < \alpha < 2$. Because σ therefore grows

at a rate faster than 'normal' Fickian dispersion, this behavior is referred to as 'super-

diffusive'. Moreover, this behavior may be characterized as being 'nonlocal' in that during

a small interval, tracers released from position x mostly move to nearby positions, but

also involve an 'unusually large' number of motions to positions far from x. Conversely,

during the same interval, tracers arriving at position x mostly originate from nearby

positions, but also involve a significant number or motions originating far from x. This

means that the behavior at a particular position does not necessarily depend only on local

(nearby) conditions, but rather may also depend on conditions upstream or downstream.

For example, because the local rate of change in tracer concentration consists of the

divergence of the flux of tracers, then if the flux is related to the bed stress, say, in the

case of sediment transport, the possibility exists that changes in the local concentration

depend on stress conditions and therefore on system configuration 'far' upstream. As

described below, the behavior of the tracer concentration $C(x, t)$ in this 'nonlocal' case

may be described by a fractional derivative version of the advection-dispersion equation.

Consider now an ensemble of tracers whose motions, after release, involve states of

motion and rest. In the case of tracer particles moving as bedload, rest states might involve

momentary disentrainment on the bed surface associated with turbulence fluctuations, or

burial beneath the surface for longer periods followed by reentrainment with scour [*Nikora

et al.*, 2002; *Parker et al.*, 2000]. Even longer rest periods might consist of storage of

tracers within the bars and floodplain of a river [*Malmon et al.*, 2002, 2003]. Similarly,

rest times of nutrients in open-channel flow might consist of temporary excursions into

so-called dead zones, or into hyporheic zone storage, with eventual release back into the

main fluid column [*Haggerty et al.*, 2002; *Gooseff et al.*, 2003]. With bioturbation of soil particles, rest times consist of the intervals between displacement by biotic activity. The rest time (or waiting time, or residence time) may be considered a random variable. If random waiting times are, in general, small, then the speed V may be considered an average virtual speed that includes periods of motion and rest, and the local rate of change in tracer concentration, $\frac{\partial C}{\partial t}$, in the ADE accommodates this. If, in contrast the waiting time density is a heavy-tailed distribution that reflects the presence of unusually long rest periods for tracers in storage, then as described below a suitable phenomenological description of the behavior of the tracer concentration $C(x,t)$ may be obtained from a fractional ADE (fADE) with a fractional time term $\frac{\partial^{\gamma} C}{\partial t^{\gamma}}$, where $0 < \gamma < 1$. Dispersion now scales at a rate proportional to $t^{\gamma/2}$. Because σ grows at a rate slower than normal Fickian dispersion, this behavior is referred to as sub-diffusion. This characterizes a long memory effect, analogous to an influence from 'far' upstream as described above with reference to the fractional space derivative; the fractional time derivative characterizes an influence from far back in time. Like variations in tracer speed, rest times contribute to dispersion. As such, these effects may not be separable without ancillary information.

As elaborated below, if both the hop length density and the waiting time density are heavy-tailed, then the fADE can include both a time and a space fractional derivative. Dispersion now scales at a rate proportional to $t^{\gamma/\alpha}$. Note that a particular value of the ratio $\frac{\gamma}{\alpha}$ is a non-unique combination of values of α and γ. This means that if the ratio is obtained empirically from measurements of tracer dispersion, say, in a flume or river, additional information empirical and/or theoretical is needed to distinguish whether the source of the anomalous dispersion resides with the hop lengths, waiting times, or both.

On this note we point out that the development below provides a mostly phenomenological viewpoint of how tracer behavior may be parsimoniously described by a fADE, consistent with evidence of super-diffusive and sub-diffusive behavior in natural and experimental systems (e.g. *Benson et al.* [2001, 2000]; *Haggerty et al.* [2002]; *Schumer et al.* [2003]). A particularly exciting opportunity resides in clarifying the underlying physics that lead to these behaviors, including the mechanisms that give rise to heavy-tail distributions of hop distances and waiting times, perhaps building on recent insights from the hydrologic literature. In the following sections, we look at details underlying the application of fractional advection-dispersion equations for describing collective behavior of particle transport at the Earth surface. Much of this manuscript is a review of basic information about fADE for non-specialists that is not available in a single reference. New material fills in theoretical gaps, compares solution characteristics and application of various fADEs, and links landscape evolution processes to the conceptual models underlying fADEs.

2. Fractional Derivatives

Before presenting fractional ADEs, we provide some useful introductory material about fractional derivatives. Three characteristics of fractional derivatives will be used to interpret and solve fractional ADEs: (1) the Grunwald fractional difference quotient that approximates a fractional derivative; (2) the fractional derivative as the convolution of an integer-order derivative with a memory function; and (3) the Fourier and Laplace transforms of fractional derivatives.

2.1. Grunwald finite difference

How can we generalize our understanding of derivatives to include non-integer order cases? The first derivative of a function $y = f(x)$ is approximated by the difference quotient $\Delta y / \Delta x$ where $\Delta x = h$ and $\Delta y = f(x) - f(x - h)$ (Fig. 1(a)). The second derivative $f''(x)$ is approximated by $\Delta^2 y / \Delta x^2$ where

$$\Delta^2 y = \Delta[f(x) - f(x - h)]$$

$$= [f(x) - f(x - h)] - [f(x - h) - f(x - 2h)] \tag{1}$$

$$= f(x) - 2f(x - h) + f(x - 2h)$$

the second difference. Continuing in this manner shows that the nth order derivative is approximated by the nth finite difference quotient

$$\frac{d^n}{dx^n} f(x) \approx \frac{1}{h^n} \sum_{j=0}^{n} \binom{n}{j} (-1)^j f(x - jh) \tag{2}$$

where the binomial coefficients (from Pascal's triangle)

$$\binom{n}{j} = \frac{n!}{j!(n-j)!} = \frac{\Gamma(n+1)}{\Gamma(j+1)\Gamma(n-j)}.$$

The Grunwald definition of the fractional derivative is the non-integer variant of (2)

$$\frac{d^\alpha}{dx^\alpha} f(x) \approx \frac{1}{h^\alpha} \sum_{j=0}^{\infty} \binom{\alpha}{j} (-1)^j f(x - jh) \tag{3}$$

where the fractional binomial coefficients are

$$\binom{\alpha}{j} = \frac{\Gamma(\alpha+1)}{\Gamma(j+1)\Gamma(\alpha-j)}$$

and the approximation becomes exact as $h = \Delta x \to 0$. To see that the integer order formula (2) is really a special case of the fractional formula (3), note that the binomial coefficients are $0 = 1/\infty$ for $j > n$, so that we could have written the sum in (2) with the upper limit of infinity.

Several interesting properties of the fractional derivative follow from the Grunwald definition (3). Most important is that a fractional derivative is a non-local operator that depends on function values far away from x (Fig. 1(b)). The importance given to a far away point $x - jh$ is measured by the Grunwald weight $w_j = (-1)^j \Gamma(\alpha + 1) / (\Gamma(j + 1) \cdot \Gamma(\alpha - j))$ so that $w_0 = 1$, $w_1 = -\alpha$, $w_2 = \alpha(\alpha - 1)/2$ and $w_{j+1} = w_j \cdot (j - \alpha)/(j + 1)$. An appeal to Stirling's approximation [*Meerschaert and Scheffler*, 2002a] shows that $w_j \approx j^{-(1+\alpha)}/\Gamma(-\alpha)$ (Figure 1). Hence the Grunwald definition is essentially a discrete convolution with a power law. The finite difference formula (3) is also the basis for numerical codes to solve the fractional ADE [*Tadjeran et al.*, 2006].

2.2. Convolution of the derivative with a power-law memory function

The continuum limit of the Grunwald finite difference formula is a convolution integral (see for example *Meerschaert and Tadjeran* [2004]). Writing (3) in terms of the Grunwald weights and using $w_j \approx j^{-(1+\alpha)}/\Gamma(-\alpha)$ shows that

$$
\begin{aligned}
\frac{d^\alpha}{dx^\alpha} f(x) &\approx \frac{1}{\Gamma(-\alpha)} \sum_{j=0}^{\infty} f(x - jh)(jh)^{-\alpha-1} h \\
&\approx \frac{1}{\Gamma(-\alpha)} \int_0^\infty f(x - y) y^{-\alpha-1} \, dy
\end{aligned}
\tag{4}
$$

which can be written in various equivalent forms [*Samko et al.*, 1993; *Mainardi*, 1997]. If $f(t)$ is defined on $t \geq 0$ then integration by parts yields the Caputo form

$$
\frac{d^\alpha}{dt^\alpha} f(t) = \frac{1}{\Gamma(1 - \alpha)} \int_0^t f'(t - y) y^{-\alpha} \, dy
\tag{5}
$$

which is very useful for fractional time derivatives. Noting that the integral (5) is a convolution we can write

$$
\frac{d^\alpha f(t)}{dt^\alpha} = \frac{df(t)}{dt} \star \frac{t^{-\alpha}}{\Gamma(1 - \alpha)}
\tag{6}
$$

which relates the fractional derivative to a power-law memory function.

2.3. Transforms of fractional derivatives

For partial differential equations, it is useful to use Fourier transforms in space and Laplace transforms in time. Since the transform of a power law is another power law, it is easy to check using (4) and (6) the Fourier and Laplace transforms pairs:

$$\frac{d^{\alpha} f(x)}{dx^{\alpha}} \leftrightarrow (ik)^{\alpha} \hat{f}(k)$$
$$\frac{d^{\gamma} f(t)}{dt^{\gamma}} \leftrightarrow s^{\gamma} \tilde{f}(s) - s^{\gamma-1} f(0) \tag{7}$$

using $t^{-\gamma}/\Gamma(1-\gamma) \leftrightarrow s^{\gamma-1}$.

2.4. Examples of fractional derivatives

Use the Laplace transform pair $t^{-\gamma}/\Gamma(1-\gamma) \leftrightarrow s^{\gamma-1}$ along with the Laplace transform (7) to check that the α-order fractional derivative of t^p is $t^{p-\alpha}\Gamma(p+1)/\Gamma(p+1-\alpha)$ which reduces to the usual form when α is an integer. The Fourier transform (7) can be used to show that the α order fractional derivative of $\sin x$ is $\sin(x + \pi\alpha/2)$, which reduces to the usual formula for integer α using standard trigonometric identities. The α order fractional derivative of e^{bx} is $b^{\alpha}e^{bx}$ and so forth. The point is that fractional derivatives are natural analogues of their integer order cousins.

