

Article

Comparison of Time Nonlocal Transport Models for Characterizing Non-Fickian Transport: From Mathematical Interpretation to Laboratory Application

Bingqing Lu ¹ , Yong Zhang ^{1,2,*}, Chunmiao Zheng ³, Christopher T. Green ⁴, Charles O'Neill ⁵, Hong-Guang Sun ² and Jiazhong Qian ⁶

¹ Department of Geological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA; blu5@crimson.ua.edu

² Department of Engineering Mechanics, Hohai University, Nanjing 210098, Jiangsu, China; shg@hhu.edu.cn

³ Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control, School of Environmental Science & Engineering, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China; zhengcm@sustc.edu.cn

⁴ U.S. Geological Survey, Menlo Park, CA 94025, USA; ctgreen@usgs.gov

⁵ Department of Aerospace Engineering and Mechanics, University of Alabama, Tuscaloosa, AL 35487, USA; croneill@eng.ua.edu

⁶ School of Resources and Environmental Engineering, Hefei University of Technology, Anhui 230009, Hefei, China; qianjiazhong@hfut.edu.cn

* Correspondence: yzhang264@ua.edu; Tel.: +1-205-348-3317

Received: 8 May 2018; Accepted: 11 June 2018; Published: 13 June 2018



Abstract: Non-Fickian diffusion has been increasingly documented in hydrology and modeled by promising time nonlocal transport models. While previous studies showed that most of the time nonlocal models are identical with correlated parameters, fundamental challenges remain in real-world applications regarding model selection and parameter definition. This study compared three popular time nonlocal transport models, including the multi-rate mass transfer (MRMT) model, the continuous time random walk (CTRW) framework, and the tempered time fractional advection–dispersion equation (tt-fADE), by focusing on their physical interpretation and feasibility in capturing non-Fickian transport. Mathematical comparison showed that these models have both related parameters defining the memory function and other basic-transport parameters (i.e., velocity v and dispersion coefficient D) with different hydrogeologic interpretations. Laboratory column transport experiments and field tracer tests were then conducted, providing data for model applicability evaluation. Laboratory and field experiments exhibited breakthrough curves with non-Fickian characteristics, which were better represented by the tt-fADE and CTRW models than the traditional advection–dispersion equation. The best-fit velocity and dispersion coefficient, however, differ significantly between the tt-fADE and CTRW. Fitting exercises further revealed that the observed late-time breakthrough curves were heavier than the MRMT solutions with no more than two mass-exchange rates and lighter than the MRMT solutions with power-law distributed mass-exchange rates. Therefore, the time nonlocal models, where some parameters are correlated and exchangeable and the others have different values, differ mainly in their quantification of pre-asymptotic transport dynamics. In all models tested above, the tt-fADE model is attractive, considering its small fitting error and the reasonable velocity close to the measured flow rate.

Keywords: time nonlocal transport model; Non-Fickian diffusion; breakthrough curve

1. Introduction

Non-Fickian or anomalous transport, where the plume variance grows nonlinearly in time, has been documented extensively for solute transport in heterogeneous aquifers [1–4], soils [5–7], and rivers [8,9]. Non-Fickian transport for dissolved contaminants can occur at all scales, varying from field scale [10] to micro- and nano-scale media [11,12]. Non-Fickian diffusion, which is characterized by slow diffusion (sometimes further classified as sub-diffusion), has long been believed to relate to sorption/desorption between mobile and immobile domains under an equilibrium assumption [13] or kinetic conditions [14,15], or mass exchange between flow regions with relatively high and low velocities [16]. Note here the term “diffusion” contains both molecular diffusion and mechanical dispersion (representing the local variation of advection from the mean velocity), and hence “non-Fickian diffusion” is used interchangeable with “non-Fickian transport” in this study.

