Time and space nonlocalities underlying fractional-derivative models: Distinction and literature review of field applications

Yong Zhang a,*, David A. Benson b, Donald M. Reeves c

a Division of Hydrologic Sciences, Desert Research Institute, 755 E Flamingo Road, Las Vegas, NV 89119, United States
b Department of Geology and Geological Engineering, Colorado School of Mines, Golden, CO 80401, United States
c Division of Hydrologic Sciences, Desert Research Institute, 755 E Flamingo Road, Las Vegas, NV 89119, United States

A R T I C L E   I N F O

Article history:
Received 6 June 2008
Received in revised form 13 January 2009
Accepted 19 January 2009
Available online 3 February 2009

Keywords:
Anomalous dispersion
Space nonlocality
Time nonlocality
Fractional ADE
Field application

A B S T R A C T

We investigate the spatiotemporal nonlocality underlying fractional-derivative models as a possible explanation for regional-scale anomalous dispersion with heavy tails. Properties of four fractional-order advection-dispersion equation (fADE) models were analyzed and compared systematically, including the space fADEs with either maximally positive or negative skewness, the time fADE with a temporal fractional-derivative \( 0 < \gamma < 1 \), and the extension of the time fADE with \( 1 < \gamma < 2 \). Space fADEs describe the dependence of local concentration change on a wide range of spatial zones (i.e., the space nonlocality), while time fADEs describe dynamic mass exchange between mobile and multiple immobile phases and therefore record the temporal history of concentration “loading” (i.e., the time nonlocality). We then applied the fADEs as models of anomalous dispersion to four extensively-studied, regional-scale, natural systems, including a hillslope composed of fractured soils, a river with simultaneous active flow zones and various dead-zones, a relatively homogeneous glacifluvial aquifer dominated by stratified sand and gravel, and a highly heterogeneous alluvial aquifer containing both preferential flowpaths and abundant aquitards. We find that the anomalous dispersion observed at each site might not be characterized reasonably or sufficiently by previous studies. In particular, the use of the space fADE with less than maximally positive skewness implies a spatial dependence on downstream concentrations that may not be physically realistic for solute transport in watershed catchments and rivers (where the influence of dead-zones on solute transport can be described by a temporal, not spatial, fractional model). Field-scale transport studies show that large ranges of solute displacement can be described by a space nonlocal, fractional-derivative model, and long waiting times can be described efficiently by a time nonlocal, fractional model. The unknown quantitative relationship between the representative nonlocality (time and/or space) for any given regional-scale anomalous dispersion process.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Dispersion of aqueous tracers in natural systems including heterogeneous soils, aquifers, and rivers, is typically observed to be non-Fickian, also called “anomalous”. Anomalous transport (despite the somewhat misleading nomenclature) is very often found at all scales – from pore scale simulations [144], soil cores [24,33,92,116], laboratory experiments [70,79,119,133], to field-scale observations (e.g., [2,11–13,45,65,83,87,143], among many others). Anomalous transport may be characterized by non-Gaussian leading or trailing edges (also called heavy tails) of a plume emanating from a point source, or nonlinear growth of the centered second moment. If the growth rate is faster than linear, the transport is anomalous superdiffusion; slower than linear growth rate is subdiffusion [21]. Far from a purely academic exercise, the extreme non-Gaussian (heavy) tails of the solute transport dominate problems associated with toxic chemicals: early arrivals pose the most risk, while late arrivals dominate the cost and strategy of cleanup. Proper modeling of non-Fickian or non-Gaussian transport of tracers is a truly important step that has not been achieved completely, even after decades of effort.

Non-Fickian transport behavior may be due to different mechanisms. First, the complex flow velocity in natural, multi-scale heterogeneous media leads to anomalous spreading of a conservative tracer that deviates significantly from a Fickian model. For example, preferential flow paths can enhance spreading of the plume front, resulting in anomalous superdiffusion. Second, mass exchange between relatively mobile and immobile zones can retard the movement of aqueous tracers and result in non-Gaussian

* Corresponding author. Tel.: +1 702 862 5314; fax: +1 702 862 5427.
E-mail address: Yong.Zhang@dri.edu (Y. Zhang).

0309-1708/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2009.01.008
concentration profile, which is typically referred to as a subdiffusive process [21]. The second mechanism is due to chemical reactions (including sorption/desorption), where the diffusion and dispersion of solute in the mobile phase can still be Fickian. See also the recent discussion by Berkowitz et al. [19].

Numerous numerical experiments indicate that anomalous dispersion cannot be described by the traditional second-order advection–dispersion equation (ADE) without extremely detailed information on the connectivity of high and low hydraulic conductivity (K) sediments [2,5,1,13]. Five alternative, nonlocal transport techniques have been developed for the use of the general hydrology community. These techniques are the Stochastic Averaging of the classical ADE (SA-ADE) method, the (single- and) multiple-rate mass transfer (MRMT) method (see the recent work [64,65,66,69], also see [25,26,30,60,67,109,134,138,149,154], among many others), the continuous time random walk (CTRW) method [18,47,48,108], the time fractional advection–dispersion equation (fADE) method (with the time scale index \(0 < \gamma < 1\)) [8,9,128,147], and the space fADE method [11–13,29,129,148]. These alternative conceptual models describe nonlocal dependence on either time or space, or both, to explain the development of anomalous dispersion. Here the nomenclature “nonlocal” follows the definition given by Cushman [35]: “If the constitutive variable depends on what is happening at a point in space–time or a very small neighborhood of the point, the variable is said to be local and derived from a local theory . . . On the other hand, if information is needed to define the constitutive variables from regions of space and time distinct from a neighborhood of the space–time point where evaluation of the variables is to be made, then the theories and constitutive variables are said to be nonlocal in character”. More specifically, the concentration change at some location in time might depend on a wide variety of locations upstream (i.e., space nonlocality) [39,106,127,148], and it also might depend on the temporal history of concentration “loading” at that location (time-nonlocality) [36,37,48,49,111,128]. The temporal nonlocality can be physically attributed to mass transfer of solute between relatively immobile and mobile phases (i.e., the sorption/desorption mechanism discussed above) [47,65] and transport in segregated regions of high and low permeability [13,15,149,155]. The spatial nonlocality might be due to the high variation and long-range dependence (or long spatial autocorrelation) of permeability [62,71,82,140,148], including the preferential flow paths discussed above.

Among the five nonlocal approaches, the most commonly studied one is the SA-ADE method that holds on a local scale and has random velocity. It is well recognized that the stochastic averaging of local ADEs leads to space and time-nonlocal effective transport equations (e.g., [34–38,41–43,59,80,81,84,85,110–112], among many others). The SA-ADE model is attractive in many ways, as reviewed recently by Neuman and Tartakovsky [113]. For example, (1) the space–time memory kernels embedded in the governing equation of solute concentration have general forms that can describe various nonlocal transport processes, (2) the SA-ADE model can employ physically observable properties (e.g., the correlation length of hydraulic conductivity fluctuations) to estimate the model parameters, and (3) it can be used for nonstationary and correlated velocity fields. The other nonlocal transport models, including the MRMT, CTRW, and the time fADE models, generally attribute anomalous transport to the very long times that solute “particles” move very slowly; while the space fADE model applies non-Gaussian motions directly to the solutes. The MRMT, CTRW and fADE models do not assume Fickian diffusion at a reference length, and the CTRW framework. We also link the fADE models to other nonlocal representations of non-Fickian transport (see Appendix). This link provides additional and helpful information about the time and space nonlocality underlying the fADEs.

This study aims to solve several critical issues related to the fADE models. First, the fundamental difference between the time and space fADEs has not been illustrated, either using numerical experiments or field applications. Possible mis-application of fADEs has increased recently. For example, the space fADE has been used to model various observations in quite different media (for an incomplete list of recent publications in hydrology journals, see [27,45,46,74,143]), no matter what the possible dominant nonlocality is likely to be. To guide the correct application of fADEs, this study focuses on the prior comparisons of fADE models and fundamental evaluations of the dominant nonlocality. In the next section, we calculate and analyze the concentration profiles and plume snapshots systematically for the time fADE and the space fADE, respectively.

The second objective is to extend and test the applicability of the time fADE. The temporal nonlocal fractional-order ADE model, which has limited range for its time scale index \(0 < \gamma < 1\) (see details in the next section), was proposed by Schum et al. [128]. Its application to solute transport modeling, however, is rather rare, probably due to the computational burden of the available subordination solution of the time fADE [128]. We apply the computationally efficient, particle tracking, approach developed recently by Zhang et al. [150] to approximate and investigate the applicability of the time fADE model. We also explore systematically a time fADE model (with the time scale index \(1 < \gamma < 2\)) proposed by Baumer et al. [5,6], where the underlying time-nonlocality can describe not only the trapping process, but also the large displacement of solutes which could be confused with a space nonlocal process. The extension of the time scale index \(\gamma\) makes it possible to describe a wider range of solute transport processes.

We then distinguish between time and space nonlocalities by revisiting four previous applications of fADEs and rebuilding the nonlocal model for each case to simulate the anomalous dispersion observed in various natural systems. We also analyze the underlying (usually limited) hydrogeological information and try to relate quantitatively the heterogeneity of variable depositional systems to the nonlocal transport parameters/processes embedded in the fADEs. Other important parameters, including the average velocity, are compared to field measurements if available. These preliminary efforts can be beneficial to the building of an eventually predictive fADE model.

Field cases and numerical results in this study show clearly the importance of distinguishing between time and space nonlocalities when using fractional-derivative techniques. We also expect that this study can provide some simple diagnostics that the practitioner can use to identify when fADE models can be useful.

2. Nonlocalities captured by the fractional-derivative models

The fractional-order differintegral, a mathematical tool initiated by Leibniz in 1695 and developed further by Liouville in
1832–1837, is almost as old as its integer-order counterpart [105,115]. The fractional-derivative operator has been introduced into the diffusion equation by the physics community, and remains an active area of research for the simulation of anomalous diffusion (see the excellent reviews by Metzler and Klafter [103,104] and the numerous references cited therein). When the ADE is generalized with a fractional-derivative in time, the underlying stochastic (particle motion) process is based on a broad waiting time distribution. Similarly, when the second-order dispersion term in the ADE is replaced by a spatial fractional-derivative, the underlying particle motion process is based on a broad jump size distribution. The underlying CTRW has independent, identically-distributed (iid) jumps and iid duration times between two jumping events (see Appendix A.2), so that the fADE models discussed in this study should not be confused with correlated CTRW processes. In the following subsections, we introduce and evaluate each fractional kinetic model.