3. The classical random walk, central limit theorem, and the ADE

Sediment transport has long been modeled as a stochastic sequence of hops and rest periods [*Einstein*, 1937; *Sayre and Hubbell*, 1965] leading to emergent properties after long time. It is also common to represent Earth surface transport process using diffusion-type models arising from the combination of a continuity equation with transport (flux) laws (since *Culling* [1960]). Random walk and diffusion models are related. Specifically, diffusion equations govern the stochastic processes that arise when the scaling limit of

random walk is taken. Here we describe the type of particle motion that may be well-modeled by a classical ADE so that the exact nature of our fractional generalizations will be clear. We can approach the advection-dispersion equation from either a Lagrangian or Eulerian point of view. By using a Lagrangian (following the progress of an individual particle with time) approach, the exact nature of the transport process is clarified.

3.1. Lagrangian approach to the ADE

Our goal is to describe the motion of an ensemble of particles as measured by its concentration (or mass) in space and time $C(x,t)$. The distribution of particles within the plume will be represented by the pdf governing the location of a single particle in space. The discrete stochastic model we choose to develop this distribution is the random walk. We write the random particle location $X(t)$ as sum of stationary random jump lengths Y_i separated by time steps of length Δt:

$$X(t) = \sum_{n=1}^{t/\Delta t} Y_n.$$

Then we can use the Central Limit Theorem (CLT) to estimate the distribution of $X(t)$ at long time. The law of large numbers (LLN) says that the average of independent and identically distributed (iid) random variables with mean μ converges to the theoretical average:

$$\frac{Y_1 + \cdots + Y_n}{n} - \mu \to 0 \tag{8}$$

as n becomes large. The CLT refines this statement by approximating the deviation between these two terms. It says that when appropriately rescaled, the sum of iid random variables with mean μ and standard deviation σ converges to a standard normal

distribution:

$$\frac{Y_1 + \cdots + Y_n - n\mu}{\sigma n^{1/2}} \rightarrow N(0,1)$$

We use this theorem to approximate the sum of iid particle jumps

$$Y_1 + \cdots + Y_n \approx n\mu + \sigma n^{1/2} N(0,1). \tag{9}$$

Let the number of jumps be the total time divided by time per jump $n = \frac{t}{\Delta t}$ and recast the CLT to suit the random sum

$$Y_1 + \cdots + Y_n \approx t\frac{\mu}{\Delta t} + \frac{\sigma}{\sqrt{\Delta t}}t^{1/2} N(0,1). \tag{10}$$

Let $v = \frac{\mu}{\Delta t}$ and $D = \frac{\sigma^2}{2\Delta t}$ and we find a Gaussian density

$$C(x,t) = \frac{1}{\sqrt{2\pi \cdot 2Dt}}e^{-\frac{(x-vt)^2}{2 \cdot 2Dt}} \tag{11}$$

that governs particle location after a number of jumps. The right-hand side of (10) is called Brownian motion with drift. It combines a deterministic advective drift with a stochastic diffusion term $Z(t) = (2Dt)^{1/2}N(0,1)$ that spreads like $t^{1/2}$ (Fickian scaling). The ADE solution $C(x,t)$ given by (11) is the probability density of the process $vt + Z(t)$ that represents the location of a randomly selected particle at time $t > 0$, assuming that particle location $x = 0$ at the initial time $t = 0$. Figure 2(a) illustrates the simple random walk. Figure 2(b) shows the same random walk at a longer time scale. Figure 2(c) shows the Brownian motion that emerges as the long-time scaling limit of this simple random walk. The continuous graph of the particle path is a random (not generated from a deterministic pattern) fractal with Hausdorff dimension 3/2 [*Mandelbrot*, 1982]. The abrupt jumps seen in the random walk disappear in the scaling limit. Figure 3 illustrates the probability density $C(x,t)$ of particle locations for the Brownian motion limit process.

3.2. Eulerian approach to the ADE

Using an Eulerian approach, we choose a specific location in space and describe particle motion through that location with time. From this perspective, we begin with a conservation of particle mass equation that equates the rate of mass change at a location with the difference between the mass of particles entering and leaving:

$$n\frac{\partial C(x,t)}{\partial t} = -\frac{\partial F(x,t)}{\partial x}, \tag{12}$$

where effective porosity $n = 1$ for transport at the surface, particle concentration C is in mass per volume, and flux F in mass per area per time. If particle flux is assumed to be by advection and Fickian dispersion:

$$F = vnC - n\mathcal{D}\frac{\partial C}{\partial x}, \tag{13}$$

where v is average particle velocity and $\mathcal{D}$ is a dispersion coefficient, then we obtain the advection-dispersion equation (also known as the Fokker Planck equation [*Feller*, 1971]):

$$\frac{\partial C(x,t)}{\partial t} = -v\frac{\partial C}{\partial x} + \mathcal{D}\frac{\partial^2 C(x,t)}{\partial x}. \tag{14}$$

The solution to the ADE can be obtained using Fourier transforms:

$$\frac{\partial \hat{C}(k,t)}{\partial t} = -v(ik)\hat{C}(k,t) + \mathcal{D}(ik)^2\hat{C}(k,t). \tag{15}$$

Solve for C and use initial conditions $C(x,0) = \delta(0)$ to find

$$\hat{C}(k,t) = e^{\left(-vt(ik) + Dt(ik)^2\right)} \tag{16}$$

the Fourier transform of a Gaussian density (11) with mean $\mu = vt$ and variance $\sigma^2 = 2\mathcal{D}t$.

It is interesting that the deterministic solution to a partial differential equation should be a pdf. This important observation means that random behavior exhibits a deterministic regularity at long time. The location $vt + Z(t)$ of a randomly selected particle is unknown,

but even so, the relative concentration of particles is predictable. We obtain the same

solution using the Eulerian perspective of Fickian particle jumps and conservation of

solute mass, or the Lagrangian random walk formulation taken to its scaling limit by

applying the CLT. The assumptions underlying the CLT must hold for a Gaussian limit

to emerge. This means that there are assumptions buried in Fickian/mass conserving

ADE that must hold for it to be applicable. Our focus here is that, in order for the ADE

to emerge, there can be only moderate deviation from the average jump size. Since it

is well known that the ADE and its solutions do not always reproduce essential features

of sediment or solute transport (e.g. *Benson* [1998]; *Ganti et al.* [2009]; *Gooseff et al.*

[2003]; *Worman et al.* [2007]; *Neuman and Tartakovsky* [2009], as well as references in

this special issue on stochastic transport and emergent scaling at the Earth surface), we

will now consider models that allow less stringent assumptions. We will also explore the

long memory case where a time component is introduced by allowing random waiting times

between jumps. In the following sections, we introduce fractional-in-space ADEs caused

by heavy-tailed velocity distributions, and fractional-in-time ADEs caused by heavy-tailed

residence times.

3.3. Characteristics of ADE solutions

If tracer transport is well modeled by a classical ADE, then long-term transport charac-

teristics (of a pulse of tracer) will resemble the Green's function (point source) solutions

[*Arfken and Weber*, 1995] to traditional integer-order advection-dispersion equations (Fig.

4). These characteristics include:

1. Spatial snapshots

(i) Gaussian (symmetric bell-shaped) concentration profiles

274   (ii) Plume edges decay rapidly

275   (iii) Snapshot width spreads like $t^{1/2}$

276   (iv) Total mass remains constant over time

277   2. Flux at position x

278   (i) Bell-shaped breakthrough curves

279   (ii) Leading and trailing edges decay rapidly

280   (iii) Breakthrough curve width grows like $x^{1/2}$

281   (iv) Area under breakthrough curve remains constant

4. Fractional-in-space ADEs

282   A variety of field and theoretical studies suggest super-diffusive non-locality in transport

283   of tracers at the Earth surface. For example, re-analysis of tracer studies in sand and

284   gravel-bed streams revealed hop length distributions with heavy tails [*Bradley et al.*,

285   2009]. *Foufoula-Georgiou et al.* [2009] argue that transport on hillslopes is non-local

286   and that sediment flux must be calculated using not just the local gradient, but also

287   that of upslope topography. *Stark et al.* [2009] use a fractionally integrated flux term to

288   capture non-local effects in a model of bedrock channel evolution that includes the effect

289   of hillslope sediment production on channel bed sediment buffering and bedrock erosion.

290   These studies all suggest fractional flux terms to represent non-locality in transport.

291   *Hill et al.* [2009] suggest the source of power law hop lengths in bedload transport. They

292   present and review experimental data to justify an exponential step length for particles

293   of a given size. Since mean step length varies with particle size, the overall step length

294   for all particles is a mixture of those exponential distributions. Using standard models

[e.g., Gamma distribution] for the pdf of grain size leads to an overall step length pdf with a heavy power law tail. Even though neither the grain size distribution nor the step length pdf for grains of a given diameter are heavy tailed, the mixture distribution turns out to have a heavy tail. This result shows that a fractional Exner equation [*Ganti et al.*, 2009] is applicable to bedload transport. It also highlights the need for additional experiments and analysis to improve estimates of grain size distribution, step length for moving particles, as well as entrainment rate for particles of varying size. Here we develop space-fractional ADEs used to model super-diffusive transport using both the Lagrangian (particle tracking) and Eulerian (conservation of mass) approach.

4.1. Lagrangian approach to fractional-in-space ADEs

If the density of jump lengths Y_i is the heavy-tailed density $p(x)$, then the integral $\sigma^2 = \int_{-\infty}^{\infty} (Y_i - \mu)^2 p(x) dx$ diverges. Then the classical CLT we use to determine the density of the sum of jump lengths $X(t)$ does not hold. A more general form of the CLT exists, however, and says that the properly rescaled sum of stationary iid random variables with even infinite variance converges in distribution to an α-stable density, denoted S_α, with mean μ, spread σ, tail parameter α, and skewness β [*Feller*, 1971; *Gnedenko and Kolmogorov*, 1968]. Here we assume that particle jumps are heavy-tailed in the direction of flow only.

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma n^{1/\alpha}} \to S_\alpha(\mu = 0, \sigma = 1, \alpha, \beta = 1). \tag{17}$$

Note that the Gaussian distribution is α-stable with $\alpha = 2$ (in which case the skewness is irrelevant), making this CLT a true generalization of the classical version.

As before, recast the generalized CLT (17) to approximate the sum of random jumps:

$$Y_1 + \cdots + Y_n \approx n\mu + \sigma n^{1/\alpha} S_\alpha \left(0, 1, \alpha, \beta\right)$$

$$Y_1 + \cdots + Y_n \approx t\frac{\mu}{\Delta t} + \frac{\sigma}{\Delta t^{1/\alpha}} t^{1/\alpha} S_\alpha \left(0, 1, \alpha, \beta\right) \tag{18}$$

and let $v = \frac{\mu}{\Delta t}$ and $\mathcal{D} = \frac{\sigma^\alpha}{\Delta t}$ to find an α-stable density

$$\hat{C}(k, t) = e^{(-vt(ik) + \mathcal{D}t(ik)^\alpha)}$$

in Fourier space. The α-stable density cannot be written in closed form in real space. A variety of methods exist to estimate the inverse transform numerically [*Nolan*, 1998].