To capture non-Fickian transport induced by solute retention, various transport models have been developed, starting from the standard advection–dispersion equation (ADE) with either equilibrium or kinetic sorption [17,18] and the two-domain or two-site models proposed originally in chemical engineering [19]. Time nonlocal transport models were then developed to capture solute retention in natural geologic media with intrinsic physical and chemical heterogeneity. There are at least three popular time nonlocal transport models, which are the multi-rate mass transfer (MRMT) model [16,20], the hydrologic version of the continuous time random walk (CTRW) developed by Berkowitz and colleagues (see the extensive review in [21] and the mathematical version of CTRW in [22]), and the tempered time fractional advection–dispersion equation (tt-fADE) model [23]. Some of these models have been compared theoretically. For example, mathematical similarity between the CTRW framework, the tt-fADE, and the MRMT model was explored by [24] and [21] (whose main conclusion is discussed below for clarification), and the numerical approximation of the MRMT model using CTRW schemes was developed by [25].

While previous studies showed that most of the time nonlocal models are identical with correlated parameters, fundamental challenges remain in real-world applications regarding model selection and parameter definition. Although various stochastic models have been developed for three decades in hydrology, they have not become routine modeling tools, due to many reasons. For example, given well-controlled laboratory transport experiments using heterogeneous sand columns and conservative tracers (which have been widely used to understand real-world diffusion and the resultant transport dynamics that are non-Fickian), a newcomer faces the challenge of model selection: which time nonlocal model should be used to capture the observed non-Fickian dynamics under specific flow/transport conditions, such as conservative tracer transport in saturated, heterogeneous columns repacked in the laboratory with a stable, relatively high water flow rate? To our best knowledge, there is, unfortunately, no literature providing such an answer for this simple question. The above time nonlocal transport models were originally built upon different physical theories and contain different quantities and types of parameters. A better understanding of the potential benefits and limitations of different time nonlocal transport models as applied to non-Fickian dynamics, therefore, is required before they can be reliably applied for real-world applications and attract new users with limited knowledge in stochasticity.

This study will systematically evaluate the above (MRMT, CTRW, and tt-fADE) theoretical treatments of non-Fickian transport of conservative solutes, with respect to their parameters and ability to represent non-Fickian dispersion, especially the late-time tailing behavior, which is typical for hydrological processes in heterogeneous geological media. Late-time dynamics of contaminant transport is also a major factor in many environmental issues, such as groundwater contamination remediation and aquifer vulnerability assessment. We will then apply all the modeling methods for laboratory column transport experiments and field tracer tests. We emphasize here that the application of a time nonlocal transport model to capture non-Fickian transport is not new. What is new in this study is the quantitative, practical comparison of nonlocal transport models given experimental data.

The rest of the work is organized as follows. In Section 2, we review the above time nonlocal transport models. In Section 3, the time nonlocal transport models are applied to non-Fickian transport in multidimensional heterogeneous porous media. The applicability of those time nonlocal models is checked using laboratory experiments and field tests. Our laboratory experiments, where a conservative tracer moves through heterogeneous lightweight expanded clay aggregate (Leca) beads, exhibit typical non-Fickian diffusive behavior with elongated late-time tails. The field tracer tests also exhibit apparent tailing behaviors. Section 4 checks and compares all the time nonlocal models for characterizing the observed non-Fickian dynamics. Further analyses are shown in Section 5, where we group the transport models and briefly discuss parameter uncertainty. Conclusions are finally drawn in Section 6.

2. Review and Evaluation of Time Nonlocal Transport Models

The core of the time nonlocal models is the appropriate definition of the memory function, which controls the distribution of waiting times for contaminants trapped by immobile zones. In this section, we focus on the theoretical background, especially the memory function, for each model and explore potential correlation of critical parameters in different models. For example, previous studies emphasized that the tt-fADE is a specific form of the CTRW framework since the tt-fADE assumes a truncated power-law memory function, which is one of the memory functions previously assumed by the CTRW framework [21]. Identical functionality was also pointed out for the MRMT and the CTRW framework (see Section 1).