2.1. The time fADE model with index 0 < γ < 1

If the waiting times of solute particles have infinite mean (where the probability distribution is assumed to decline algebraically with index 0 < γ < 1) and the jump sizes have finite mean and variance, then the scaling limit of such a CTRW has a probability density $C(x,t)$ and variance, then the scaling limit of such a CTRW has a probability density $C(x,t)$ that evolves according to the following fractional kinetic equation [8,97]:

$$\frac{\partial C(x,t)}{\partial t} = -V \frac{\partial C(x,t)}{\partial x} + D \frac{\partial^2 C(x,t)}{\partial x^2} + \frac{\delta(x) t^{-\gamma}}{\Gamma(1-\gamma)},$$

which was first proposed (without identifying the limit process) by Zaslavsky [142] for Hamiltonian chaos. The probability density is equal to the normalized solute concentration. Parameters $V$ and $D$ in (1) denote the fractional velocity and fractional dispersion coefficient, respectively; $\Gamma(\cdot)$ is the Gamma function. The term $\frac{\partial^\gamma}{\partial t^\gamma}$ represents the fractional derivative, which is most easily defined in terms of Laplace transforms: $\frac{\partial^\gamma}{\partial t^\gamma} f(t)$ has Laplace transform $\mathcal{L}\{f(t)^n\}$, where $f(u) = \int_0^\infty e^{-ut}f(t)dt$, generalizing the familiar formula for integer-order derivatives.

The model (1) was regarded as a subset of the CTRW framework [18], and it describes solute particles sticking or trapping in relatively immobile domains. As also shown by Metzler and Klafter [103], the CTRW corresponding to (1) has a long-tailed waiting time pdf with long time asymptotic decline $w(t) \sim t^{-\gamma}$, so the order of the fractional time derivative ($\gamma$) governs the probability tails of the delay between jumps. The development of model (1) using the generalized Master equation is shown in Appendix A.1. A small index $\gamma$ in the time fADE (1) models a relatively high probability for long trapping times, and thus it describes solute transport through a strongly heterogeneous medium where the trapping times have a broad distribution.

Previous research [20,47,128] established a connection between CTRW and MRMT models. In particular, Dentz and Berkowitz [47] linked the transition time distribution of a CTRW framework (see their Eq. (16)) to the defining transition time distribution of the MRMT model (see their Eq. (33)) by defining an identical memory function (see their Eq. (42) and Appendix C). They also provided the mobile/immobile CTRW models by dividing the transition time into mobile and immobile parts (Eq. (11)), where the mobile time is constant. Benson and Meerschaert [14] further interpreted the MRMT memory function as the rate at which particles of a certain age, measured by residence time in the immobile zone, exit to become mobile once again. The stochastic process that exactly corresponds to the MRMT equations requires exponentially distributed times spent by mobile particles. The separation of mobile and immobile times provides an efficient approximation of the MRMT model using random walks. To distinguish explicitly the mobile and immobile status using the fractional dynamics, Schumer et al. [128] developed the following fractional-order, mobile/immobile (MIM) model for the total concentration:

$$\frac{\partial C_{\text{tot}}}{\partial t} + \beta \frac{\partial C_{\text{tot}}}{\partial x} = -V \frac{\partial C_{\text{tot}}}{\partial x} + D \frac{\partial^2 C_{\text{tot}}}{\partial x^2},$$

(2)

where $C_{\text{tot}}$ denotes the solute concentration in the total (mobile + immobile) phase, and $\beta$ is the fractional capacity coefficient. Here $V$ and $D$ are the velocity and dispersion coefficient for the mobile phase (and hence may be directly measured). Compared to (1), the time fADE (2) is more flexible due to the new term $\beta \frac{\partial C_{\text{tot}}}{\partial t}$ and the additional parameter $\beta$. The time drift term $\beta \frac{\partial C_{\text{tot}}}{\partial t}$ is added to describe the motion time and thus helps to distinguish the status of particles conveniently (see also the discussion in [14]). The corresponding fractional-derivative governing equations for the mobile and immobile phases can be found in [128] (see their Eqs. (15) and (16)). When $\gamma = 1$, the time fADE (2) reduces to the ADE with a retardation factor $1 - \beta$ (see also [128]).

The model (2) is identical to the one presented by Carrera et al. [26] and the MRMT model [65] with a non-truncated power-law distribution of rate coefficients. The MRMT model is the advection–dispersion–reaction equation that assumes a number of kinetic rates (first-order reversible kinetic sorption) to characterize the diffusion between a mobile region and parallel immobile zones of different sizes. The development of the multi-domain diffusion model has a long history, starting probably from the Griddings–Eyring two-site adsorption chromatography model [e.g., 63]). The two-site model assumes a single mobile domain with two reaction sites – one is an equilibrium sorption reaction, and the other is a kinetic-controlled reaction. The two-domain model was then developed by Coats and Smith [30] and extended by many other researchers, including van Genuchten and Wierenga [138], Rao et al. [120], and Brusseau et al. [25]. In this model, the total domain is divided into mobile and immobile sub-domains and the transport between the sub-domains is controlled by first-order reversible kinetics or equilibrium sorption (see also Nkedi-Kizza et al. [114] for the equivalence between the two-site and two-domain models). To predict the slower than exponential decline of the late-time solute concentrations observed in the field, Haggerty and Gorelick [64] developed the multi-domain model by replacing the single mass transfer term with a series of rates. Haggerty et al. [65] provided various possible distributions of rate coefficients, where the power-law distribution is of particular interest. Schumer et al. [128] further showed that the power-law distribution of rate coefficients leads to a solution that is scale invariant in time and is identical to an advection–dispersion–reaction equation with a fractional time derivative, as shown by (2).

The solution of the time fADE (2) can be approximated by the non-Markovian random walk approach developed by Zhang et al. [150,151]. We first approximated model (2) with variable indexes $\gamma$. An instantaneous source was released in the mobile domain, representing a typical contamination process. The influence of time-nonlinearity on solute transport can be seen from the decline of mobile mass (for details, see [128]), which is also well-documented in field experiments (see for example, the excellent discussion in [69]). In the case where the mobile mass is not available, the snapshot of the total concentration (whose mass will not decline) still reveals the existence of time-nonlocal transport, such as the retention of solute in the immobile domains. Smaller values of $\gamma$ describe a higher probability for large waiting times between particle jumps and an increased retention signature near the source (as demonstrated by the simulated snapshots shown in Fig. 1a). Double peaks (in space) can form if the waiting/trapping time tail is heavy enough (e.g., $\gamma = 0.1$). The leading edge of the
The snapshot appears very similar to the un-retarded motion process (in this case a Gaussian curve) (Fig. 1b), because the motion of individual solute particles follow Fick’s law with a mean drift. In addition, the location of the plume peak does not change significantly with the index \( \gamma \), although the magnitude of peak concentration is controlled by \( \gamma \); since lower values of \( \gamma \) describe greater retention and more solute mass behind the peak.

Another important feature of time-nonlocal transport is the “transfer” of the decline rate of large waiting times to the late tail of the flux concentration (\( C_f \)) based profile (i.e., the breakthrough curve or BTC). The latter, as shown by Fig. 1c, also follows the power-law trend \( C_f(t) \sim t^{-\gamma} \). Note the resident concentration \( (C_R) \) declines as \( C_R(t) \sim t^{-\gamma} \), much slower than \( C_f \), since the flux concentration is defined as the time derivative of the resident concentration \( C_f(x,t) = \frac{1}{\gamma} \int_0^t C_R(x,t) \, dt \), as shown by Kreft and Zuber ([86], see their Eq. (T1)) and Zhang et al. ([147], Eq. (2)). Note the direct comparison of flux and resident concentrations can be misleading if users do not pay attention to the different physical meanings underlying these two quantities: the resident concentration denotes a distribution density in space, while the flux concentration denotes a distribution density in time [147].

The power-law decline of the late tail on a BTC is the most apparent property of the subdiffusive process described effectively by the fractional time-nonlocal model (2). However, we emphasize here that the duration of observation may not be long enough to obtain completely the late-time tail of BTCs, as demonstrated by some field data (Experiment 2) discussed in the next section.

We then approximated the time \( fADE(2) \) with variable parameter \( \beta \) (Fig. 2). A large fraction of immobile material, corresponding to large \( \beta \), delays the average motion of particles, and consequently, the plume peak. The resulting spatial snapshot of concentration in the total phase skews more positively, simply due to more mass trapped near the source. The power-law index of the late-time concentration breakthrough curve, however, is independent of \( \beta \). This is because the tail of the waiting time pdf is defined by the index \( \gamma \).

![Fig. 1. Simulated concentrations in the total phase using the random walk method. (a) The snapshot of tracers whose transport is governed by the time \( fADE(2) \) with the time scale index \( \gamma = 0.1–0.9 \). The other parameters are: \( \beta = 0.1, V = 1, D = 0, T = 20 \), and the initial source is located at \( x_0 = 10 \). The classical ADE (dashed line, with fractional velocity \( V_f = V/(1 + \beta) \) and fractional dispersion coefficient \( D_f = D/(1 + \beta) \)) is also shown for comparison. (b) The semi-log version of (a). (c) The concentration profile (i.e., the evolution of the resident concentration with time). The control point is 5 unit lengths downstream. The correspondent flux concentration (the long-dashed line) is shown for the case of \( \gamma = 0.9 \).](image1)

![Fig. 2. Concentration in the total phase. (a) The snapshot of tracers whose transport is governed by the time \( fADE(2) \) with the time scale index \( \gamma = 0.5 \) and the capacity coefficient \( \beta = 0.01–1.2 \). The other parameters are: \( V = 1, D = 0, T = 60 \), and the initial source is located at \( x_0 = 10 \). The classical ADE (dashed line, with velocity \( V \) and dispersion coefficient \( D \)) is also shown for comparison. (b) The semi-log version of (a). (c) The concentration profile. The control point is 5 unit lengths downstream. The correspondent flux concentration (the long-dashed line) is shown for the case of \( \beta = 0.01 \).](image2)
The prior estimation of the fractional capacity coefficient $\beta$ remains as an open question. The classical capacity coefficient was defined by Haggerty and Gorelick [64] as the ratio of mass in the immobile and mobile phases at equilibrium. In the absence of chemical sorption, Schumer et al. [128] defined $\beta$ as the ratio of porosity in the immobile versus mobile zones. The Monte Carlo experiments conducted by Zhang et al. [149] demonstrated that $\beta$ is proportional to the ratio of volumes of the immobile versus mobile zones. The most recent work of Benson and Meerschaert [14] revealed further that $\beta$ can actually be defined by the ratio of expected time in the immobile versus mobile zones. Theoretically, the fractional capacity coefficient varies with space, especially for the nonstationary heterogeneous medium where the fractional capacity coefficient varies with space, especially for the nonstationary heterogeneous medium where the volumetric proportion of different sediments and/or the heterogeneity structure (that governs solute transport) can evolve in space. Field measurement of the variation of the fitting parameter $\beta$ is difficult, since it is difficult to sample the average time that solute particles spend in the immobile domains.