Figure 5(a) illustrates a simple random walk with heavy tailed particle jumps. Figure 5(b) shows the same random walk at a longer time scale. Figure 5(c) shows the stable Lévy motion that emerges as the long-time scaling limit of this heavy tailed random walk. The graph of the particle path is a random fractal with dimension $2 - 1/\alpha$. The large jumps seen in the random walk persist in the scaling limit. Figure 6 illustrates the probability density $C(x, t)$ of particle locations for the stable Lévy motion limit process. This particle motion process exhibits a super-diffusive spreading rate, skewness, and power law right tail (early arrivals downstream).

The right-hand side of (18) is called α-stable Lévy motion with drift. This limit process $vt + Z(t)$ represents the location of a randomly selected particle at time $t > 0$, assuming that particle location $x = 0$ at the initial time $t = 0$. The diffusion term $Z(t) = t^{1/\alpha} S_\alpha \left(0, 1, \alpha, \beta\right)$ exhibits non-Fickian scaling, a characteristic often seen in laboratory and field studies [*Benson et al.*, 2000, 2001]. Since $\alpha < 2$ the scaling factor $t^{1/\alpha}$ grows faster (the plume spreads faster) than the classical ADE case $\alpha = 2$, so this model is often called super-diffusive. Aside from its super-Fickian scaling, Lévy motion has other features that distinguish it from its close cousin, Brownian motion. Lévy motion proba-

bility densities are positively skewed, with a long leading tail [*Samorodnitsky and Taqqu*, 1994]. In fact, the probability mass a distance r units or more from the plume center of mass falls off like $x^{-\alpha}$ so that particles are much more likely to race ahead of the mean.

4.2. Eulerian approach to fractional-in-space ADEs

The Eulerian approach accounts for bulk mass movements into and out of a finite size control volume of edge size $\Delta x = h$ during a short time step Δt, as in Figure 7. In a fractal porous medium (and perhaps, in a fractal river network), the velocity of different particles in the control volume can vary widely, so that a particle can enter the control volume from long distances upstream via high velocity paths. The fractal medium imposes a power-law velocity distribution, leading to a fractional Fick's Law $F = -\mathcal{D}\partial^{\alpha-1}C/\partial x^{\alpha-1}$ for dispersive flux [*Schumer et al.*, 2001], since dispersion comes from variations in velocity. The Grunwald weights in Figure 1 code the velocity distribution. A fractional dispersive flux in the solute mass conservation equation (12) leads to a fractional-in-space ADE

$$\frac{\partial C(x,t)}{\partial t} = -v\frac{\partial C(x,t)}{\partial x} + \mathcal{D}\frac{\partial^\alpha C(x,t)}{\partial x^\alpha}. \tag{19}$$

Use Fourier transforms to solve the space-fADE (19):

$$\frac{\partial \hat{C}(k,t)}{\partial t} = -v(ik)\hat{C}(k,t) + \mathcal{D}(ik)^\alpha \hat{C}(k,t)$$

$$\hat{C}(k,t) = \exp\left[-vt(ik) + \mathcal{D}t(ik)^\alpha\right]$$

which is the Fourier transform of an α-stable density with shift $\mu = vt$, spread $\sigma^\alpha = \mathcal{D}t$, skewness $\beta = 1$, and tail parameter $1 < \alpha \le 2$. More extreme deviations from the mean velocity are represented by heavier tailed jump distributions and, in turn, are governed by fractional ADEs with smaller fractional derivatives on the flux term [*Clarke et al.*, 2000].

An alternative Eulerian derivation of the space-fADE uses a fractional conservation of mass with a traditional Fickian flux [*Meerschaert et al.*, 2006]. *Ganti et al.* [2009] develop a

356 probabilistic Exner equation for sediment transport, which is applicable when particle step

357 lengths follow a heavy tailed (power law) pdf. This asymptotic governing equation for the

358 movement of entrained particles is valid at scales long enough to encompass many particle

359 jumps. Using an active layer approach, neglecting porosity, and assuming equilibrium

360 (steady, uniform) bedload transport of grains of uniform size over a bed, the probabilistic

361 Exner equation in the presence of a heavy tailed step length is $\frac{L_a}{E}\frac{\partial f_a}{\partial t} = -v\frac{\partial f_a}{\partial x} + D\frac{\partial^\alpha f_a}{\partial x^\alpha}$,

362 where f_a is the fraction of tracer particles in the active layer, v, D are as in the space-

363 FADE (19), L_a is the thickness of the active layer, and E is the entrainment rate. Dividing

364 both sides by L_a/E yields an equation mathematically identical to the space-FADE (19),

365 so that the same solution methods apply.

4.3. Characteristics of space fADE solutions

366 Characteristics of the Green's function solutions to the advection-dispersion equation

367 with a fractional-order term in space (Fig. 4) are:

368   1. Spatial snapshots

369   (i) α-stable concentration profiles are skewed with long right tail

370   (ii) Plume leading edges decays like a power law $x^{-\alpha-1}$

371   (iii) Snapshot width spreads like $t^{1/\alpha}$

372   (iv) Total mass remains constant over time

373   2. Flux at position x

374   (i) Asymmetric breakthrough curves

375   (ii) Long leading edge

376   (iii) Breakthrough curve width grows like $x^{1/\alpha}$.

₃₇₇   (iv) Area under breakthrough curve remains constant.

5. Fractional-in-time ADEs

₃₇₈   Tracer particles at the Earth surface spend more time at rest than in motion [*Sadler*,

₃₇₉ 1981; *Leopold et al.*, 1964; *Tipper*, 1983]. For example, frequency distributions of bedload

₃₈₀ path length are positively skewed, indicating that many particles do not move from the

₃₈₁ point of tracer input over the measurement period [*Hassan and Church*, 1992; *Pryce and*

₃₈₂ *Ashmore*, 2003]. Small deviations in waiting times will not affect long term dispersion

₃₈₃ rates, but heavy-tailed waiting times will. This affects overall bedload rates [*Singh et al.*,

₃₈₄ 2009] and the resulting transport characteristics can be accommodated by time-fractional

₃₈₅ ADE, which will be developed from both the Lagrangian and Eulerian view. Evidence for

₃₈₆ heavy-tailed waiting times in transport is captured in the depositional record [*Schumer*

₃₈₇ *and Jerolmack*, 2009]. A promising avenue for future research is to examine waiting times

₃₈₈ between entrainment events, to determine whether a time-fractional Exner equation may

₃₈₉ be useful to incorporate particles outside the active layer, including deeply buried particles,

₃₉₀ those incorporated into the hyporheic zone, and particles outside the normal flowpath of

₃₉₁ a river (sandbars, floodplain - e.g. see discussion in *Ganti et al.* [2009]).

5.1. Lagrangian approach to fractional-in-time ADEs

₃₉₂   A continuous time random walk (CTRW) imposes a random waiting time J_n before the

₃₉₃ particle jump Y_n occurs. The particle location at time t is

₃₉₄
$$X(t) = Y_1 + \cdots + Y_n$$

where $n = N(t)$ is the number of jumps by time t. Suppose the particle jumps have mean zero so that the centered (generalized) CLT applies:

$$Y_1 + \cdots + Y_n \approx n^{1/\alpha} Z$$

where Z is stable (normal if $\alpha = 2$). The sum $J_1 + \cdots + J_n$ gives the time of the nth particle jump. If the waiting times have a finite mean ν then the law of large numbers (8) shows that the nth jumps happens at time $t = T_n \approx n\nu$ and then $X(t) \approx Y_1 + \cdots + Y_{t/\nu}$ leads to the classical ADE. On the other hand, if the waiting times J_n are heavy-tailed with power-law index $0 < \gamma < 1$ and scale $\sigma = 1$ then the generalized CLT yields

$$\frac{J_1 + \cdots + J_n - n\nu}{n^{1/\gamma}} \to W, \tag{20}$$

where W is distributed like $S_\gamma(\mu = 0, \sigma = 1, \gamma, \beta = 1)$. Since $\gamma < 1$, we find $n\nu/n^{1/\gamma} \to 0$, and so the centering term in (20) can be neglected. Then the nth jump occurs at time

$$t = J_1 + \cdots + J_n \approx n^{1/\gamma} W \tag{21}$$

for large n. Solving for n shows that the particle location $X(t) = Y_1 + \cdots + Y_n \approx n^{1/\alpha} Z$ where $n \approx (t/W)^\gamma$ for large n. Putting this all together, we see that

$$X(t) \approx (t/W)^{\gamma/\alpha} Z$$

and this approximation is exact in the scaling limit, where the number of jumps tends to infinity [*Meerschaert and Scheffler*, 2004]. If $g(t)$ is the pdf of W, a change of variables [*Meerschaert and Scheffler*, 2004] shows that $u = (t/W)^\gamma$ has pdf

$$q(u, t) = \frac{t}{\gamma} u^{-1-1/\gamma} g(tu^{-1/\gamma})$$

for $u > 0$. If $f(x, u)$ is the pdf of $u^{1/\alpha} Z$ then taking a weighted average with respect to u [*Meerschaert and Scheffler*, 2004] shows that the limit particle location x at time t has

pdf

$$C(x,t) = \int_0^\infty f(x,u)q(u,t)\, du \tag{22}$$

a scale mixture of stable densities. Since

$$\hat{f}(k,u) = e^{\mathcal{D}u(ik)^\alpha}$$

$$\tilde{q}(u,t) = s^{\gamma-1}e^{-us^\gamma} \tag{23}$$

(for the Laplace transform formula, see [*Meerschaert and Scheffler*, 2002b]) a simple integration shows that the Fourier-Laplace transform of $C(x,t)$ is

$$\bar{C}(k,s) = \frac{s^{\gamma-1}}{s^\gamma - \mathcal{D}(ik)^\alpha}$$

which will lead to the space-time fractional ADE. Rearrange to get $s^\gamma \bar{C}(k,s) - s^{\gamma-1} = \mathcal{D}(ik)^\alpha \bar{C}(k,s)$ and invert the Laplace and Fourier transforms using (7) to get

$$\frac{\partial^\gamma C(x,t)}{\partial t^\gamma} = \mathcal{D}\frac{\partial^\alpha C(x,t)}{\partial x^\alpha}$$

using the Caputo derivative in time, where the point source initial condition implies $\hat{C}(k,t=0) = 1$. Adding an advective drift yields the space-time fractional ADE

$$\frac{\partial^\gamma C(x,t)}{\partial t^\gamma} = -v\frac{\partial C(x,t)}{\partial x} + \mathcal{D}\frac{\partial^\alpha C(x,t)}{\partial x^\alpha} \tag{24}$$

which governs the CTRW particle density in the long-time limit. It is important to note that the parameter α codes the large particle jumps (early arrivals) and γ controls the long waiting times (residence times). A more detailed analysis using the CLT leads to higher order temporal derivatives [*Baeumer et al.*, 2005; *Baeumer and Meerschaert*, 2007].