2.1. Multi-Rate Mass Transfer Model

The MRMT model describes mass transfer between a mobile domain and any number of immobile domains with varying properties. The linear, multi-rate, first-order solute transport equations in the absence of sources/sinks can be written as [20]:

$$\frac{\partial C_m}{\partial t} + \sum_{j=1}^n \beta_j \frac{\partial C_{im,j}}{\partial t} = -\nabla \cdot [v C_m - D \nabla C_m], \quad (1)$$

$$\frac{\partial C_{im,j}}{\partial t} = \alpha_j [C_m - C_{im,j}], \quad j = 1, 2, \dots, n, \quad (2)$$

where C_m and $C_{im,j}$ [ML^{-3}] represent the aqueous concentrations in the well-mixed mobile zone and the j -th well-mixed immobile zone, respectively; β_j [dimensionless] is the capacity coefficient usually defined as the ratio of porosities of the j -th immobile and the mobile phases; v [LT^{-1}] is the velocity vector; D [L^2T^{-1}] is the dispersion coefficient tensor; n [dimensionless] is the number of distinct immobile phases; and α_j [T^{-1}] is the first-order mass transfer rate (also called the rate coefficient) associated with the j -th immobile zone. When $n = 1$, Equation (1) reduces to the single-rate mass transfer model.

The summation term in the left-hand side of (1) can be expressed as a convolution, leading to the time-nonlocal form [16,26]:

$$\frac{\partial C_m}{\partial t} + g(t) * \frac{\partial C_{im,j}}{\partial t} = -\nabla \cdot [v C_m - D \nabla C_m], \quad (3)$$

where the symbol “*” denotes convolution, and $g(t)$ [T^{-1}] is a memory function defined by the weighted sum of the exponential decay from individual immobile zones [16]:

$$g(t) = \int_0^\infty \alpha b(\alpha) \exp(-\alpha t) d\alpha, \quad (4)$$

where $b(\alpha)$ [T] is a density function of first-order rate coefficients.

In terms of similarities with the other nonlocal methods discussed below, the MRMT model captures the time nonlocality caused by the diffusion-limited transport of solutes in immobile zones. A practical advantage of this approach is that the memory function has explicit hydrogeological meaning and thus it may be calculated, fitted, or even predicted [27]. Finally, the parameters v , D , and $g(t)$ can be spatially variable, so the MRMT method can capture the local variation of solute transport velocity caused by heterogeneity.

2.2. Tempered Time Fractional Advection–Dispersion Equation Model

The tt-fADE is one analytic technique that accounts for the time nonlocality of the medium, and simultaneously accounts for convergence of a stochastic solute particle motion process (i.e., a CTRW) to a limit distribution. In particular, if the distribution of trapping times between the movement of solute particles has an infinite mean, then the overall transport equation has one fractional-order derivative representing “dispersion” in time, leading to the standard time fractional advection–dispersion equation (t-fADE) model. Assuming a power-law memory function to describe random waiting times in the immobile zones [28]:

$$g(t) = \frac{t^{-\gamma}}{\Gamma(1-\gamma)}, \quad (5)$$

where $\Gamma(\cdot)$ is the Gamma function, and the exponent $0 < \gamma \leq 1$ (when γ approaches to 1, $\Gamma(1-\gamma)$ approaches to infinity, making no memory effects $g(t) = 0$, and the model would behave like the traditional ADE model with a retardation coefficient $1 + \beta$). Then by definition,

$$\frac{\partial C_m(x, t)}{\partial t} * g(t) = \frac{\partial C_m(x, t)}{\partial t} * \frac{t^{-\gamma}}{\Gamma(1-\gamma)} = \frac{\partial^\gamma C_m(x, t)}{\partial t^\gamma}, \quad (6)$$

is a Caputo fractional derivative of order γ . Inserting (6) into the MRMT model (1) and assuming that the solute is initially placed in the mobile zone only, one obtains the following standard t-fADE describing the mobile and immobile solute transport:

$$\frac{\partial C_m}{\partial t} + \beta \frac{\partial^\gamma C_m}{\partial t^\gamma} = -\nabla \cdot [v C_m - D \nabla C_m] - \beta C_m(x, t = 0) \frac{t^{-\gamma}}{\Gamma(1-\gamma)}, \quad (7)$$

$$\frac{\partial C_{im}}{\partial t} + \beta \frac{\partial^\gamma C_{im}}{\partial t^\gamma} = -\nabla \cdot [v C_{im} - D \nabla C_{im}] + C_m(x, t = 0) \frac{t^{-\gamma}}{\Gamma(1-\gamma)}, \quad (8)$$

where C_{im} denotes the overall chemical concentration in all immobile domains.