Similar transport has also been studied by Dentz et al. [48], see their Figs. 2, 3, 6 and 7) using the CTRW framework with an exponentially truncated power-law memory function. The time fractional ADE model (1) is functionally equivalent to the CTRW model developed by Dentz et al. [48], see their Eq. (8), if the CTRW memory function can be approximated by a power-law with the exponent between 0 and 1. It is noteworthy that the time fractional ADE (2) has an additional factor (i.e., the fractional capacity coefficient) and the capability of separating the mobile and immobile phases, compared to model (1).

2.2. The time fractional ADE model with index $1 < \gamma < 2$

When the waiting times of trapped solute particles have infinite variance, the corresponding mean may not always diverge. Becker-Kern et al. [9] and Baeumer et al. [5] found that the scaling limit of the CTRW with finite mean and infinite variance waiting times is the following fractional-in-time equation:

$$\frac{\partial C_{\text{tot}}}{\partial t} - \beta \frac{\partial^\gamma C_{\text{tot}}}{\partial x^\gamma} = -V \frac{\partial C_{\text{tot}}}{\partial x} + D \frac{\partial^2 C_{\text{tot}}}{\partial x^2} + f(x)\delta(t), \quad (3)$$

where the function $f$ is uniquely determined by the requirement of (3) having a mild solution [5].

Note that the time fractional ADEs (Eqs. (3) and (2)) divide the total time into two components, including the mobile time and an immobile waiting time, to distinguish the solute status (see also Appendix A.2). The time for particle to move is the mobile time only, so that the solute velocity is not the ratio of the displacement to the total time. The kinematic relationship (where velocity $= \text{displacement}/\text{total time}$) providing the Lagrangian velocity correlation was used by Dentz et al. [50] to build a correlated random walk model, where the total time was not separated into mobile and immobile components.

Here waiting times have heavy tails of order $1 < \gamma < 2$, and the probability of waiting longer than $t$ declines like $t^{-\gamma}$. When $1 < \gamma < 2$, the scaling limit of the operational time process is split into two parts, the mean waiting time and the deviation about the mean, to preserve the near- and intermediate-time behavior [9]. The number of jumps is associated with the max-process [14], since deviations from the mean waiting time may be positive or negative. A negative waiting time physically means that a particle makes a number of simultaneous jumps until the accumulated deviations once again become positive, so that the max always increases [150]. When $0 < \gamma < 1$, however, the waiting time is always positive and thus no particles can jump more than once within each operational time step.

The time-nonlocal model (3) has not been applied to any practical problem yet. Here we explore the nonlocal properties characterized by this model, and then test its applicability in the next section. Particles moving successively during one operational time step form the fast moving front (Fig. 3a), describing the local variation of velocity (especially the relatively high velocities). Although the corresponding leading edge is heavier than the model with infinite mean waiting times (i.e., model (2) and Fig. 1a and b), both leading edges decline exponentially versus distance because the dispersion of mobile particles in both models follows the Fick's law. The plume peak moves faster with decreasing $\gamma$ (Fig. 3b), due to the property of the corresponding $\gamma$-stable density [126] defined for the waiting times. Lower values of $\gamma$ increase the tendency for negative deviations about the mean waiting time, resulting in a greater number of simultaneous jumps. Similarly, early breakthrough is much heavier (Fig. 3c) than that of model (2) (Fig. 1c), but it is not as heavy as a power-law.

On the other hand, the retention of particles, especially at the source, is still preserved in the plume snapshot, and the resultant late-time tail of BTC declines as power-law (Fig. 3c), similar to the process described by model (2). Note that model (3) is applicable for a plume with a total solute late-time BTC slope between $-2$
and \(-3\), while model (2) is limited to a slope between \(-1\) and \(-2\). A different model (CTRW, MRMT) is required if an observed BTC for total concentration exhibits a slope steeper (i.e., less-heavy) than \(-3\).

We then explore anomalous behavior caused by the capacity coefficient \(b\) in (3). Similar to model (2), greater values of \(b\) enhance the positive skewness of plumes by retarding more mass near the source (Fig. 4). When the system has very limited immobile material (i.e., \(b \rightarrow 0\)), (3) reduces to the classical ADE and the resultant plume is symmetric (as demonstrated by the case of \(b = 0.01\) in Fig. 4).

2.3. The space fADE model with maximally positive skewness

The time-nonlocal transport models discussed above describe the trapping of particles in an immobile phase, where the solutes, when mobile, follow the classical, second-order ADE, similar to the other time-nonlocal techniques including the CTRW and MRMT. While sampling heterogeneous soils/aquifers, solute particles can experience various velocity zones. If the complex heterogeneity structure, such as the spatial connectivity, can facilitate movement of particles within a certain scale, fast motions may no longer obey the classical Fick’s law and may indeed have a probability density function that follows a power-law [71]. Superdiffusion is one possible form of fast motions.

A modeling equation that extends the classical ADE

\[
\frac{\partial P}{\partial t} = -V \frac{\partial P}{\partial x} + D \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial(-x)^2} + D \frac{\partial^2 P}{\partial x^2}
\]

was introduced by Saichev and Zaslavsky [124], Benson [10], and Chaves [28], extending earlier pioneering work [31,32,102,118,131,142] on anomalous diffusion that arises from space nonlocality. \(P\) is the density of the particle location after release at time zero at the origin. The parameter \(0 \leq q \leq 1\) called “skewness” represents the proportion of high-velocity “jets” in the direction of flow, and \(\partial^q/(\partial x)^q\) fractional derivatives of order \(1 < q < 2\). The solution to (4) follows \(P(x,t) = (Dt)^{-\alpha/\beta} \Phi_0 [x-Vt] (Dt)^{-\alpha/\beta} \), so the scaling parameter \(H = 1/\alpha\) is often called the Hurst index. When \(x = 2, (4)\) models the density of a diffusive Brownian motion. When \(x < 2, (4)\) models superdiffusive growth according to a stable Lévy motion, see for example the review [103]. The derivation of (4) is also shown in Appendix.

To simulate the anomalously rapid transport of contaminants in heterogeneous systems, we choose \(q = 1\), and thus (4) reduces to the following space fADE:

\[
\frac{\partial C(x,t)}{\partial t} = -V \frac{\partial C(x,t)}{\partial x} + D \frac{\partial^2 C(x,t)}{\partial x^2} + \frac{\partial^2 C(x,t)}{\partial(-x)^2}
\]

where the dispersive term \(D\partial^2 C(x,t)/\partial x^2\) follows the fractional (also called the generalized) Fick’s law [28,77,117,127].

The choice \(q = 1\) is mainly based on the observation that the largest particle motions through natural media are usually in front of the mean (see also the explanation in [10,11,13,127]). The distribution of the greatest groundwater velocities at the fine-scale is also assumed to be power-law with a positive skewness. Large “jumps” behind the mean are most likely due to immobilization and are properly described by the trapping models in Section 2.1. Pareto distributions of hydraulic conductivity (with a heavy tail for large values) observed in the field (see for example, [11]) supports the assumption of positively skewed particle displacements.

If the net inflow of solute flux at one point depends on not only the current position but also properties of a wide range of upstream zones, then the concentration change at the local point is a space nonlocal process [35]. If the space dependence declines as power-law with distance, then the corresponding transport process can be described by the space fADE (5) [148].

We approximated the space fADE (5) using random walks following the Markovian-type particle-tracking scheme developed recently by Zhang et al. [145,146]. Four dominant characteristics of the model are observed. First, in plume snapshots (Fig. 5a), the speed of the plume peak decreases with a decreasing \(x\). Since more solute mass is transferred to the leading plume edge as values of \(x\) decrease (more superdiffusive), the peak lags behind to compensate for the redistribution of solute mass. The influence of peak lag can also be observed in the concentration profile from a downstream control point (Fig. 5c).

Second, backward dispersion is weak in the maximum skewed space fADE. Third, the leading edge of plumes declines as a power-law in space: \(C(x_{\text{large}}) \sim x^{-1-a}\) (Fig. 5b), which is the same as the jump size pdf for the largest jumps. Finally, the early-time tail of the resident concentration based breakthrough curve increases linearly with time: \(C(t_{\text{small}}) \sim t\) (Fig. 5c).

The index \(a\) can be estimated by hydraulic conductivity statistics, as demonstrated by Benson et al. [13]. The estimation formula was given by Meerschaert and Schffler (see Eq. (2.1) in [95]).
Fig. 5. (a) The snapshot of tracers whose transport is governed by the space fADE (5) with maximally positive skewness \( q < 1 \). The space scale index \( \alpha \approx 1.1-1.9 \). The other parameters are \( V = 1, D = 1, T = 40 \), and the initial source is located at \( x_0 = 20 \). (b) The semi-log version of (a). (c) The concentration profile. The control point is 20 unit lengths downstream. In (a)–(c), the classical ADE (dots, with velocity \( V \) and dispersion coefficient \( D \)) is also shown for comparison.

\[
\frac{\partial C(x,t)}{\partial t} = -V \frac{\partial C(x,t)}{\partial x} + D \frac{\partial^2 C(x,t)}{\partial (x)^2},
\]

(7)

which physically means that the local variation of solute concentration is caused by a wide range of downstream zones. Simulated concentration snapshots and profiles, as shown in Fig. 6, look similar to those of the time fADE model (2) with \( 0 < q < 1 \) (Fig. 1). This similarity might be the motivation behind the application of the space nonlocal model to simulate the retention process (see for example, [12,45,67,143]).