The limit process $X(t) = Z(T(t))$ is a subordinated Lévy (or Brownian) motion. The outer process is the random walk limit discussed previously. The inner process $T(t) = (t/W)^\gamma$ adjusts for the time a randomly selected particle spends in motion (e.g., see recent work by *Ganti et al.* [2009] who propose a subordinated Brownian motion model for

sediment transport). We call $u = T(t)$ the operational time, since it counts the number of

particle jumps by time t. Since the operational time process incorporates memory effects,

it is non-Markovian, and it is simpler to understand its inverse process $t = W(u)$ that

maps operational time back to clock time t. The process $W(u) = u^{1/\gamma}W$ comes from

(21) and it is another Lévy motion with index $\gamma < 1$. Then the operational time process

$u = T(t)$ is the inverse of this Lévy process. The scaling $W(u) = u^{1/\gamma}W$ is consistent with

the inverse scaling $T(t) = t^{\gamma}T$ already noted. The scaling $X(t) = Z(T(t)) = Z(t^{\gamma}T) =$

$t^{\gamma/\alpha}Z(T) = t^{\gamma/\alpha}X$ shows that particles following the space-time fractional ADE spread at

rate $t^{\gamma/\alpha}$ which is slower than the ADE if $\alpha = 2$. This model is often called sub-diffusive.

The random time change T has finite moments of all orders, so in the heavy tailed case

$\alpha < 2$ it does not affect the tails of $X(t)$ which still fall off like $x^{-\alpha}$ [*Meerschaert and*

Scheffler, 2004]. The mixture integral (22) is a weighted average of the PDF $f(x, u)$ of

the random walk limit $Z(u)$ according to the PDF $q(u, t)$ of the inverse Lévy process $T(t)$.

Figure 8(a) illustrates a CTRW with heavy tailed waiting times. Figure 8(b) shows the

same CTRW at a longer time scale. Figure 8(c) shows the continuous particle path of the

subordinated Brownian motion that emerges as the long-time scaling limit of this CTRW.

The long resting times (particle retention) seen in the CTRW persist in the scaling limit.

5.2. Eulerian approach to fractional-in-time ADEs

The Eulerian approach accounts for bulk mass movements into and out of a control

volume of edge size $\Delta x = h$ during a short time step Δt, as in Figure 7. The time-

fADE includes memory effects, allowing particle residence times for long periods. This

means that particles can enter the control volume at the current time step from locations

upstream in time as depicted in Figure 10. A power-law memory function [*Haggerty and*

Gorelick, 1995; *Haggerty et al.*, 2000, 2002] implies that the time-flux can be coded using a discrete convolution in time $\sum C(x, t - jh)w_j$ using the power-law Grunwald weights from Figure 1. This leads to a time-fractional analogue $n\partial^\gamma C(x,t)/\partial t^\gamma = -\partial F(x,t)/\partial x$ to the conservation of solute mass equation (12) that combines with the usual advective and dispersive flux equation (13) to produce the time-fADE

$$\frac{\partial^\gamma C(x,t)}{\partial t^\gamma} = -v\frac{\partial C(x,t)}{\partial x} + \mathcal{D}\frac{\partial^2 C(x,t)}{\partial x^2}. \tag{25}$$

If power-law residence time (memory function) is combined with power-law velocity distribution, we recover the same space-time fADE (24) derived from the Lagrangian model. The time-fractional parameter γ codes retention, since it implies that the memory function $t^{-\gamma}$ is a power law, and so it can be estimated from the late-time tail of the breakthrough curve [*Schumer et al.*, 2003].

The solution $C(x,t)$ to the time-fADE (25) is the probability density of the subordinated process $X(t) = Z(T(t))$ that models the location of a randomly selected particle. The solution formula (22) mixes the density of $Z(u)$ according to the operational time process $u = T(t)$. The outer process $Z(u)$ is a Brownian motion connected with the right-hand side of (25), and the inner process compensates for the memory effects of the fractional time derivative on the left-hand side of (25).

5.3. Characteristics of time (and space) fADE solutions

Characteristics of the Green's function solutions to the advection-dispersion equation with a fractional-order term in time and possibly space are:

1. Spatial snapshots

(i) Subordinated α-stable concentration profiles are skewed with long right tail if $\alpha < 2$

(ii) Plume leading edges decays like a power law $x^{-\alpha-1}$ if $\alpha < 2$

(iii) Snapshot width spreads like $t^{\gamma/\alpha}$

(iv) Total mass remains constant over time

2. Flux at position x

(i) Breakthrough curve tail decays as $t^{-\gamma-1}$

(ii) Breakthrough curve width grows like $x^{\gamma/\alpha}$

(iii) If $\alpha < 2$, immediate spike in flux then strong leading edge.

(iv) Area under breakthrough curve is constant with time

5.4. Fractional ADEs for mobile and immobile zones

The fractional-in-time ADE describes the evolution of a tracer plume, but does not distinguish between particles moving at the Earth surface (with mobile concentration denoted C_m) and particles that have been immobilized (C_{im}), where $C = n_m C_m + n_{im} C_{im}$ (e.g. *Coats and Smith* [1964]), mobile porosity $n_m = 1$ at the surface and immobile porosity $n_i m$ may be less than 1 if availability of pore space is a control on immobilization of sediment or solute. This is significant because typical sampling methods for surface water and groundwater solute concentration permit only observation of mobile solute concentration. Mobile solute equations can predict absolute solute concentration rather than relative or normalized concentration C/C_0 because they account for mass loss to immobile zones with time [*Schumer et al.*, 2003].

Traditional mobile-immobile equations assume that particles move between the mobile and immobile phases at an instantaneous rate proportional to the difference in concentration. If particles begin in an immobile phase, this implies that immobile concentration decays exponentially in time, as reflected in the late-time breakthrough curve. A multiple-rate mass transfer model (MRMT) uses a memory function to govern release from the immobile phase [*Haggerty and Gorelick*, 1995], and has been used to fit late time solute breakthrough curves in groundwater aquifers [*Haggerty et al.*, 2000]. Following [*Schumer et al.*, 2003] we write the MRMT equations for transport:

$$\frac{\partial C_m}{\partial t} + \beta \frac{\partial C_m}{\partial t} \star g(t) = LC_m - C_m(x, t = 0)\beta g(t)$$
$$\frac{\partial C_{im}}{\partial t} + \beta \frac{\partial C_{im}}{\partial t} \star g(t) = LC_{im} + C_m(x, t = 0)g(t)$$

(26)

where the capacity coefficient $\beta = n_{im}/n_m$, "$\star$" denotes convolution, $g(t)$ is the memory function, and we write $L = -v\partial/\partial x + D\,\partial^2/\partial x^2$ for brevity. Equations (26) assume that all solute begins in the mobile phase. If power law memory is of form $g(t) = t^{-\gamma}/\Gamma(1-\gamma)$, the memory function convolution becomes the Caputo fractional derivative in time (6), and (26) becomes the fractal mobile-immobile equations

$$\frac{\partial C_m}{\partial t} + \beta \frac{\partial^\gamma C_m}{\partial t^\gamma} = LC_m - C_{m,0}(x)\frac{\beta t^{-\gamma}}{\Gamma(1-\gamma)}$$
$$\frac{\partial C_{im}}{\partial t} + \beta \frac{\partial^\gamma C_{im}}{\partial t^\gamma} = LC_{im} + C_{m,0}(x)\frac{t^{-\gamma}}{\Gamma(1-\gamma)}$$

(27)

with $C_{m,0}(x) = C_m(x, t = 0)$ as in [*Schumer et al.*, 2003]. Figure 9 shows the characteristic behavior of fractal mobile-immobile transport. The late-time behavior of the breakthrough curve is governed by the order of the time-fractional derivative. The key characteristics of the mobile zone FADE that differ from those of the fractional-in-time ADE are:

1. Total mass in the mobile zone decays with time as $\frac{t^{\gamma-1}}{\Gamma(\gamma)}$

519   2. If $\alpha < 2$, the breakthrough curve has a long leading edge but no immediate spike in

520 flux.

6. Scaling properties of fractional ADEs

521   The limiting stochastic process governed by any (traditional or fractional) ADE is self-

522 similar and has useful scaling properties. For example, the traditional ADE governs

523 Brownian motion with drift, so that relative concentration $C(x,t)$ in (11) gives the prob-

524 ability density for this stochastic process. There are actually two scales in this equation.

525 The plume center of mass moves linearly, proportional to t, and the plume spreads more

526 slowly, proportional to $t^{1/2}$ as t increases. To focus on the spread, adopt a moving coordi-

527 nate system with origin at the plume center of mass. This converts to $v = 0$ in the ADE,

528 so that it reduces to the diffusion/dispersion equation

$$529 \qquad \frac{\partial C(x,t)}{\partial t} = \mathcal{D}\frac{\partial^2 C(x,t)}{\partial x^2} \tag{28}$$

530 whose point source solution is

$$531 \qquad C(x,t) = \frac{1}{\sqrt{2\pi \cdot 2\mathcal{D}t}}e^{-\frac{x^2}{2\cdot 2\mathcal{D}t}} \tag{29}$$

532 a mean zero normal density with variance $2\mathcal{D}t$. A snapshot graph of this density curve

533 for any fixed t is bell-shaped, centered at the origin, and as t increases, the snapshots

534 are all the same shape. This scaling property is expressed mathematically as $C(x,t) =$

535 $t^{-1/2}C(xt^{-1/2}, 1)$ which clearly shows that the concentration snapshot spreads away from

536 its center of mass like $t^{1/2}$ and the plume peak concentration decreases at the same rate.

537 This scaling is also evident in the Fourier transform $\hat{C}(k,t) = \exp[-\mathcal{D}tk^2]$ since evidently

538 $\hat{C}(k,t) = \hat{C}(t^{1/2}k, 1)$. For the space-fADE, the same moving coordinate system leads

539 to $C(x,t) = t^{-1/\alpha}C(xt^{-1/\alpha}, 1)$ and again with $\hat{C}(k,t) = \exp[\mathcal{D}t(ik)^\alpha]$ we have $\hat{C}(k,t) =$

$\hat{C}(t^{1/\alpha}k, 1)$. Thus the space-fADE snapshot, a stable density with index α, spreads like $t^{1/\alpha}$ while its peak falls at the same rate. Solutions to the space-time fADE (24) with $v = 0$ can be expressed as $\hat{C}(k,t) = E_\gamma(t^\gamma \mathcal{D}(ik)^\alpha)$ [*Mainardi and Gorenflo*, 2000] where the Mittag-Leffler function

$$E_\gamma(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(1 + \gamma k)}$$

is an extension of the exponential (the special case $\gamma = 1$) used in the theory of time-fractional equations. It is easy to see that $\hat{C}(k,t) = t^{\gamma/\alpha}\hat{C}(k,1)$ and so the space-time fADE is also scale-invariant after mean centering: The plume snapshot spreads like $t^{\gamma/\alpha}$ and the plume peak falls at the same rate [*Meerschaert et al.*, 2002].