Meerschaert et al. [23] generalized the t-fADE (7) and (8) by introducing an exponentially truncated power-law function, which is an incomplete Gamma function, as the memory function:

$$g(t) = \int_t^\infty e^{-\lambda s} \frac{\gamma s^{-\gamma-1}}{\Gamma(1-\gamma)} ds, \quad (9)$$

where $\lambda > 0$ [T^{-1}] is the truncation parameter in time. This modification leads to the tt-fADE:

$$\frac{\partial C_m}{\partial t} + \beta e^{-\lambda t} \frac{\partial^\gamma}{\partial t^\gamma} [e^{\lambda t} C_m] - \beta \lambda^\gamma C_m = -\nabla \cdot [v C_m - D \nabla C_m] - \beta C_m^0 \int_t^\infty e^{-\lambda \tau} \frac{\tau^{-\gamma-1}}{\Gamma(1-\gamma)} d\tau \quad (10)$$

$$\frac{\partial C_{im}}{\partial t} + \beta e^{-\lambda t} \frac{\partial^\gamma}{\partial t^\gamma} [e^{\lambda t} C_{im}] - \beta \lambda^\gamma C_{im} = -\nabla \cdot [v C_{im} - D \nabla C_{im}] + C_m^0 \int_t^\infty e^{-\lambda \tau} \frac{\tau^{-\gamma-1}}{\Gamma(1-\gamma)} d\tau \quad (11)$$

where $C_m^0 = C_m(x, t = 0)$ denotes the initial source located only in the mobile phase. At a time

$$t \ll 1/\lambda, \quad (12)$$

the tail of the mobile-phase breakthrough curve (BTC) declines as a power law function:

$$C_m(x, t) \propto t^{-1-\gamma}, \quad (13)$$

while at a much later time $t \gg 1/\lambda$, the slope of the mobile-phase BTC approaches negative infinity (i.e., the late-time BTC tail declines exponentially). Therefore, the value of λ controls the transition of the BTC late-time tail from a power-law function to exponential function [29].

The use of the tt-fADE (10) and (11) to model mobile/immobile anomalous solute transport is motivated by four factors: (1) the equation governs the limits of known stochastic processes; (2) it describes a combination of first-order mass transfer models and reduces to known mobile/immobile equations in the integer order case; (3) the equation has tractable solutions that model the significant features of solute plume evolution in time and space; and (4) the equation is parsimonious, with no more parameters than the standard MRMT model (1) with multiple pairs of rate and capacity coefficients. To summarize, the tt-fADE has one major limitation compared to the MRMT model (1) and the CTRW model discussed below: the memory function embedded in the tt-fADE is a specific form (9), while the memory function used in the MRMT and the CTRW model can have different forms. Potential effects of this difference on the models' abilities to estimate real-world physical behavior is explored below, in Section 4, and discussed in Section 5.

2.3. Continuous Time Random Walk Framework

Derivation of the CTRW model in hydrologic sciences [21] starts from the generalized master equation with the kernel Φ defined as (in Laplace space ($t \rightarrow s$)) [30]:

$$\tilde{\Phi}(x, s) = \frac{s \tilde{\Psi}(x, s)}{1 - \tilde{\phi}(s)}, \quad (14)$$

where the tilde “~” denotes the Laplace transform, $\tilde{\Psi}(x, s)$ is the Laplace transform of the joint density of jump length and duration, and $\tilde{\phi}(s)$ is the Laplace transform of the transition time or duration density $\phi(t) = \int_{-\infty}^{+\infty} \Psi(x, t) dx$ used in the master equation [31]

$$\tilde{p}(k, s) = \frac{1 - \tilde{\phi}(s)}{s} \frac{1}{1 - \tilde{\phi}(s) \hat{f}(k)}, \quad (15)$$