Further analysis, however, indicates that the plumes described by the space fADE (7) are fundamentally different from the time fADE model (2). First, solute transport governed by the space fADE (7) has significant backward dispersion, due to the power-law distribution of large jumps in the backward direction (i.e., opposite to the main flow direction). Second, the concentration peak can move much faster than the second-order ADE, and the peak is not necessarily smaller than that of the second-order ADE. Third, the plume snapshot always skews negatively, while the skewness of the time fADE (2) varies with the capacity coefficient (which can change from negative to positive, as demonstrated by Fig. 2). Finally, the space nonlocal transport model does not discern the status of particles – all particles are mobile for all times.

The power-law distribution of jump sizes used by the space fADE model assumes that there exist arbitrarily long conductivities where particles can move with high velocities. In other words, the \( K \) structure has infinite correlation lengths consistent with, for example, fractional Brownian motion (see [10,71]). Note, however, that the probability of long jumps drops off according to a power-law so that very long jumps may be exceedingly rare and (just like Brownian motion) the probability of an infinite jump is zero. Successful field applications of the space fADE model (see for example [13]) imply that, in field experiments typically with a relatively short scale (in space and time), the solute transport may not reach the medium bounds, and thus the fADE is a useful model for this situation.
studies analyzed systematically the space/time nonlocality to
guide the selection of nonlocal models. When revisiting these tests,
we are especially interested in the hydrogeological conditions that
are believed to induce the anomalous dispersion. In each case, we
distinguish between the time and the space nonlocal models by fit-
ting the measured plumes (snapshots or concentration profiles)
with each model. We also compare the best-fit transport para-
ters (especially the velocity) used by the nonlocal models to mea-
sured field values. Results show that either a different or additional
nonlocality than considered by previous researchers is needed for
each of the four experiments.

3.1. Experiment 1: the hillslope tracer experiment near the River Dee

A tracer test was conducted in a saturated hillslope (with a
slope 4°–5°) near the River Dee, Scotland [143]. Adjacent catch-
ments contained well-connected fractures, and measurements of
the saturated hydraulic conductivity ($K_s$) show strong spatial vari-
ability [137] that is as high as four orders of magnitude (see [44]
and their Fig. 5). The variance of $\ln(K_s)$ can be larger than 5.3. A
pulse of NaCl was injected into a trench at the top of the plot, and
then flushed for 140 h using stream water. Water samples for Cl
were taken from two arrays of piezometers installed down-
slope, with one 26.3 m and the other 40 m downgradient from the
injection trench. Samples were collected every 4 h up to 64 h, and
then every 8 h for 80 h. The interconnected fractures may provide
preferential pathways for the NaCl tracer, while the soil matrix
may significantly retard some tracer particles (as discussed by
Zhang et al. [143], see their Fig. 7). We emphasize here that (1)
the sampling round at the beginning may be too “coarse” to catch
the possible early arrival of solutes, and (2) snapshot data (concen-
tration distribution in space at one time) were not collected in this
test (concentration-time data is much easier to collect in the field).
The snapshot actually can be very helpful in distinguishing the space-
and/or time-nonlocality underlying the fADE models.

The observed concentration breakthrough at the two locations
all contain apparent late tails (see the symbols shown in Fig. 7).
Zhang et al. [143] assumed that the positively skewed profiles (in
time) were caused by a negatively skewed plume in space, and
therefore they applied the space fADE (7) with negative skewness
to fit the concentration profiles. We tested their assumption, and
the corresponding space fADE simulation is shown in Fig. 7. The
calibrated parameters, including $\alpha = 1.4$, $V = 0.40$ m/h, and
$D = 2.8$ m$^2$/h, are similar to those calibrated by Zhang et al.
[143] (which are $\alpha = 1.51$, $V = 0.81$ m/h, and $D = 2.8$ m$^2$/h).

The application of a space nonlocal model for hillslope transport
dominated possibly by a retention process, however, may cause
two problems. First, although the space fADE (7) does describe a
long-tail concentration profile (as discussed above and see also
Fig. 6c), it requires an infinite upstream boundary where signifi-
cant amount of mass disperses backward. A space nonlocal model
assumes that the late tail of a concentration breakthrough is not
caused by trapping in relatively immobile zones (i.e., the soil ma-
trix), but caused by large backward dispersion. The long-distance
and long-term backward dispersion is physically unlikely. Second,
the space nonlocal model cannot describe the decline of solute
mass. Note that it was difficult to calculate the full mass recovery
at this test, due to the limited number of piezometers and the dif-
ficulty of accurately estimating the water flow rate, as explained by
Zhang et al. [143]. However, the simulated concentration profiles
using either the space fADE (7) or the classical ADE suggest that
the total measured Cl mass may decline with transport distance –
the best-fit results underestimate the peak concentration (and also
the full mass) at the upslope piezometers, while they overes-
timate the peak concentration (and also the full mass) at the down-
slope piezometers (Fig. 7).
We now model the late BTC tail with a time-nonlocal (i.e., retention) process. First, we assumed that the observed concentration is the resident concentration in the total phase, and the simulated results using the time-nonlocal model (2) are shown in Fig. 8. The effect of fractures on the plume leading edge remains unknown, due to the weak and possibly undetected early arrivals in concentration profiles.

The above transport solution to the hillslope concentration profiles, denoted as “Time fADE #1”, however, ignores the possible decline of mass because it is tracking total concentration. To address a scenario in which the measured concentration represents the resident concentration in the mobile phase (that declines with time), we developed an additional time fADE model (denoted as “Time fADE #2” and shown in Fig. 7) that can describe both the peak and late-time behavior of concentration profiles measured at each monitoring location. The calibrated time scale index, $\gamma = 0.28$, is very close to the $\gamma (0.30)$ calibrated by Schumer et al. [128] for fluorescent dye transport in a mountain stream. Note that the time

![Fig. 7. Comparison between the measured (diamonds) and the simulated (lines) Cl$^-$ concentrations at two downslope locations (denoted as $L_1$ and $L_2$) of the hillslope tracer experiment [143]. The nonlocal models are the space fADE with negative skewness (7) (representing the total resident concentration) and the time fADE (2) (representing the resident concentration in the mobile phase). Solutions from the second-order ADE are also listed (dashed line). (c) and (d) is the log–log version of (a) and (b), respectively. The best-fit parameters are shown in Table 1, where the time fADE model is denoted as “Time fADE #2”.

![Fig. 8. Comparison between the measured (diamonds) and the simulated (lines) Cl$^-$ concentrations at two downslope locations (denoted as $L_1$ and $L_2$) of the hillslope tracer experiment [143]. In the legend, “$RC_{tot}$” and “$RC_m$” denote the resident concentration in the total and mobile phase, respectively; “$FC$” denotes the flux concentration. The nonlocal model is the time fADE (2). (c) and (d) is the log–log version of (a) and (b), respectively. The best-fit parameters are shown in Table 1 (denoted as “Time fADE #1”).

a $L_1 = 26.3$ m (Time fADE #1)

b $L_2 = 40$ m (Time fADE #1)
FADE model (2) can describe the positively skewed concentration profile, with a positively skewed spatial plume snapshot.

The best-fit velocity for the time FADE (2), 1.15 m/h (Table 1), is the same order as the field measured and averaged velocity, \( \sim 0.80 \text{ m/h} \) [143]. For the case of the space FADE (7), the nonlocality is described by the fractional dispersion, and thus its velocity \( (V = 0.40 \text{ m/h}) \) is relatively smaller than either the time FADE model or the measured velocity.

### 3.2. Experiment 2: the dye test at Missouri River

The dye data measured on four reaches of the Missouri River between Sioux City, Iowa, and Plattsmouth, Nebraska [141], remain attractive to various researchers. Czernuszenko et al. [40] and Seo and Cheong [130] fit the measured concentration profiles to a dead-zone model (also called the storage zone model, see Eqs. (2) and (3) in [130]). This storage zone model is similar to the single-rate mass transfer model [30,134], and it is assumed to describe long tails that are often attributed to trapping effects of dead-zones. Deng et al. [45] revisited the test data recently and found that the single-rate mass transfer model tends to overestimate the decline of late tails of measured concentration breakthrough curves. They then applied the space FADE with negative skewness to model the measured concentrations.

A single transfer rate may not sufficiently describe the complex mass transfer processes attributed to the wide spectrum of dead-zones in natural rivers and streams. As pointed out by Deng et al. [45], dead-zones can have many forms, such as reverse flows induced by bends and pools, side pockets, zones between dikes, turbulent eddies, and wakes behind bed irregularities and roughness elements (ripples, sand-dunes, cobbles, boulders, etc.). A nonlocal model therefore should be used to describe the multi-scale storage-release zones. However, as discussed above, the space nonlocal model with negative skewness is a reasonable physical model only if there is significant backward dispersion (at a scale much larger than the forward/downgradient movement). In natural rivers and streams, whether any of the above dead-zones can cause such significant backward motions of solutes remains an open question. However, as an alternative, the multi-rate mobile/jomobile model developed by Haggerty et al. [65] models particles that are immobilized for extended periods of time and has a more sound physical basis than either upstream dispersion or the single-rate counterpart to describe mass transfer between active flow and dead-zones in natural systems (see for example, [61,66]).

We did not apply the MRMT model for this case directly. Indeed, the time FADE (2) with \( 0 < \gamma < 1 \) is a MRMT model with infinite mean power-law memories (see [128] and also the recent discussion by Benson and Meerschaert [14]). Here we first fit the Missouri River dye data using the time FADE (2). The simulated late-time tails of concentration profiles are too heavy to match the measured ones. We then explored the suitability of the time FADE (3) for this problem. Note the model (3) is a logical extension of (2). The simulated concentration profiles generally match the measured ones (Fig. 9b,d), although the match is not as good as the space FADE (7) (Fig. 9a and c) at very early or late periods.

### 3.3. Experiment 3: the natural gradient tracer test at Cape Cod

The well-known natural gradient tracer test at Cape Cod, Massachusetts, monitored the transport of several tracers, including...
the relatively nonreactive bromide (Br\(^-\)) ion, through a sand and gravel aquifer from 1985 to 1986 [58,88]. A 7.6 m\(^3\) slug of tracer solution was injected into the aquifer and then collected through 656 multilevel samplers for 16 cycles over 511 days, with the furthest observation well 280 m downgradient from the injection location. The unconfined aquifer is composed of glaciofluvial deposits, mainly stratified sand and gravel with minor fractions of silt and clay. Borehole flowmeter tests indicate that the variation of hydraulic conductivity is approximately one order of magnitude of silt and clay. Borehole flowmeter tests indicate that the variation of hydraulic conductivity is approximately one order of magnitude [72,88], with an estimated \(\ln K\) variance of only 0.26. The Cape Cod test is one of few field experiments conducted with sufficient detail, control, and time to fully examine the field-scale (i.e., hundreds of meters) dispersion process [88].