If a transport model reveals scale dependence of parameters, it is likely that the model is scale dependent, not the transport process itself. One option is to use a model with time-dependent parameters, but this is difficult to apply in practice. Using a model that includes the correct scaling is useful for prediction because its parameters will be constant in time. In the fractional-in-space and time case, the scaling rate $t^{\gamma/\alpha}$ can be obtained with non-unique $0 < \gamma \leq 1$ and $1 < \alpha \leq 2$ (for example, a scaling rate of $t^{0.56}$ can be obtained with $\alpha = 1.8$ and $\gamma = 1$ or $\alpha = 1.5$ and $\gamma = 0.83$). The tail properties of plume snapshots or breakthrough curves will be necessary to obtain the order of the space and time fractional derivatives. An analysis of plume moments may also be necessary for choosing the appropriate transport model [*Zhang et al.*, 2008]. The scaling rate for plumes following a mobile fractional ADE is [*Schumer et al.*, 2003] $C_m(x,t) = t^{\gamma-1}t^{-\frac{\gamma}{\alpha}}C_m(t^{-\frac{\gamma}{\alpha}}x, 1)$. In summary, fADE plumes spread out from the center of mass according to precise scaling rates, that are determined by the order of the space or time derivative. A smaller index α on the space derivative implies faster (super-dispersive)

⁵⁶³ spreading, while a smaller γ index on the time derivative yields a slower (sub-dispersive)

⁵⁶⁴ spread. In practice, the super-dispersive effect comes from high velocity contrasts, while

⁵⁶⁵ the sub-diffusive behavior reflects long residence times.

7. CTRW derivation of the space-time fADE

⁵⁶⁶ The Lagrangian derivation of the space-time fADE is closely connected to the underlying

⁵⁶⁷ CTRW model for particle transport [*Berkowitz et al.*, 2002, 2006]. Each random particle

⁵⁶⁸ jump Y_n follows a random waiting time J_n. Suppose that Y_n has mean v and $P(Y_n - v >$

⁵⁶⁹ $x) \approx x^{-\alpha}$ and $P(J_n > t) \approx t^{-\gamma}$ for large x and t. Then the jump pdf $p(x)$ of Y_n has

⁵⁷⁰ Fourier transform $\hat{p}(k) = 1 - ikv + \mathcal{D}(ik)^\alpha + \cdots$ and the waiting time pdf $\psi(t)$ of J_n

⁵⁷¹ has Laplace transform $\tilde{\psi}(s) = 1 - s^\gamma + \cdots$ where we emphasize the dominant terms for

⁵⁷² small k and s. The master equation (*Montroll and Weiss* [1965]; *Scher and Lax* [1973],

⁵⁷³ also used in *Furbish et al.* [2009a, b] to describe probability of soil particle flux) gives the

⁵⁷⁴ Fourier-Laplace transform of the CTRW particle density :

⁵⁷⁵
$$\bar{C}(k, s) = \frac{1 - \tilde{\psi}(s)}{s} \frac{1}{1 - \hat{p}(k)\tilde{\psi}(s)} \tag{30}$$

⁵⁷⁶ After substituting power-law space and time jump pdfs into the master equation, we find

⁵⁷⁷
$$\bar{C}(k, s) \approx \frac{s^{\gamma-1}}{s^\gamma + ikv - \mathcal{D}(ik)^\alpha} \tag{31}$$

⁵⁷⁸ which rearranges to

⁵⁷⁹
$$s^\gamma \bar{C}(k, s) - s^{\gamma-1} = -ikv\bar{C}(k, s) + \mathcal{D}(ik)^\alpha \bar{C}(k, s).$$

⁵⁸⁰ Invert the Fourier and Laplace transforms to recover the space-time fADE (24). For

⁵⁸¹ light-tailed particle jumps, the Fourier transform $\hat{p}(k) = 1 - ikv + \mathcal{D}(ik)^2 + \cdots$ and the

⁵⁸² second derivative in space reappears. For light-tailed waiting times, the Laplace transform

583 $\tilde{\psi}(s) = 1 - s + \cdots$ leading to the first derivative in time. The asymptotic approximation

584 is identical to our earlier discussion, and it is only in the limit that the neglected terms

585 vanish to produce the fADE (or the ADE). If the jump length and residence time densities

586 are known exactly, it is possible to model transport with a discrete CTRW without taking

587 the long time or scaling limit simply by using those densities in Equation (30). In this

588 case, a pre-ergodic solution is reached that does not necessarily have scale invariance.

8. Relationship with other stochastic models

589 Fractional ADEs are related to other general stochastic process models of random par-

590 ticle motion. Here we give a brief discussion to provide context. The CTRW master

591 equation (30) inverts to an integral equation for particle density $C(x,t)$ which is also

592 known as the (generalized) master equation [*Klafter and Sibley*, 1980]. The convolution

593 integral in this master equation can approximate fractional derivatives in space and/or

594 time. The solution to this master equation gives the pdf of the CTRW particle den-

595 sity. The underlying particle motion process consists of space-time jumps as shown in

596 Figures 2(a), 5(a), or 8(a). Fractional advection-dispersion equations arise in the scaling

597 limit. A different approach based on statistical mechanics leads to a continuously evolving

598 stochastic model for particle motion, where particle flux is represented as a convolution

599 integral [*Cushman*, 1997]. The space-time convolution can capture memory effects (time

600 convolution) and non-local super-diffusive particle jumps (space convolution). The space-

601 time fractional diffusion equation is a special case of this stochastic model, in which the

602 convolution kernel is a power law [*Cushman and Ginn*, 2000]. A more complex but corre-

603 spondingly general theory is reviewed in *Neuman and Tartakovsky* [2009]. That paper also

604 compares and contrasts the various theories used in stochastic hydrology in more detail.

The added generality allows a velocity field that varies in space and time [*Neuman*, 1990]. The variable coefficient fractional ADE represents the special case of a power-law memory kernel in space and time. In any given situation, the practitioner will make an intelligent model selection that takes into account a trade-off between simplicity and generality. For example, it is possible to generate stochastic processes with non-stationary increments that have heavy tails and satisfy classical ADEs with time dependent dispersion tensors [*Berkowitz et al.*, 2002]. In our view, fractional advection dispersion equations represent the simplest model that incorporates the anomalous spreading and skewness often seen in transport at the Earth surface. They complement CTRW and other more detailed stochastic models with a simplified (limit) case appropriate for plume modeling at the large scale.

9. Summary- Use of fractional ADEs in modeling Earth surface transport

Characterizing the collective behavior of particle transport on the Earth surface is a key ingredient for describing landscape evolution. Alternatives to local, diffusive transport laws are sought because these classical models do not always capture essential features of transport on hillslopes, valleys, river channels, or river networks. Fractional ADEs that can incorporate non-local effects in space are now used to describe anomalous transport of sediment which can arise from particle size [*Ganti et al.*, 2009; *Hill et al.*, 2009] and flow field variation [*Stark et al.*, 2009; *Bradley et al.*, 2009] in streams or disturbance events on hillslopes that mobilize colluvium over a wide range of scales [*Foufoula-Georgiou et al.*, 2009]. Fractional ADEs that incorporate non-local effects in time can reproduce power-law residence times and loss to immobile zones that arise because tracer particles spend more time at rest than in motion at the Earth surface. These macroscopic equations arise

from characteristic particle 'hop' behaviors (Figure 11). In the Lagrangian picture, when the probability of a long hop falls off like $x^{-\alpha}$ then the α-order fractional derivative in space emerges. If the likelihood of a long residence time falls off like $t^{-\gamma}$ then a γ-fractional derivative in time appears. From an Eulerian viewpoint, fractional derivatives in space allow the possibility of high velocity contrast, and fractional time derivatives model long power-law residence times. Space or time terms in ADEs with fractional derivatives result in solutions with flexible scaling rates. In real applications, the fractional paradigm often allows the practitioner to replace scale-dependent ADE parameters by scale-independent their fADE analogues.

References

Arfken, G., and H. Weber (1995), *Mathematical Methods for Physicists*, Academic Press, San Diego.

Baeumer, B., and M. M. Meerschaert (2007), Fractional diffusion with two time scales, *Physica A-Statistical Mechanics and Its Applications, 373*, 237–251.

Baeumer, B., D. Benson, and M. Meerschaert (2005), Advection and dispersion in time and space, *Physica A, 350* (2-4), 245–262.

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

Benson, D., S. Wheatcraft, and M. Meerschaert (2000), The fractional-order governing equation of Lévy motion, *Water Resour. Res., 36* (6), 1413–1423.

Benson, D. A. (1998), The fractional advection-dispersion equation: development and application, Ph.d., University of Nevada.

Benson, D. A., R. Schumer, M. M. Meerschaert, and S. W. Wheatcraft (2001), Fractional dispersion, Lévy motion, and the made tracer tests, *Trans. Por. Media, 42*, 211–240.

Berkowitz, B., J. Klafter, R. Metzler, and H. Scher (2002), Physical pictures of transport in heterogeneous media: Advection-dispersion, random-walk, and fractional derivative formulations, *Water Resour. Res.*, 38(10), 1191, doi:10.1029/2001WR001030.

Berkowitz, B., A. Cortis, M. Dentz, and H. Scher (2006), Modeling non-Fickian transport in geological formations as a continuous time random walk, *Rev. Geophys.*, 44, RG2003, doi:10.1029/2005RG000178.

Bradley, D., G. Tucker, and D. Benson (2009), Anomalous dispersion in a sand-bed river, *Journal of Geophys. Res. - Earth Surface, this issue.*

Clarke, D.D., M.M. Meerschaert, and S.W. Wheatcraft (2005), Fractal Travel Time Estimates for Dispersive Contaminants, *Ground Water 43*(3), 401–407.

Coats, K. H., and B. Smith (1964), Dead-end pore volume and dispersion in porous media, *J. Soc. Pet. Eng., 4*, 73–84.

Culling, W. (1960), Analytical theory of erosion, *J. Geology, 68*, 336–344.

Cushman, J. H. (1997) *The physics of fluids in hierarchical porous media: Angstroms to miles.* Kluwer Academic, New York.

Cushman, J. H., and T. R. Ginn (2000) Fractional Advection-Dispersion Equation: A Classical Mass Balance With Convolution-Fickian Flux. *Water Resour. Res.* 36(12), 3763-3766.

Cushman, J., and T. Ginn (1993), Nonlocal dispersion in media with continuously evolving scales of heterogeneity, *Trans. Por. Media, 13*, 123–138.