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

Berkowitz et al. [21] defined a memory function M to replace $\tilde{\phi}(s)$ in (15)

$$\tilde{M}(s) = \frac{t_1 s \tilde{\phi}(s)}{1 - \tilde{\phi}(s)}, \quad (16)$$

where $t_1 [T]$ denotes a “typical median transition time” for particles. Inserting (16) into (15), and taking the Laplace and Fourier inverse transform, the following well-known CTRW framework was obtained [21]:

$$\frac{\partial p(x, t)}{\partial t} = \int_0^t M(t - \tau) \left[-v_\psi \frac{\partial}{\partial x} + D_\psi \frac{\partial^2}{\partial x^2} \right] p(x, \tau) d\tau, \quad (17)$$

where v_ψ denotes the average velocity, and D_ψ is the dispersion coefficient. When deriving the CTRW (17) from the master Equation (15), the jump size density $\tilde{f}(k)$ in (15) needs to be expanded as

$$\hat{f}(k) \approx 1 - \mu i k + \frac{\alpha^2}{2} (i k)^2, \quad (18)$$

resulting in the spatially-averaged velocity v_ψ and dispersion coefficient D_ψ in (17):

$$v_\psi = \frac{\mu}{t_1}, \quad (19)$$

$$D_\psi = \frac{\sigma^2}{2t_1}, \quad (20)$$

Therefore, here the “typical median transition time” t_1 is the mean waiting time:

$$t_1 = \int_0^{+\infty} t \phi(t) dt, \quad (21)$$

The CTRW memory function M defined by (16) has various forms to capture various breakthrough curves (BTCs). One popular form is the exponentially truncated power-law, defined by the transition time ϕ (see Equation (16) in [32]):

$$\phi(t) = B \frac{\exp(-t/t_2)}{(1 + t/t_1)^{1+\xi}}, \quad (22)$$

where the factor $B = \{t_1 \tau_2^{-\xi} \exp(\tau_2^{-1}) \Gamma(-\xi, \tau_2^{-1})\}^{-1}$ (with $\tau_2 = t_2/t_1$) keeps the integral of $\phi(t)$ to be 1. For simplicity and direct comparison between models, we constrain the exponent ξ to be $0 < \xi < 1$ in this study. The Laplace transform of (22) is (see Equation (17) in [32])

$$\tilde{\phi}(s) = \frac{\Gamma(-\xi, \tau_2^{-1} + t_1 s)}{\Gamma(-\xi, \tau_2^{-1})} (1 + \tau_2 s t_1)^\xi \exp(t_1 s), \quad (23)$$

A simple manipulation shows the relationship between the MRMT memory function $g(t)$ and the CTRW memory function $M(t)$ in Laplace space:

$$\tilde{M}(s) = \frac{1}{\theta_M + \theta_I \tilde{g}(s)}, \quad (24)$$

where θ_M [dimensionless] and θ_I [dimensionless] are the porosity in the mobile and (total) immobile domains, respectively. Inserting (16) and (23) into (24), we obtain:

$$\begin{aligned} \tilde{g}(s) &= \frac{1 - \tilde{\phi}(s)}{\theta_I t_1 s \tilde{\phi}(s)} - \frac{\theta_M}{\theta_I} \\ &= \frac{1}{\theta_I t_1 s} \frac{\Gamma(-\xi, \tau_2^{-1})}{\Gamma(-\xi, \tau_2^{-1} + t_1 s) (1 + \tau_2 s t_1)^\xi \exp(t_1 s)} - \frac{1}{\theta_I s t_1} - \frac{\theta_M}{\theta_I}, \end{aligned} \quad (25)$$