Measured bromide plumes show subtle anomalous behaviors. The 1D snapshots of the bromide plume's core were analyzed by Benson et al. [11] and shown by symbols in Fig. 10. Most parts of the spatial distributions of plumes in either 1D (Fig. 10a) or 3D (see Figs. 10–13 in [88]) look symmetric. Garabedian et al. [58] found further that the bromide longitudinal variance increased nonlinearly during the first 26 m of travel distance (corresponding to a travel time between 55 and 83 days) and then reached a linear trend (see also the symbols in Fig. 11b and c). The 1D snapshot measured on day 349 (Fig. 10), however, shows clearly a concave-upward shape of tails, one anomalous dispersion character that cannot be explained by the traditional, second-order ADE.

Benson et al. [11] applied the space fADE (4) with a symmetric skewness (i.e., \(q = 0.5\)) to fitting the 1D snapshots. We first adopted their model and approximated the space fADE (4) with \(\alpha = 1.8\) and \(q = 0.5\) using our Markovian random walks. Note here \(\alpha = 1.8\) is the value estimated by Benson et al. [11] based on the measured plume statistics (including the mean travel distance and the measured plume variance), and thus we did not modify it. The solution does describe the concave-upward shape of the plume (Fig. 10b), but it overestimates both tails of the snapshot. As also noted by Benson et al. [11], the degree of overestimation can be decreased by calibrating the scale index \(\alpha\) (i.e., by increasing

---

**Table 2**

Best-fit parameters used for the Missouri River dye test. Note “Ri-space” denotes the space fADE (7) with negative skewness used for reach \(i (i = 1–4)\), and “Ri-time” denotes the time fADE (3) with \(1 < \gamma < 2\).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distance (km)</th>
<th>(\alpha)</th>
<th>(\gamma)</th>
<th>(\rho) (m/s)</th>
<th>(V) (km/h)</th>
<th>(D) (km(^2)/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R_i)-space</td>
<td>65.65</td>
<td>1.8</td>
<td>N/A</td>
<td>N/A</td>
<td>4.92</td>
<td>1.3</td>
</tr>
<tr>
<td>(R_i)-time</td>
<td>65.65</td>
<td>2</td>
<td>1.7</td>
<td>0.03</td>
<td>4.92</td>
<td>1.3</td>
</tr>
<tr>
<td>(R_{i-1})-space</td>
<td>134.35</td>
<td>1.7</td>
<td>N/A</td>
<td>N/A</td>
<td>5.2</td>
<td>1.5</td>
</tr>
<tr>
<td>(R_{i-1})-time</td>
<td>134.35</td>
<td>2</td>
<td>1.6</td>
<td>0.024</td>
<td>5.1</td>
<td>1.8</td>
</tr>
<tr>
<td>(R_{i-2})-space</td>
<td>186.64</td>
<td>1.7</td>
<td>N/A</td>
<td>N/A</td>
<td>5.4</td>
<td>1.7</td>
</tr>
<tr>
<td>(R_{i-2})-time</td>
<td>186.64</td>
<td>2</td>
<td>1.6</td>
<td>0.022</td>
<td>5.4</td>
<td>2.0</td>
</tr>
<tr>
<td>(R_{i-3})-space</td>
<td>227.35</td>
<td>1.8</td>
<td>N/A</td>
<td>N/A</td>
<td>5.5</td>
<td>2.5</td>
</tr>
<tr>
<td>(R_{i-3})-time</td>
<td>227.35</td>
<td>2</td>
<td>1.6</td>
<td>0.020</td>
<td>5.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

---

**Fig. 10.** (a) The measured (symbols) versus the simulated (lines) Cape Cod Bromide snapshots using the space and time fADE (8). (b) The space fADE (4) with a symmetric fractional-order operator (\(q = 0.5\)) versus the space and time fADE (8) with a maximum skewed fractional-order operator. The classical ADEs (dashed lines) are listed for comparison. \(C_{\text{tot}}, C_m,\) and \(C_{\text{im}}\) in (b) denote the concentration in the total, mobile, and immobile phases, respectively.

**Fig. 11.** Measured (symbols) versus simulated Cape Cod Bromide spatial moments (zeroth to the second orders) using the space and time fADE (8). (a) Mass; (b) mean travel distance versus the measured plume variance; (c) is the log–log plot of (b). Solutions of the second-order ADE are added for comparison in (b) and (c).
However, as discussed above, the space nonlocality described by a fractional dispersion with skewness $q < 1$ contains backward dispersion stronger than the Fickian dispersion. A strong upstream dispersion, however, may not be hydrologically possible in ambient conditions in natural aquifers, and thus an alternative is needed to explain the heavy tails.

The stratified architecture of the Cape Cod aquifer, where particle transport through continuous high-velocity zones could possibly be a space-nonlocal transport process [101] and slow particle transport in low-velocity zones could possible be a weak time-nonlocal process, motivated us to introduce both space and time-nonlocalities into a single transport model. In particular, the space fADE (5) with the maximally positive skewness is a physically more reliable alternative to Benson et al.’s model [11] for describing the leading edge of plumes in space. To simulate the trailing edge of plumes, we also considered a weak time nonlocality due to the relatively low-velocity zones that the solute may encounter. The resultant fADE model is a combination of space and time nonlocalities (see Appendix A.2 and A.3 for its development):

$$\frac{\partial C(x,t)}{\partial t} + \beta \frac{\partial^\gamma C(x,t)}{\partial x^\gamma} = -V \frac{\partial C(x,t)}{\partial x} + D \frac{\partial^3 C(x,t)}{\partial x^3}.$$
The observed noise shown in Fig. 11 may be simply due to the ADE model, the space and time fADE model can more closely describe measurement errors in bromide concentrations, and/or the temporal sampling rounds [58], and a similar phenomenon was observed again by Hess et al. [73] in a later test at the same site (with a larger area). The observed noise shown in Fig. 11 may be simply due to measurement errors in bromide concentrations, and/or the interpolation/extrapolation of the measured data [58]. Compared to the ADE model, the space and time fADE model (8) can more closely match the first and second moments of plumes (Fig. 11b and c) at early times, although the linear growth of plume variance at later times tends to follow Fick’s law and favors the traditional local SA-ADE [58].

An effective transport model was also developed by Dentz et al. [50] for stratified random media. The projected superdiffusive transport of solutes through a two-dimensional stratified media, where the perturbation of vertical velocity is a Gaussian white noise, can be described efficiently using a correlated random walk characterized by its Lagrangian velocity correlation. Le Borgne et al. [91] investigated the stationary velocity field caused by lognormally distributed conductivity field and defined an effective Lagrangian statistical model to describe the solute transport. In these works, the correlated CTRW was found to perform better than the classical CTRW without correlation (see for example, Fig. 2 in [91]). In a future study, we will extend the iid Lévy motion to the coupled one, and investigate numerically the influence of a fractional Brownian motion field on solute transport.

To draw a conclusion, although the classical ADE model can describe the symmetric plume peak, spatiotemporal nonlocality is necessary to describe plume tailing behaviors and nonlinear growth rates of the first and second moments at the Cope Cod test. Anomalous behavior is not uncommon for solute transport in relatively homogeneous media. For example, at the sandy aquifer in Borden, Ontario, which contains also glaciofluvial sediments with a small variation of K (the variance of lnK is as small as 0.29), the measured concentration profiles for a nonreactive tracer (iodide) have both an early leading edge and a slow-declining late tail (as indicated by Fig. 6a in [94] and Fig. 4 in [139]). The applicability of the spatiotemporal fADE (8), however, is not limited to homogeneous aquifers, but can be extended to highly heterogeneous ones, as demonstrated by the next example.

### 3.4. Experiment 4: the tritium transport at the MADE-1 test

Natural aquifers, such as those created by alluvial depositional systems, can have high heterogeneity characterized by a high variation of K [57]. The best-known example may be the MADE site. Natural gradient tracer tests conducted at the Columbus Air Force Base in northeastern Mississippi, commonly known as the Macrodispersion Experiment (MADE) test site, produced anomalous transport data that have been of continuous interest [2,22,23,123] – probably due to the strong influence of high subsurface heterogeneity on solute transport (the variance of lnK is 4.5). The 11-m-thick, unconfined aquifer is composed of alluvial deposits, containing poorly sorted to well-sorted sandy gravel and gravelly sand with minor amounts of silt and clay. Adjacent exposures indicate that soil facies occur as irregular lenses and layers having horizontal dimensions ranging up to 8 m [123]. A mixed zone of fine- and coarse-grained sediments is located at the 0–2 m depth interval, while much less fines are present in deeper intervals (the volume is as low as 7%) [22].

Three dominant characteristics of anomalous transport are present in the tracer test data including a near-source peak, a heavy leading plume edge (which can be seen by symbols shown in Fig. 12 for the longitudinal plume transects), and large apparent loss of mobile mass [2,68,128]. In response to the positively skewed tritium plumes which cannot be described by the classical ADE with a coarse-scale flow field [2], researchers have attempted to simulate the tracer plumes using many different methods in the last decade (see the review by Molz et al. [107] and Zang et al. [148]). The commonly accepted conclusion is that the classical ADE with a fine-scale (such as decimeter-scale) flow field [51,93,125,152,153,155], or the dual-domain approach with a similarly fine resolution of flow field [53,69,75,107], may describe the anomalous plumes at the MADE site. However, the practicality of the first method is questionable due to limitations of current characterization techniques in differentiating between small-scale heterogeneities. The tradeoff between the dual-domain approach and the fine-scale flow field approach remains an open research question. A time-nonlocal model embedded in a CTRW framework, which is functionally the same as the time fADE model (1), was suggested by Berkowitz and Scher [15,16] as an alternative to simulate the MADE plumes. On the contrary, Benson et al. [13]...
Zhang et al. [148] found that the space nonlocality underlying the space fADE (5) dominates the transport of tritium. The solute transport properties of realistic, large-scale alluvial sediments have been analyzed recently. Fogg et al. [55,56] found that the ancient high-permeable channel tends to interconnect spatially and form regional-scale preferential paths for solutes in typical alluvial aquifers. Sufficient evidence also indicates that aquitards surrounding the contaminant source play a key role in the retention of solute mass (see for example, LaBolle and Fogg [87]). Their findings suggest that spatiotemporal nonlocality can dominate solute transport in typical alluvial depositional systems. Particularly, the presence of preferential flowpaths may be a possible cause for the fast leading edge of MADE plumes (as also concluded by Salamon et al. [125]) and can be characterized by a space nonlocal process, while trapping in aquitards may explain the near-source peak and can be modeled as a time nonlocal process. The existence of both space- and time-nonlocalities at the MADE alluvial aquifers was also verified analytically by Schumer et al. [128] using the scaling properties of fractional derivatives.