Dietrich, W. E., D. Bellugi, A. Heimsath, J. Roering, J. Sklar, and J. Stock (2003), Geomorphic transport laws for predicting landscape form and dynamics, in *Prediction in Geomorphology, Geophsical Monograph Series*, vol. 136, edited by P. Wilcock and R. Iverson, AGU, Washington. D.C.

Einstein H. (1937), Bed load transport as a probability problem (in German), PhD thesis, Eidgenoess. Tech. Hochsch, Zürich, Switzerland. English translation by W.W. Sayre in *Sedimentation*, edited by H.W. Shen, Appendix C, Fort Collins, Colo., 1972.

Feller, W. (1971), *An Introduction to Probability Theory and Its Applications, Wiley Series in Probability and Mathematical Statistics*, vol. Vol. II, 2nd ed., Wiley, New York.

Foufoula-Georgiou, E., V. Ganti, and W. E. Dietrich (2009), A non-local theory for sediment transport on hillslopes, *Journal of Geophys. Res. - Earth Surface, this issue.*

Furbish, D., E. Childs, P. Haff, and M. W. Schmeeckle (2009a), Rain splash of soil grains as a stochastic advection-dispersion process, with implications for desert plant-soil interactions and land-surface evolution, *JGR-ES*, *114*.

Furbish, D., P. Haff, W. Dietrich, and A. Heimsath (2009b), Statistical description of slope-dependent soil transport and the diffusion-like coefficient, *JGR-ES*, *114*.

Ganti, V., M. M. Meerschaert, E. Foufoula-Georgiou, E. Viparelli, and G. Parker, Normal and Anomalous Dispersion of Gravel Tracer Particles in Rivers, *J. Geophys. Res. Earth Surface*, in review.

Ganti, V., A. Singh, P. Passalacqua, and E. Foufoula-Georgiou (2009), Subordinated brownian motion model for sediment transport, *Phys. Rev. E*, *79*(1), in press.

Gnedenko, B., and A. Kolmogorov (1968), *Limit Distributions for Sums of Independent Random Variables*,Translated from the Russian, annotated, and revised by K.L. Chung. With appendices by J.L. Doob and P.L. Hsu. Revised edition, Addison-Wesley, Reading, Mass.

Gooseff, M. N., S. M. Wondzell, R. Haggerty, and J. Anderson (2003), Comparing transient storage modeling and residence time distribution (rtd) analysis in geomorphically varied reaches in the lookout creek basin, oregon usa, *Advances in Water Resources*, *26*, 925–937.

Haggerty, R. and S.M. Gorelick (1995), Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity, *Water Resour. Res.*, 31(10), 2383–2400.

Haggerty, R., S. A. McKenna, and L. C. Meigs (2000), On the late-time behavior of tracer test breakthrough curves, *Water Resour. Res., 36*(12), 34673479.

Haggerty, R., S. M. Wondzell, and M. A. Johnson (2002), Power-law residence time distribution in the hyporheic zone of a 2nd-order mountain stream, *Geophys. Res. Lett.*, 29(13), 1640, doi:10.1029/2002GL014743.

Haggerty, R., C. F. Harvey, C. F. von Schwerin, and L. C. Meigs (2004), What controls the apparent timescale of solute mass transfer in aquifers and soils? a comparison of experimental results, *Water Resources Research*, *40*(1).

Hassan, M.A. and M. Church (1992) The movement of individual grains on the streambed, In: Billi, P., Hey, R.D., Thorne, C.R., Tacconi, P. (Eds.), Dynamics of Gravel-Bed Rivers. Wiley, Chichester, UK, pp 159–175.

Hill,K.M., L. DellAngelo, and M.M. Meerschaert, Particle Size Dependence of the Probability Distribution Functions of Travel Distances of Gravel Particles in Bedload Transport, *J. Geophys. Res. Earth Surface*, in review.

Klafter, J. and R. Silbey (1980) Derivation of continuous-time random walk equation. *Phys. Rev. Lett.* 44(2), 55–58.

Leopold, L., M. Wolman, and J. Miller (1964), *Fluvial processes in geomorphology*, W.H. Freeman, San Francisco, CA.

Mainardi, F. (1997), Fractional calculus: some basic problems in continuum and statistical mechanics, in *Fractals and Fractional Calculus in Continuum Mechanics*, edited by A. Carpinteri and F. Mainardi, pp. 292–348, Springer, New York.

Mainardi, F. and R. Gorenflo, On Mittag-Leffer-type functions in fractional evolution processes, *J. Comput. Appl. Math.* 118 (2000), 283–299.

Malmon, D. V., T. Dunne, and S. L. Reneau (2002), Predicting the fate of sediment and pollutants in river floodplains, *Environmental Science & Technology*, *36*(9), 2026–2032.

739 Malmon, D. V., T. Dunne, and S. L. Reneau (2003), Stochastic theory of particle trajec-

740 tories through alluvial valley floors, *Journal of Geology, 111*(5), 525–542.

741 Mandelbrot, B. (1982), *The Fractal Geometry of Nature*, Freeman, San Francisco.

742 Meerschaert, M.M. and H.P. Scheffler (2001) *Limit Distributions for Sums of Independent*

743 *Random Vectors: Heavy Tails in Theory and Practice.* Wiley Interscience, New York.

744 Meerschaert, M.M. and H.P. Scheffler (2002) Semistable Lévy Motion, *Fractional Calculus*

745 *and Applied Analysis* **5**, 27–54.

746 Meerschaert, M.M., D.A. Benson, H.P. Scheffler and B. Baeumer (2002) Stochastic solu-

747 tion of space-time fractional diffusion equations. *Phys. Rev. E* **65**, 1103–1106.

748 Meerschaert, M., and H.-P. Scheffler (2004), Limit theorems for continuous time random

749 walks with infinite mean waiting times, *J. Applied Probability, 41*, 623–638.

750 Meerschaert, M., D. Benson, H. Scheffler, and P. Becker-Kern (2002), Governing equations

751 and solutions of anomalous random walk limits, *Phys. Rev. E, 66*(6), 060,102(4), 102–

752 105.

753 Meerschaert, M. and C. Tadjeran (2004) Finite difference approximations for fractional

754 advection-dispersion flow equations. *J. Comput. Appl. Math.* **172**, 65–77.

755 Meerschaert, M., H.P. Scheffler and C. Tadjeran (2006) Finite difference methods for

756 two-dimensional fractional dispersion equation. *J. Comput. Phys.* **211**, 249–261.

757 Meerschaert, M. M., J. Mortensen, and S. W. Wheatcraft (2006), Fractional vector cal-

758 culus for fractional advection-dispersion, *Physica A-Statistical Mechanics and Its Ap-*

759 *plications, 367*, 181–190.

760 Montroll, E., and G. Weiss (1965), Random walks on lattices. ii, *J. Math Phys., 6*(2),

761 167–181.

762 Neuman, S. P., and D. M. Tartakovsky (2009), Perspective on theories of non-fickian

763   transport in heterogeneous media, *Advances in Water Resources*, *32*(5), 670–680.

764 Neuman, S. P. (1990) Universal scaling of hydraulic conductivities and dispersivities in

765   geologic media. *Water Resour. Res.* 26, 1749-1758.

766 Nikora, V., J. Heald, D. Goring, and I. McEwan (2001), Diffusion of saltating particles in

767   unidirectional water flow over a rough granular bed, *Journal of Physics A-Mathematical*

768   *and General*, *34*(50), L743–L749.

769 Nikora, V., H. Habersack, T. Huber, and I. McEwan (2002), On bed particle diffusion in

770   gravel bed flows under weak bed load transport, *Water Resources Research*, *38*(6).

771 Nolan, J. P. (1998), Multivariate stable distributions: approximation, estimation, simula-

772   tion and identification, in *A Practical Guide to Heavy Tails: Statistical Techniques and*

773   *Applications*, edited by R. Adler, R. Feldman, and M. Taqqu, Birkhauser, Boston.

774 Parker, G., C. Paola, and S. Leclair (2000), Probabilistic exner sediment continuity equa-

775   tion for mixtures with no active layer, *Journal of Hydraulic Engineering-Asce*, *126*(11),

776   818–826.

777 Pryce, R.S. and P.E. Ashmore (2003) The relation between particle path length distribu-

778   tions and channel morphology in gravel-bed streams: a synthesis, *Geomorphology*, 56,

779   167–187.

780 Sadler, P. M. (1981), Sediment accumulation rates and the completeness of stratigraphic

781   sections, *Journal of Geology*, *89*(5), 569–584.

782 Samko, S. G., A. A. Kilbas, and O. I. Marichev (1993), *Fractional Integrals and Deriva-*

783   *tives: Theory and Applications*, Gordon and Breach, London.

784 Samorodnitsky, G., and M. S. Taqqu (1994), *Stable Non-Gaussian Random Processes:*

785 *Stochastic Models with Infinite Variance*, Stochastic Modeling, Chapman Hall, London.

786 Sayre, W. and D. Hubbell (1965), Transport and dispersion of labeled bed material: North

787 Loup River, Nebraska. US Geological Survey Proffesional Paper 433-C.

788 Scher, H., and M. Lax (1973), Stochastic transport in a disordered solid. i. theory, *Physical*

789 *Review B*, *7*(10), 4491–4502.

790 Schumer, R., D. A. Benson, M. M. Meerschaert, and S. W. Wheatcraft (2001), Eulerian

791 derivation for the fractional advection-dispersion equation, *Journal of Contaminant*

792 *Hydrology*, *48*, 69–88.

793 Schumer, R., D. A. Benson, M. M. Meerschaert, and B. Baeumer (2003), Fractal mo-

794 bile/immobile solute transport, *Water Resources Research*, *39*(10).

795 Schumer, R., and D. Jerolmack (2009), Real and apparent changes in sediment deposition

796 rates through time, *Journal of Geophys. Res. - Earth Surface*, *in review.*

797 Singh, A., K. Fienberg, D. J. Jerolmack, J. Marr, and E. Foufoula-Georgiou (2009),

798 Experimental evidence for statistical scaling and intermittency in sediment transport

799 rates, *Journal of Geophysical Research-Earth Surface*, *114*.

800 Stark, C., E. Foufoula-Georgiou, and V. Ganti (2009), A nonlocal theory of sediment

801 buffering and bedrock channel evolution, *J. Geophys. Res.*, *114*(F01029).

802 Tadjeran, C., M.M. Meerschaert, and H.P. Scheffler (2006) A second order accurate nu-

803 merical approximation for the fractional diffusion equation. *J. Comput. Phys.* **213**,

804 205–213.

805 Tipper, J. C. (1983), Rates of sedimentation and stratigraphical completeness, *Nature*,

806 *302*(5910), 696–698.

807   Worman, A., A. I. Packman, L. Marklund, J. W. Harvey, and S. H. Stone (2007), Fractal

808   topography and subsurface water flows from fluvial bedforms to the continental shield,

809   *Geophysical Research Letters*, *34* (7).