There is no known analytical solution for g in real time t , except for the following asymptote at late time $t \gg t_1 \tau_2$:

$$g(t) \propto \exp\left(-\frac{t}{t_1 \tau_2}\right) = \exp\left(-\frac{t}{t_2}\right), \quad (26)$$

Based on (26), we obtain the late-time growth rate for p

$$p(t_{late}) \propto -\frac{\partial g(t)}{\partial t} \propto \exp\left(-\frac{t}{t_2}\right), \quad (27)$$

Therefore, for time $t \gg t_2$, the non-Fickian transport transitions to Fickian diffusion. The cutoff time scale t_2 in (22), as explained by [32], “corresponds to the largest heterogeneity length scale”. In another numerical study by Willmann et al. [33], t_2 was also called “the late cutoff time”. The above analysis shows that t_2 is functionally equivalent to the inverse of the truncation parameter λ used in the tt-fADE model (10) and (11).

For the intermediate time $t_1 \ll t \ll t_2$, one can obtain the memory function g by solving (25) numerically (e.g., using the numerical inverse Laplace transform). This was done by [32], who found that the transition probability scales as a power-law function

$$p(t) \propto t^{-\xi-1}, \text{ where } t_1 \ll t \ll t_2 \quad (28)$$

Comparing (28) and (13), we find that the power-law exponent ξ in the CTRW framework is functionally equivalent to the scale index γ in the tt-fADE model (10) and (11).

Hence, the parameters in the CTRW framework (17) (e.g., ξ and t_2) are related to the parameters in the tt-fADE model (10) and (11) (γ and λ), except for t_1 in (17), which may be estimated by the mean diffusive time. Parameters predicted by one model (such as the tt-fADE) may also help to improve the estimated parameters of the other (i.e., the CTRW framework).

3. Applications: Capturing Non-Fickian Transport in Multidimensional Porous Media

Here we apply the above time nonlocal transport models to capture non-Fickian transport observed in multidimensional, heterogeneous porous media. Well-controlled laboratory experiments of sand column transport were conducted, to provide data to evaluate the three nonlocal transport models reviewed above. We used a cylindrical organic glass (polymethyl methacrylate) tube filled with non-uniform “lightweight expanded clay aggregate” (Leca) beads to monitor solute transport in saturated porous media. Leca beads were selected since they contain high intra-granular porosity (and potential for sorption on the large surface area of clay particles), which can lead to retention for solute transport as often occurs in the field. The length of the Leca column was 100 cm with an internal diameter of 4 cm. The experimental apparatus was composed of inflow and outflow water tanks, a porous medium column, tubing, and a detection device (Figure 1). The diameter of the Leca beads varied from 1.0 to 2.0 mm. A pulse of Brilliant Blue FCF (BBF), which is a conservative organic compound typically used as colorant for foods, with a volume of 5 mL was injected into the column (from the bottom of the glass tube set vertically) at a concentration of 0.1 g/L, representing an instantaneous point source. For the relatively short time-scales of these experiments, we believe the mechanical dispersion of BBF is predominant, and effects of molecular diffusion are negligible. An ultraviolet visible light spectrophotometer was used to measure the absorbance of solute, and the absorbance was then converted to concentration. Continuous sampling provided tracer BTCs used to check the applicability of the nonlocal transport models.

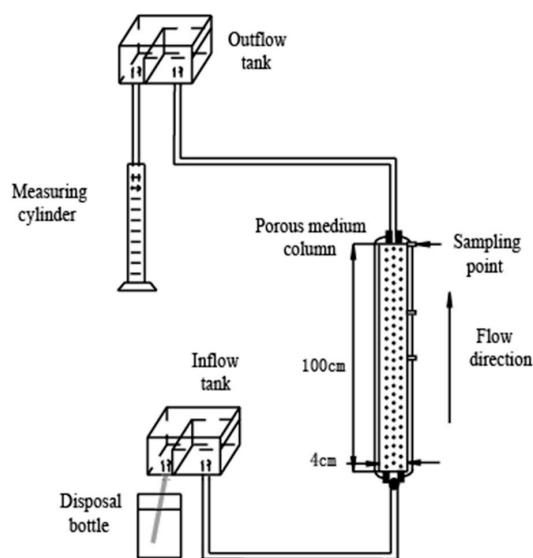


Figure 1. Experimental setup.