We tested all of the above assumptions by fitting the MADE-1 tritium 1D snapshots using the time fADE (2) with $0 < \gamma < 1$, the time fADE (3) with $1 < \gamma < 2$, the space fADE (5) with the maximum skewness, and the space and time fADE (8). The space fADE (7) with less than maximally positive skewness was not selected, because (1) it models strong backward dispersion, and (2) it cannot describe a positively skewed snapshot. Best-fit parameters for each of the four models are listed in Table 4. The simulated plumes are shown in Fig. 12.

The four fADE models simulate different plume behaviors to match the observed, normalized snapshots (the total mass for each snapshot is normalized to be one). In a double-linear plot, all models seem to fit well the skewed tritium snapshots (see the left column in Fig. 12) although the time fADE (3) with $1 < \gamma < 2$ slightly overestimates the peak concentration. The semi-log (shown by the middle column in Fig. 12) or double-log plot (the right column in Fig. 12) reveals the discrepancy between the simulated leading edges. The time fADE (2) with $0 < \gamma < 1$ cannot explain the heavy leading edge after 27 days because the motion in (2) is limited to Fickian growth. The space fADE (5) has similar leading edges as the space and time fADE (8), implying that the space nonlocality dominates the power-law leading edges in both models (note both models use the same space index $x = 1.1$). The time fADE (3) with $1 < \gamma < 2$ describes a heavier leading edge than the time fADE (2) with $0 < \gamma < 1$, but this leading edge still declines faster than the observed power-law rate at large distances.

We then compared the measured [2] and modeled mass fraction for the mobile phase (Fig. 13). As expected, the space fADE (5) cannot explain the decline of mass. The time fADE (3) with $1 < \gamma < 2$ does describe the measured mass fraction, but the initial mobile fraction reaches an extreme value (138, Table 4), that is much higher than previous estimations using either the single-rate mass transfer model (4.7, see [69]) or the fractal model (5.0, see [128]). The high initial mobile fraction of the time fADE (3) with $1 < \gamma < 2$ is due to the best-fit, large capacity coefficient ($\beta = 5$) – in order to describe the fast leading edge using any time fractional model, the capacity coefficient must be large enough – and therefore the mass in the mobile phase declines very fast. By process of elimination, this leaves the space and time fADE (8) as the best large-scale fractional transport model for MADE-1 tritium. Note the best-fit velocity for the space and time fADE (8) is in the range of field measured velocities (0.12–0.36 m/d), while the best-fit $V$ for the time fADE (3) with $1 < \gamma < 2$ (0.01 m/d) is one order of magnitude lower.

Fig. 14 shows the simulated snapshots (in real concentration units) for solute in each phase by using the space and time fADE (8). Plumes in different phases have similar (positively skewed)

Table 4

<table>
<thead>
<tr>
<th>Model</th>
<th>$\alpha$</th>
<th>$\gamma$</th>
<th>$\beta$ (d$^{-1}$)</th>
<th>$V$ (m/d)</th>
<th>$D$ (m$^2$/d)</th>
<th>IMF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space fADE (5)</td>
<td>1.1</td>
<td>N/A</td>
<td>N/A</td>
<td>0.14</td>
<td>0.14</td>
<td>1.0</td>
</tr>
<tr>
<td>Space and time fADE (8)</td>
<td>1.1</td>
<td>0.35</td>
<td>0.05</td>
<td>0.22</td>
<td>0.22</td>
<td>2.4</td>
</tr>
<tr>
<td>Time ADE (2)</td>
<td>2</td>
<td>0.50</td>
<td>1.0</td>
<td>0.50</td>
<td>0.50</td>
<td>5.0</td>
</tr>
<tr>
<td>Time ADE (3)</td>
<td>2</td>
<td>1.25</td>
<td>5.0</td>
<td>0.01</td>
<td>0.01</td>
<td>138.0</td>
</tr>
</tbody>
</table>

Fig. 14. Comparison between the measured (circles) and the simulated (lines) MADE-1 tritium snapshots (actual concentrations, without normalization). Concentrations in all phases (i.e., total, mobile, and immobile) are simulated using the space and time fADE (8).
shapes. Note the mobile mass is originally higher than the immobile mass for the first snapshot (since the mass was initially injected into the mobile domain), and then declines with transport distance until the mobile mass is much less than the immobile mass for the last sampling cycle. This suggests an increased retention of solute particles in aquitards during the observation period and is consistent with the analysis of Harvey and Gorelick [69].

4. Discussion and conclusions

4.1. Distinction between space and time-nonlocality

The space and time fADE models are nonlocal in character and they account for different transport process. The concentration change at some location and time might depend on a wide variety of locations in space (i.e., the space nonlocality), and also might depend on the temporal history of concentration “loading” at that location (the time nonlocality). The space fADE models (5) and (7) describe a space nonlocal transport process, while the time fADEs (2) and (3) describe a time-nonlocal transport process. The space and time nonlocality is also shown by the space and time integrals shown in (A.6).

The space and time fADEs differ in three main ways. First, while both the space and the time-nonlocal forms model non-Fickian transport, large ranges of solute displacement can be modeled by a space nonlocal, fractional-derivative model, and long waiting times can be modeled efficiently by a time-nonlocal, fractional model. Second, the time nonlocality can explain mass decline because it describes the dynamic partitioning of solute mass into the immobile phase, while the space nonlocality cannot distinguish the status of solutes. Third, the spatial moments of a solute plume governed by the time fADE are finite, while the space fADEs produce finite sample moments. Third, the spatial moments of a solute plume governed by the time fADE are finite, while the space fADEs produce finite sample moments. This suggests an increased retention of solute particles in aquitards during the observation period and is consistent with the analysis of Harvey and Gorelick [69].

4.2. Extension and applicability of the time fADE model

The time fADE with scale index $0 < \gamma < 1$ describes a power-law late tail of the (flux concentration based) breakthrough, with slope $-1$ to $-2$ in a log–log plot. The time fADE with scale index $1 < \gamma < 2$ extends the former by describing a power-law late tail with slope $-2$ to $-3$, and a relatively heavier early tail due to successive jumps corresponding to negative waiting times. So the time fADE with $1 < \gamma < 2$ can describe simultaneously the fast moving front through relatively high flow regions and the trapping of solutes in relatively immobile domains.

All four experiments reviewed in this study show the applicability of the time fADE model, implying that the natural flow system is typically heterogeneous with intrinsic immobile domains. In coupled surface/subsurface systems such as hillslope or river flow, abundant dead-zones can retard the transport of tracers significantly and induce a time-nonlocal process. In subsurface systems such as the stratified aquifer or aquifer/aquitard mixture (i.e., the alluvial environment), solutes can have both large motions and long trapping periods, and transport can be explained by space and time-nonlocalities.

The applicability of the time fADE (3) with $1 < \gamma < 2$ on transport through natural rivers needs further investigation. Schumer et al. [128] found that the tracer transport through a second-order mountain stream (where the relatively complete late-time tail of breakthrough was recorded) follows the time fADE (2) with $\gamma = 0.3$. In this study, the observed concentration profiles of the dye test at Missouri River only contain part of the late-time tails, due to the limited experimental time at each reach (Fig. 9). We used the model (3) with $\gamma = 1.6$ because of the relatively steep late-time tail compared to the one in Schumer et al.’s study [128]. However, it is likely that the observed late-time tracer tail at Missouri River did not reach “late-time” yet, and thus the best-fit parameter $\gamma = 1.6$ could be too large. In addition, in either case, there is no power-law leading edge or space nonlocal transport process. It might imply that the large flow rates within the natural rivers/streams do not deviate enough to form a power-law distribution (or in other words, fast moving solutes have similar velocities), so that the distribution of large particle jumps has a thin tail. Further analysis for the above hypothesis is needed.

4.3. Predictability of fADE models

The fADE model can be developed by using the CTRW scheme (see also the Appendix), where the moments of solute particle jump sizes and waiting times may be finite or infinite. The quantitative link between the moments of random walking particles (which is difficult, if not impossible, to measure at present) and the properties of natural media (which might be quantified), however, remains unknown. At present, the fADE model largely remains a fitting tool, although Benson et al. [13] showed that the space index $x$ can be estimated by the hydraulic conductivity statistics, and Zhang et al. [149] found that $\gamma$ and $\beta$ may be estimated by looking at the distribution of fine-grained facies. The weak...
predictability of the model and its main parameters is the main limitation of the fADE model in applications.

4.3.1. Numerical investigation of model predictability by previous studies

A direct link between non-Fickian transport theories and the underlying natural media structure is an area of ongoing research. Such a link may not be constructed reliably, if the medium heterogeneity cannot be measured with certainty at all relevant scales. Numerical experiments have therefore been used to explore the predictability of fractional-derivative model parameters given the knowledge of subsurface heterogeneity. For example, Herrick et al. [71] and Kohlbecker et al. [82] found that the power-law distributed hydraulic conductivity gives rise to heavy tailed velocity fields, and Grabasnakj [62] further investigated the resultant solute particle displacements. These studies concluded that the heavy tailed and long-range dependent $K$ field is responsible for the heavy tailed solute displacement, supporting the application of the space fADE model. Monte Carlo numerical simulations conducted by Zhang et al. [149] further revealed that a broad distribution of the thickness of low-permeable deposits guarantees the heavy late tail of solute breakthrough, implying the applicability of the time fADE model.