810   Zhang, Y., D. A. Benson, M. M. Meerschaert, and H. P. Scheffler (2006) On using random

811   walks to solve the space-fractional advection-dispersion equations. *Journal of Statistical*

812   *Physics*, **123(1)**, 89–110.

813   Zhang, Y., D.A. Benson, and B. Baeumer (2008) Moment analysis for spatiotemporal

814   fractional dispersion. *Water Resources Research,44* (4).
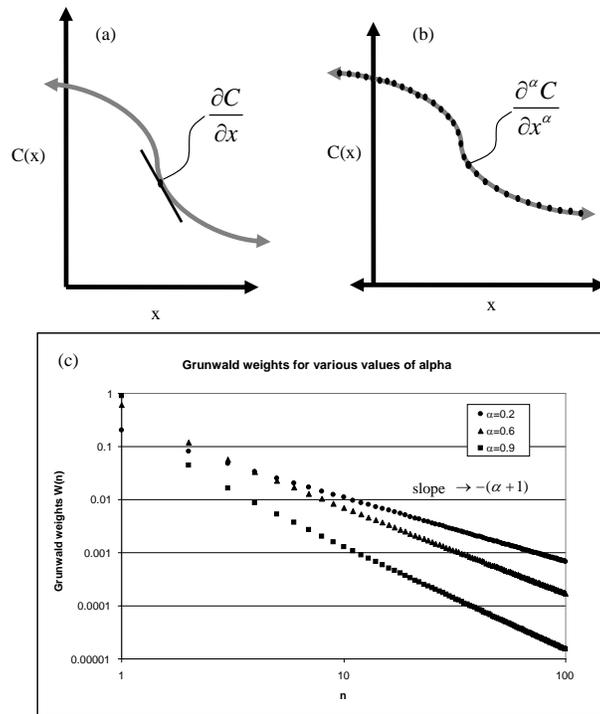
Figure 1. The usual derivative (a) is local but the fractional derivative (b) is non-local.

The log-log plot (c) demonstrates the power law decay in Grunwald weights.
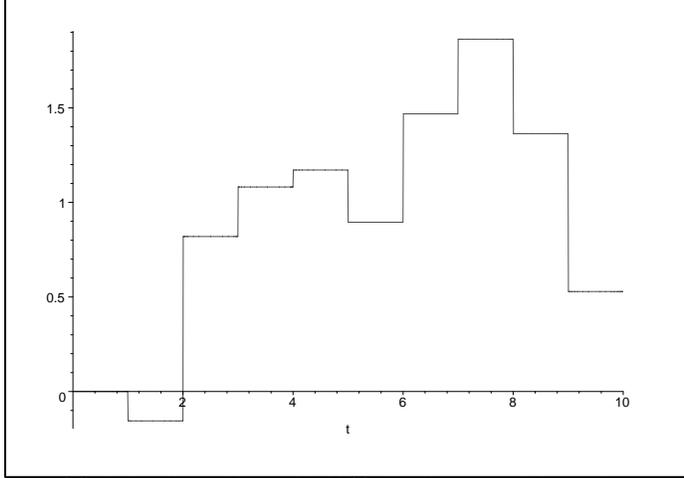
(a) A simple random walk $X(t)$ simulates particle motion.
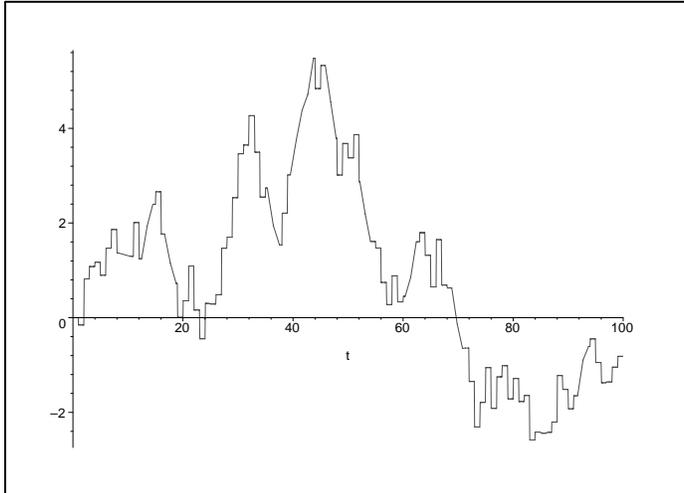
(b) Simple random walk $X(t)$ at a longer time scale.
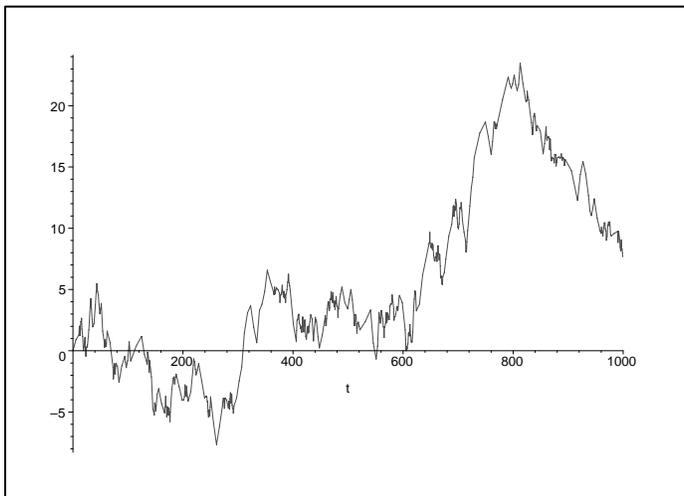
(c) Brownian motion as the scaling limit of a random walk $X(t)$. Particle

graph is a random fractal with dimension $3/2$.

Figure 2.   Comparison of random walk traces at increasing time scale and in the scaling

limit as a Brownian motion.
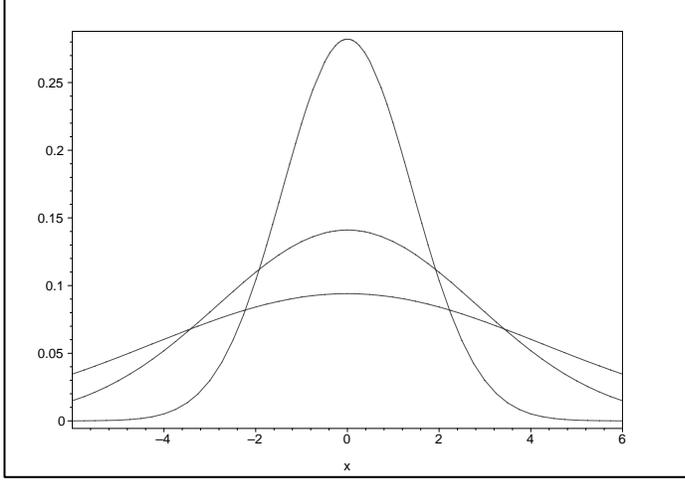
Figure 3. Brownian motion density function $C(x,t)$ describing particle spreading away from plume center of mass at time $t = 1, 4, 9$ in the scaling limit. Note the square root spreading rate and fast tail decay.
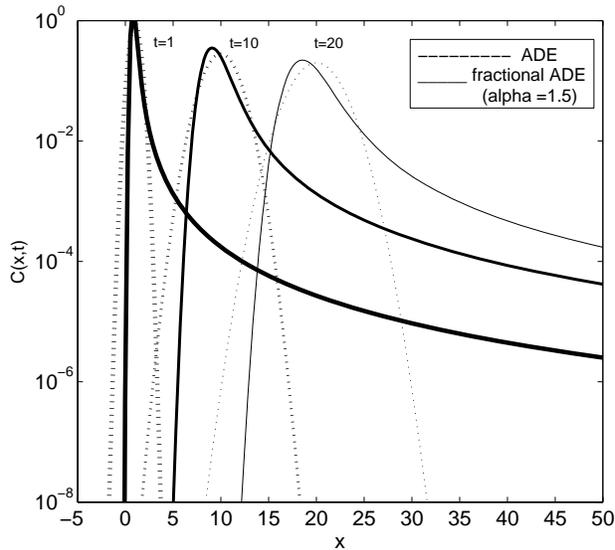
Figure 4. Comparison of the solutions in space to integer-order and fractional ADEs at time $t = 1$, $t = 10$, and $t = 20$. For both models, $v = 1$ and $D = 0.1$. The heavy leading edges of the α-stable solutions to the fractional ADE decay as $C(x) \approx x^{-\alpha-1}$.
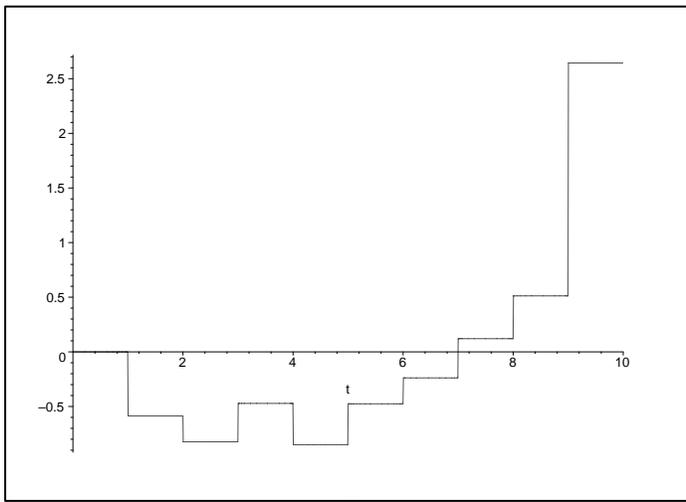
(a) A heavy tailed random walk $X(t)$ simulates anomalous particle motion.
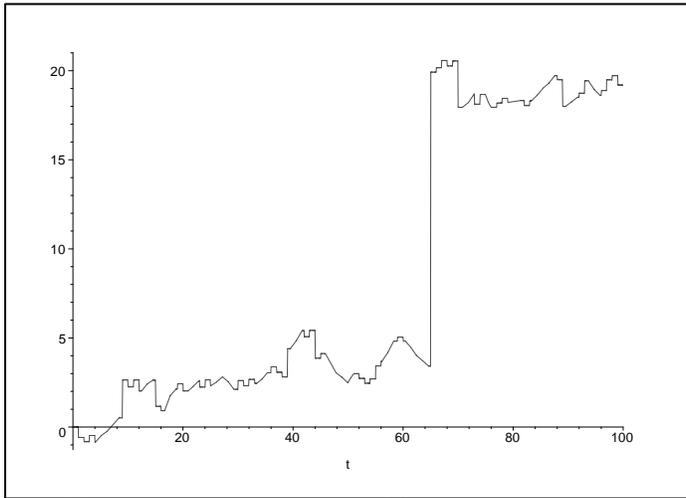
(b) Heavy tailed random walk $X(t)$ at a longer time scale.
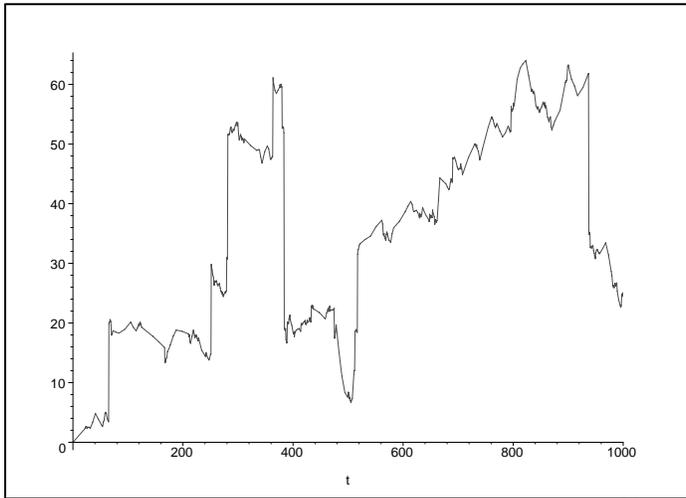
(c) Stable Lévy motion as the scaling limit of a random walk $X(t)$. Particle graph is a random fractal with heavy tailed jumps.