Recent work by Le Borgne, Dentz, and their colleagues extend the classical CTRW to correlated CTRW. For example, Le Borgne et al. [89] investigated the velocity field correlation structure for multinormal log permeability fields. The simulated Lagrangian velocity distribution (i.e., the velocities sampled along a streamline at equal space intervals) led to a correlated CTRW model (see Le Borgne et al. [91]). Similar transport behavior was found for random walkers through an exponentially correlated hydraulic conductivity field (see Fig. 4 in [90]) and stratified flow field (see Dentz et al. [50]), where the Lagrangian velocity shows correlation to (or “memory of”) its initial velocity in a certain time range (see their Fig. 4). Only after a long time will the velocity “memory” disappear. Therefore, the non-Fickian models based on the CTRW with iid jumps and waiting times, including the classical CTRW [18] and the fractional-derivative models investigated in this study, may need to be refined by considering correlation of jumps, such as the correlated CTRW developed by Meerschaert et al. [100] and Jurlewicz et al. [76]. This topic is beyond the scope of this study, and we will leave it for future research.

4.3.2. Relationship between model parameters and medium heterogeneity revealed by this study

One of the main motivations of this study is to reveal, and to quantify if possible, the variation of main nonlocal parameters in various natural systems. In this study, the nonlocal parameters (including the space scale index $\alpha$, the time scale index $\gamma$, and the fractional capacity coefficient $\beta$) used by the fADE model are defined by fitting the available plume data (breakthrough curve or snapshot). Results of a literature review shows that $\gamma$ decreases with the increase of heterogeneity represented quantitatively by the variability of hydraulic conductivity $K$ (Fig. 15). The opposite is true for $\beta$. The space index $\alpha$, however, may not always decrease with the increase of variance of $\ln(K)$, since the space nonlocal process may also depend on the structure/correlation of $K$ (which is consistent with the assumptions in [10,13] and further numerical verification in [71]). In addition, the transport parameter, including the velocity, however, may be measured in the field. All four experiments revisited in the last section show that the best-fit velocity in the fADE models are in the range of field measurements, implying that field measured velocity can be a reliable starting point for calibrating fADE models.

Further investigations, including theoretical, numerical, laboratory, and field experiments, are needed in the future to establish the important link between the nonlocal parameters and the heterogeneity property. Some preliminary tests have been done [149], and we will analyze them in the next study.

Analysis, comparison and application of various fADEs in this study are intended to provide a general guidance for model selection. Natural geological deposits with highly contrasting permeability may form mobile and relatively immobile zones, where the potential mass exchange between mobile and immobile zones results in a wide time distribution for solute “trapping”. The transport process, combined with the distinct particle status, can be characterized efficiently by the time-nonlocal model, including the time fADE, the MRMT model, and the CTRW framework with mobile/immobile distinction and appropriate memory function. If the high-permeable material tends to form preferential flow paths, such as the interconnected paleo-channels observed in alluvial depositional systems, then the solute transport may show a heavy leading edge, which can be described by the space fADE with maximally positive skewness. In real-world subsurface media such as the MADE and the Cape Cod aquifer systems, the superdiffusive process (i.e., particle jumps deviating from the Fick’s law) can appear simultaneously with subdiffusion (i.e., solute retention due to sorption or trapping in immobile domains), where a mixed space and time-nonlocal model (such as the spatiotemporal fADE) can be used.

Acknowledgements

This material is based upon work supported by the National Science Foundation (NSF) under Grant Nos. DMS-0539176, EAR-0748953, and EAR-0749035, and the US Department of Energy (DOE) Basic Energy Sciences grant DE-FG02-07ER15841. Y. Zhang was also partially supported by the faculty startup package at Desert Research Institute (DRI). Any opinions, findings, conclusions or recommendations do not necessary reflect the views of the NSF, DOE, or DRI. We thank two anonymous reviewers for their insightful suggestions which improved this work significantly.

Appendix A. A macroscopic and microscopic dynamics underlying the fractional-derivative models

To reveal further the representative nonlocal property underlying the fractional-derivative models, we derive the space and time fADE by considering the spatiotemporal averaged, macroscopic dynamics. For comparison, we then explore the microscopic dynamics underlying the fractional dispersion in both space and time. For simplicity, in the following we only show the case of time scale index $0 < \gamma < 1$. For the case of $1 < \gamma < 2$, the derivation...
A.1. Comparison of the fADE, the generalized Master equation and the SA-ADE

We first show the connection between the spatiotemporal fADE and the generalized Master equation (GME), and then compare the resultant model to the space and time-nonlocal kernels defined in the SA-ADE theory. The link between fADE and GME has been discussed in many studies (see [103] and the numerous references cited therein), so the following derivation of fADE using the GME is a short review of published work.

The well-known generalized Master equation given by Klafter and Silbey [78] (see also Eq. (48) in [132]) is of the form

\[
\frac{\partial \psi(x,t)}{\partial t} = \int_{-\infty}^{\infty} \Phi(y, t) \psi(y,t)d\tau dy - \int_{-\infty}^{\infty} \Phi(x, t - \tau) \psi(y, t)d\tau dy,
\]

where the kernel \( \Phi \) is space and time-nonlocal, and also spatiotemporally stationary. The kernel can be defined as (in Laplace space)

\[
\Phi(s, x) = \tilde{\psi}(s, x) \frac{1}{1 - \tilde{\phi}(s)},
\]

where the tilde “~” denotes the Laplace transform, \( \tilde{\psi}(s, x) \) is the Laplace transform of the CTRW joint density of jump length and duration, and \( \tilde{\phi}(s) \) is the Laplace transform of the duration density \( \phi(t) = \int_{0}^{\infty} \psi(x, t)dt \). Take the Laplace transform (\( t \to s \)) and Fourier transform (\( x \to k \) of (A.1)), and use the definition of (A.2) to derive the Montroll–Weiss Master equation [108]

\[
\tilde{\psi}(s, k) = \frac{1 - \tilde{\phi}(s)}{s} \tilde{\phi}(s) = \frac{1 - \tilde{\phi}(s)}{s} \tilde{\phi}(s) \tilde{f}(k),
\]

where the right-hand term is for independent jump size and waiting time, and \( \tilde{f}(k) \) denotes the jump density.

Empirical density functions can be defined in (A.3) when fitting real data, as shown by Berkowitz et al. [18]. Similarly, the fADE can be developed by assuming case-specific density functions in (A.3). In particular, we assume that (1) the long-tailed jump length density \( f(x) \) shows the asymptotic behavior \( f(x) \sim Ax^{-\gamma}x^{-1-\gamma} \) (where \( x > \sigma > 0.1 < a < 2 \), and \( \sigma \) is a dimensional constant) ([103], Eq. (56); [17], Eq. (37)), and (2) the long-tailed duration density shows the asymptotic behavior \( \phi(t) \approx Bt^{1+1/1-\gamma} \) (where \( t > \tau > 0.0 < \gamma < 1 \), and \( \tau \) is a dimensional constant [17, Eq. (36)]). Substituting \( f(k) \) (Fourier space asymptotic) and \( \phi(s) \) (Laplace space asymptotic) into (A.3), taking into account the expansions for small \( k \) and the usual long time limit, and taking the inverse Fourier–Laplace transform, one gets the spatiotemporal fADE

\[
\frac{\partial^2 \psi(x,t)}{\partial t^2} - \frac{t^{-\gamma}}{\Gamma(1-\gamma)} p_0(x) = D \frac{\partial^2 \psi(x,t)}{\partial x^2},
\]

which is the same as the fADE developed by Meerschaert and Scheffler [2004, Eq. (5.4)]. Here \( D = \sigma^2/\tau \) is held constant. The drift displacement of particles can be added into (A.4) by assuming a shift jump size distribution (i.e., \( f'(x - Vt) = f(x) \)), or using the Galilei invariant assumption (see [103, Eqs. (79)–(86)]):

\[
\frac{\partial^2 \psi(x,t)}{\partial t^2} - \frac{t^{-\gamma}}{\Gamma(1-\gamma)} p_0(x) = -V \frac{\partial \psi(x,t)}{\partial x} + D \frac{\partial^2 \psi(x,t)}{\partial x^2}.
\]

To compare (A.4) to the SA-ADE model, we adopt the CTRW memory function \( M(t) \) defined by Dentz and Berkowitz [48] (see their Eq. (6)), which has a Laplace transform

\[
\tilde{M}(s) = \frac{t_1s^\phi(s)}{1 - \phi(s)},
\]

where \( t_1 \) denotes the mean transition time. Then we take the inverse Fourier–Laplace transform of (A.3) to obtain:

\[
\frac{\partial p(x,t)}{\partial t} = \int_0^t M(t) \left[ \frac{\partial^2 p(x, t - \tau)}{\partial x^2} \right] d\tau
\]

\[
= \frac{\partial}{\partial x} \int_0^t \frac{M(t)DH(y)}{T(2 - 2y^{\alpha+1})} \frac{\partial p(x, y, t - \tau)}{\partial x} dy d\tau,
\]

where \( H(y) \) is the Heaviside function on \( (0, \infty) \). It is a special case of the one-dimensional simplification of the nonlocal dispersive constitutive theory (e.g., one SA-ADE theory) proposed by Cushman et al. ([37], Eq. (56)) (note here the advection term is not included for simplicity)

\[
\frac{\partial G(x,t)}{\partial t} = -\frac{\partial}{\partial x} \int_0^t B_1(y, t, \tau) \tilde{G}(x, y, t - \tau) dy d\tau + \frac{\partial}{\partial x} \int_0^t B_2(y, t, \tau) \tilde{G}(x, y, t - \tau) dt dy,
\]

if the SA-ADE memory kernels are defined as

\[
B_1(y, y, \tau) = \frac{M(t)DH(y)}{T(2 - 2y^{\alpha+1})}, \quad B_2(y, t, \tau) = \frac{\partial G(x, y, t - \tau)}{\partial x}.
\]

The above discussion shows that, from the Eulerian point of view, the local variation (at time \( t \) and location \( x \)) of solute concentrations is due to a weighted average of dispersive flux at all upgradient zones (including the local position \( x \)) and previous times (including the current time \( t \)). The weight, defined by the SA-ADE memory kernel \( B_0 \), declines as a power-law in space and time.

It is noteworthy that the specific FADE model (A.4) is a scaling limit (at a long time and a large distance) of the CTRW with power-law distributed jumps and durations. At early times, the solution of the fADE model differs sufficiently from the corresponding CTRW framework (see also the discussion by Barkai et al. [7]), and strictly speaking, the fADE is not a “subset” of CTRW. In addition, the fADE model (A.4) derived by the GME is a specific SA-ADE model with a stationary time and space nonlocal memory kernel. Only when the GME kernel \( \Phi \) is extended to describe also the local position and time properties (i.e., \( \Phi \) depends on not only the interval \( x - x' \), but also \( x \) itself) (see for example, [102]), will the fADE derived by the GME approach be extended to describe solute transport through nonstationary media.