Figure 5. Comparison of heavy tailed random walk traces at increasing time scale and
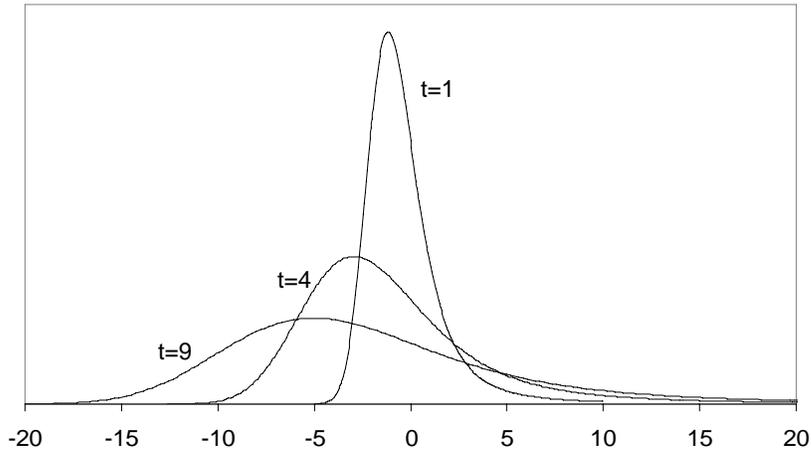
Figure 6. Lévy motion density function $C(x, t)$ describing particle spreading away from plume center of mass at time $t = 1, 4, 9$ in the scaling limit. Note the super-diffusive spreading rate, skewness, and power law tail.
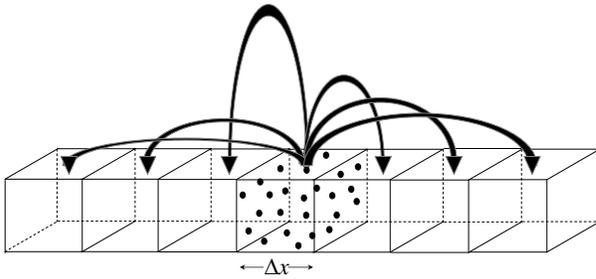

Figure 7. When governed by a fractional-in-space ADE, particle hop length during a small time step Δt may be much larger than the mean size. In finite-difference implementations for fractional ADEs, probability of particle hop length decays as a power law with distance. (After *Schumer et al.* [2001].)
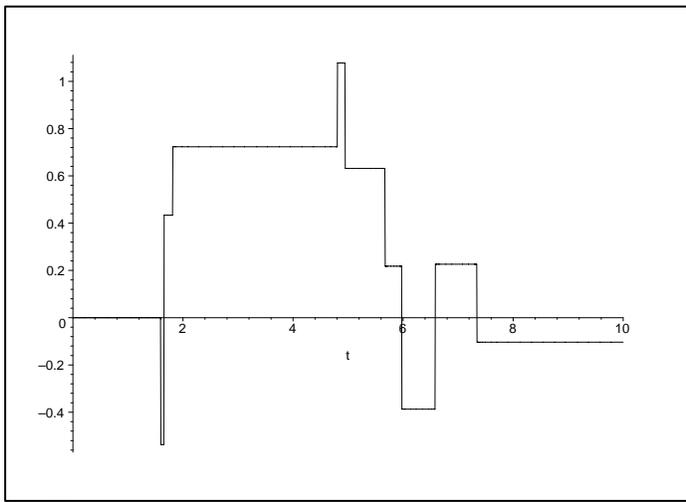
(a) CTRW with heavy tailed waiting times simulates sub-diffusive particle motion.
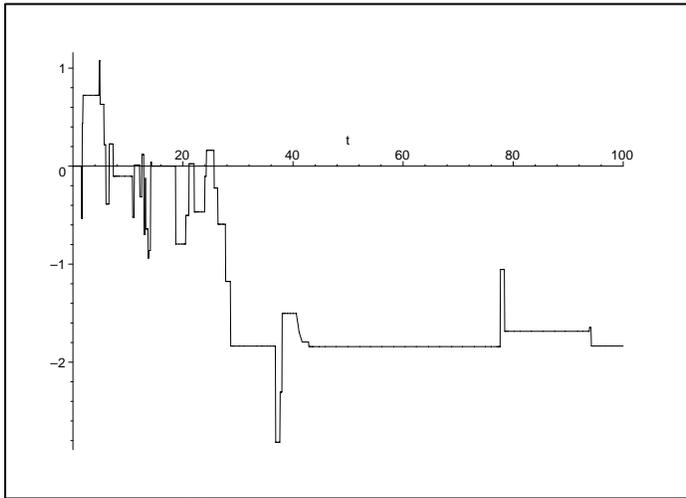
(b) CTRW at a longer time scale.
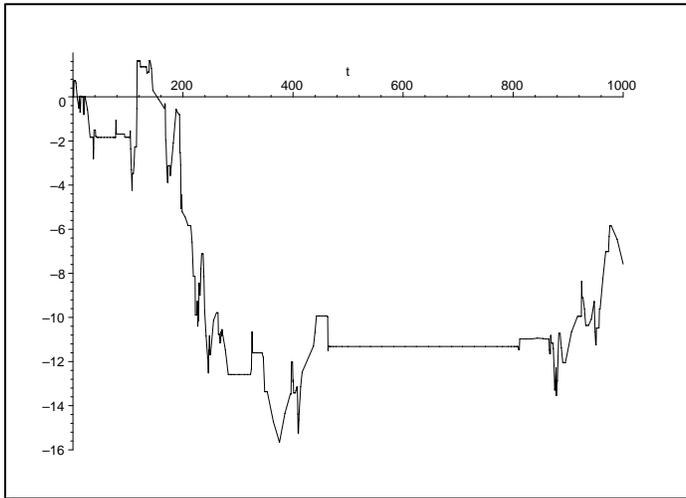
(c) Subordinated Brownian motion as the CTRW scaling limit. Long resting times persist in the limit process.

Figure 8. Comparison of heavy tailed waiting time CTRW traces at increasing time

scale and in the scaling limit as a subordinated Brownian motion.
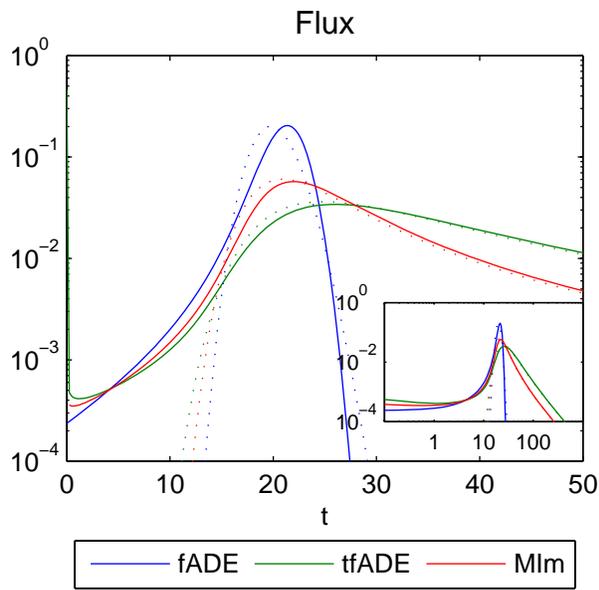
Figure 9. Flux at a distance of $x = 20$ vs. time in semi-log and log-log plots for solutions to equations (19), (23) and (26) with $v = 1$, $D = 0.1$, $\gamma = 0.8$ and $\beta = 1$ where applicable. Dotted lines represent the flux with $\alpha = 2$, solid lines have $\alpha = 1.5$.
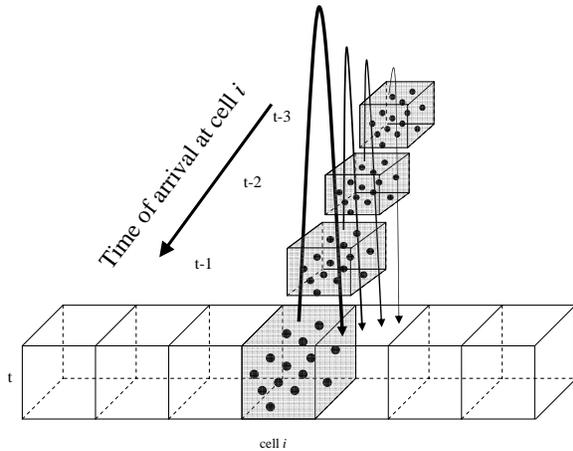

Figure 10. When governed by a fractional-in-time ADE, particles have memory of the time that they arrive at a given point. Their probability of release decays as a power law from arrival time.

limiting stochastic process ↓ governing equation ↓ solution	finite-variance jump length distribution	infinite-variance jump length distribution
finite-mean waiting time distribution	Brownian motion ↓ ADE ↓ Gaussian density	Lévy motion ↓ fractional-in-space ADE ↓ α-stable density
infinite-mean waiting time distribution	Brownian motion subordinated to inverse Lévy Process ↓ fractional-in-time ADE ↓ subordinated Gaussian density	Lévy motion subordinated to inverse Lévy Process ↓ fractional-in-time and space ADE ↓ subordinated α-stable density

Figure 11. CTRW converge to different limiting stochastic process depending on the tail characteristics of the jump length distribution and waiting time distribution. Particles undergoing the limiting stochastic process are governed by ADEs with either integer-order or non-integer-order derivatives with solutions that are related to probability density functions. A CTRW with finite-mean waiting time distribution converges in the scaling limit to the same stochastic process as the analogous random walk.