The fADE model (8) can be derived by taking the form

\[
\tilde{f}(k) = e^{-\omega/\text{M}(\text{ik})^\alpha} \text{and } \tilde{\phi}(s) = e^{-\omega/\text{M}(\text{ik})^\alpha}
\]

in the Montroll–Weiss Master equation (A.3):

\[
\tilde{p}(k, s) = \frac{1 - e^{-\omega/\text{M}(\text{ik})^\alpha}}{s[1 - e^{-\omega/\text{M}(\text{ik})^\alpha} - \text{M}(\text{ik})^\beta]}.
\]

It has the scaling limit (in the long time and long-range limit, or correspondingly, \( s, k \to 0 \))

\[
\tilde{p}(k, s) \approx \frac{1 + \beta s^{-1}}{s \beta s^{-1} + \text{M}(\text{ik})^\beta}
\]

whose Laplace and Fourier inverse transform is the model (8).

A.2. The scaling limit of heavy tailed CTRW and the subordinated process

Here we derive the governing equation for the scaling limit of heavy tailed CTRW, by combining the subordinated scheme developed by Benson and Meerschaert [14] and the Lévy representation scheme developed by Baumer and Meerschaert [4]. We consider a
with independent, identically-distributed (iid) jumps $Y_1, Y_2, \ldots, Y_n$, and iid inter-jump duration times $t_1, t_2, \ldots, t_n$. The duration time $t$ between two jump events, which is also called the "clock/total time", contains two parts – operational time $u$ and waiting time $w$, so that $t = u + w$. The random process $U(t)$, usually called the "hitting time process", denotes the operational time $u$ that a particle spends in motion. The random process for the total clock time $t$ is denoted as $T(u)$, which is an infinitely divisible process. The process $U(t)$ is the inverse of the total clock time process $T(u)$, and vice versa. The density of $T(u)$, denoted as $l(t, u)$, represents the density of hitting time process for any given operational time $u$. The density of hitting time process, denoted as $h(u, t)$ (representing the density of operational time $u$ at a total time $t$), is therefore given by [14]

$$h(u, t) = \frac{d}{dt} P(U(t) \leq u) = \frac{d}{dt} P(T(u) \geq t) \quad \text{for} \quad u \leq t,$$

(A.11)

The clock time density $l(t, u)$ has Laplace transform $\hat{l}(s, u) = e^{-ut}$ (so that $\hat{s} \hat{l}(s, u) = -L(s)\.\hat{l}(s, u)$). The Laplace transform of (A.11) gives [14]

$$\hat{h}(s, u) = -d \left(1/s \hat{l}(s, u)\right) - \frac{1}{s} L(s) e^{-ut},$$

(A.12)

so we have

$$\frac{d}{du} \hat{h}(u, s) = -L(s) \hat{h}(u, s).$$

(Benson and Meerschaert [14] found that $L(s)$ is related to the probability of waiting time $w$ greater than $t$ (in Laplace space) via the relationship

$$L(s) = s + s\mu \hat{P}(w > t),$$

(A.14)

Here we assume that the particle waiting time probability decreases as a power-law of the clock time $t$

$$P(w > t) = C \frac{t^{-\gamma}}{\Gamma(1 - \gamma)}.$$  

(A.15)

Leading (A.15), whose Laplace transform is $s^{\gamma-1}/\mu$ and (A.14) into (A.13), and then taking the Laplace inverse transform, we obtain the governing equation for the hitting time process $h(u, t)$

$$\frac{d}{du} h(u, t) = -L(u) h(u, t) - \frac{\partial}{\partial t} \hat{h}(u, t),$$

(A.16)

with initial condition $h(u, 0, t) = \delta(t) + \frac{\partial}{\partial t} \delta(u, t)$.

The CTRW $X_t = S_t$ has the scaling limit process $A(U(t))$, which is a subordination that replaces the time parameter $t$ in the scaling limit jump process $A(t)$ by the hitting time process $U(t)$ [98]. The density $p(x, t)$ of $A(U(t))$ can be calculated via the integration based on the time subordination formula (see Eq. (7.13) in [54]; see also [99])

$$p(x, t) = \int_0^\infty w(x, u) h(u, t) du,$$

(A.17)

where $w$ represents the transition probability in operational time $u$.

In the following we explore the governing equation for $w$.

Baeumer and Meerschaert [4] derived the governing equation of the scaling limit of a CTRW with heavy tailed jumps. Here we adopt their methodology. Suppose that the iid jumps $Y_1$ and its probability distribution are infinitely divisible [96]. The Lévy representation states that a probability measure is infinitely divisible if and only if we can write its characteristic function in the form of $e^{\psi(k)}$ where

$$\psi(k) = ik + \int_{|y|>0} \left(e^{iky} - 1 - \frac{iky}{1 + |y|}\right) \phi(dy),$$

(A.18)

where $a$ is a real number and $\phi$ is a Lévy measure. The Lévy measure specifies the probability of particle jumps of a given size [129], which can be defined as the following forms if the probability follows a power-law [4]:

$$\phi(r, \infty) = \frac{qD^\nu(x-1)}{\Gamma(2 - \alpha)} r^{-\alpha}, \quad \phi(-\infty, -r) = \frac{(1 - q)D^\nu(x-1)}{\Gamma(2 - \alpha)} r^{-\alpha},$$

(A.19)

where $r > 0, D^\nu > 0$, and $1 < \alpha < 2$. Leading (A.19) into (A.18) and after a little algebra, one gets [4]

$$\varphi(k) = -ikV' + qD^\nu(k)^2 + (1 - q)D^\nu(-ik)^2,$$

(A.20)

where $V' = -\frac{\mu^{(1-2\nu)(\nu-1)}}{\Gamma(1-\nu)} \left(\frac{\nu}{\nu-1}\right) \gamma 2^{\nu-1} dy$.

Defining a generator in the operational time domain $u$, we have the following governing equation for the transition probability:

$$\frac{\partial \varphi(k, u)}{\partial u} = \varphi(k) \hat{w}(k, u),$$

(A.21)

Replacing $\varphi(k)$ in (A.21) with (A.20), and then taking the Fourier inverse transform, we get

$$\frac{\partial \varphi(w, u)}{\partial u} = -V \frac{\partial \varphi(w, u)}{\partial x} + qD^\nu \frac{\partial^2 \varphi(w, u)}{\partial x^2} + (1 - q)D^\nu \frac{\partial^2 \varphi(w, u)}{\partial (-x)^2},$$

(A.22)

Combining (A.22), (A.17), and (A.16), we obtain the spatiotemporal FDE model

$$\frac{\partial p(x, t)}{\partial t} + \beta \frac{\partial^2 p(x, t)}{\partial x^2} = -V \frac{\partial p(x, t)}{\partial x} + qD^\nu \frac{\partial^2 p(x, t)}{\partial x^2} + (1 - q)D^\nu \frac{\partial^2 p(x, t)}{\partial (-x)^2},$$

(A.23)

So the FDE model (A.23) requires a CTRW with power-law distributed jump size and waiting time.

A.3. Langevin approach and time subordination

The Langevin picture can shed some light on the foundation on microscopic dynamics underlying the fractional dispersion [103]. We propose the following Langevin equation by extending the nonlinear Langevin approach proposed by Zhang et al. [145]

$$dx(t) = A(x) du + q^{1/2} B_t(x) du + (1 - q)^{1/2} B_t(x) du,$$

(A.24)

where $dx(u)$ is a differential distance of travel, $A$ is a drift term, $X = X(u)$ is the current particle location, $du$ is a differential unit of time, $B_t(x)$ is a term defining the strength of diffusion, and $dL_t$ and $dL^*_t$ are independent maximum-positively skewed and maximum-negatively skewed $\alpha$-stable noise rescaled by $(du)^{1/\alpha}$, respectively. Following the argument in Ethier and Kurtz [52, p. 379], Stroock [136], and Zhang et al. [145], we obtain the backward equation corresponding to the specific Markov process (A.24)

$$\frac{\partial \varphi}{\partial u} = A(x) \frac{\partial \varphi}{\partial x} + D q \varphi(x) + D[1 - q] \varphi(x) \frac{\partial^2 \varphi}{\partial (-x)^2},$$

(A.25)

where $D = -1/\cos(\pi x/2)$.

Using the fractional-adjoint approach proposed by Zhang et al. [145], we obtain the forward equation corresponding to (A.25)

$$\frac{\partial \varphi}{\partial u} = \frac{\partial}{\partial x}[A(x)g(x, u)] + \frac{\partial}{\partial x}[D^*_t \varphi(x)g(x, u)],$$

(A.26)

where $D^*_t \varphi(x) = D q \varphi(x)$ and $D^*_t \varphi(x) = D[1 - q] \varphi(x)$. 


Similarly, we define the Time-Langevin equation

$$dT = e du + \beta^{1/2} \frac{1}{C_21} dW(u),$$  \hspace{1cm} (A.27)

where $e = 1$ or $0$, and $dW(u)$ denotes the maximum-positively skewed stable noise rescaled by $(du)^{1/2}$. The Time-Langevin analysis proposed by Zhang et al. [151] reveal the forward equation corresponding to the time Markovian process (A.27)

$$\frac{\partial}{\partial u} h(u, t) = - \frac{\partial}{\partial x} \left[ A(x)p(u,x) \right] + \frac{\partial^2}{\partial x^2} \left[ D^*(x)p(u,x) \right].$$  \hspace{1cm} (A.28)

Similar to the derivation of (A.23), we obtain the fADE model by combining (A.28), (A.26), and (A.17):

$$\frac{\partial}{\partial u} p(x, u) = - \frac{\partial}{\partial x} \left[ A(x)p(x,u) \right] + \frac{\partial^2}{\partial x^2} \left[ D^*(x)p(x,u) \right].$$  \hspace{1cm} (A.29)

When $e = 0, \beta = 1$, and $q = -1$, Eq. (A.29) reduces to the fractional Fokker–Planck equation with variable velocity and fractional dispersion coefficient proposed by Srokowski and Kamińska [135]. When $e = 1$ and $D_u$ and $D_x$ are constant, (A.29) reduces to the fADE (A.23). This analysis significantly extends the fADE gained from CTRW scaling, because the coefficients may have any degree of spatial heterogeneities – a feature (due to ensemble averaging) that CTRW lack [113].

References