Six flow rates (increasing from 0.4, 0.6, 0.8, 1.0, 1.2, to 1.4 mL/s; see Table 1) were carried out in the experiment. For each flow rate, three experimental runs were conducted where the BBF BTCs were collected at three sections of the sand column, varying from 50, 70, to 100 cm. During each run with the similar flow rate, the hydraulic gradient between the inlet and outlet was kept constant in time, and therefore the flow velocity for the three experimental runs should be close to each other.

Table 1. Measured and fitted parameters (using the advection–dispersion equation (ADE) model with equilibrium sorption) for Brilliant Blue FCF (BBF) transport through the “lightweight expanded clay aggregate” (Leca) beads at different flow rates and travel distances and field tests. In the legend, Q represents the flow rate; L denotes the travel distance (i.e., the length of the column); v^* ($=v/R$) is the average flow velocity divided by the retardation coefficient; D^* ($=D/R$) is the dispersion coefficient divided by the retardation coefficient; θ is the porosity; and $RMSE$ stands for root mean square error between observed values and predicted values.

Experiment	Q (mL/s)	L (cm)	v^* (mm/s)	D^* (mm ² /s)	Θ	$RMSE$
	Measured	Measured	Fitted	Fitted	Measured	Calculated
Lab	0.4	50	1.20	5.0	0.39	2.75
		70	1.12	7.0	0.39	1.54
		100	1.07	5.3	0.39	2.30
	0.6	50	1.45	6.8	0.39	2.16
		70	1.55	15	0.39	0.78
		100	1.63	9.2	0.39	1.99
	0.8	50	2.23	15	0.39	1.57
		70	2.20	22	0.39	0.97
		100	2.07	14	0.39	1.55
	1	50	2.67	20	0.39	1.85
		70	2.85	24	0.39	1.43
		100	2.90	16	0.39	2.92
	1.2	50	3.25	23	0.39	2.80
		70	3.22	28	0.39	1.53
		100	3.04	19	0.39	2.40
	1.4	50	3.48	40	0.39	1.09
		70	3.85	50	0.39	0.90
		100	3.70	38	0.39	1.33
Field	83.33	200	0.003	0.694	0.3	n/a
		400	0.004	0.984	0.3	n/a
		600	0.004	0.926	0.3	n/a

The measured BTCs all exhibit late-time tailing behavior (see Figure 2 and Figures S1–S5 in the Supplementary File), which is one of the major characteristics of non-Fickian transport.

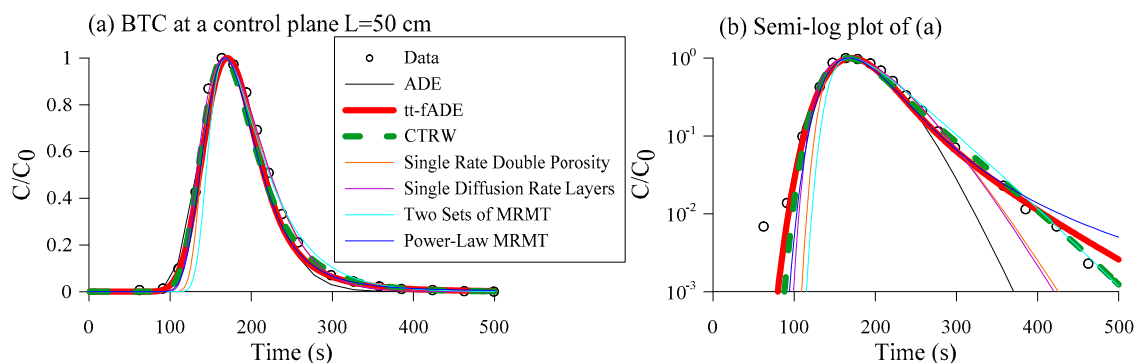


Figure 2. Cont.

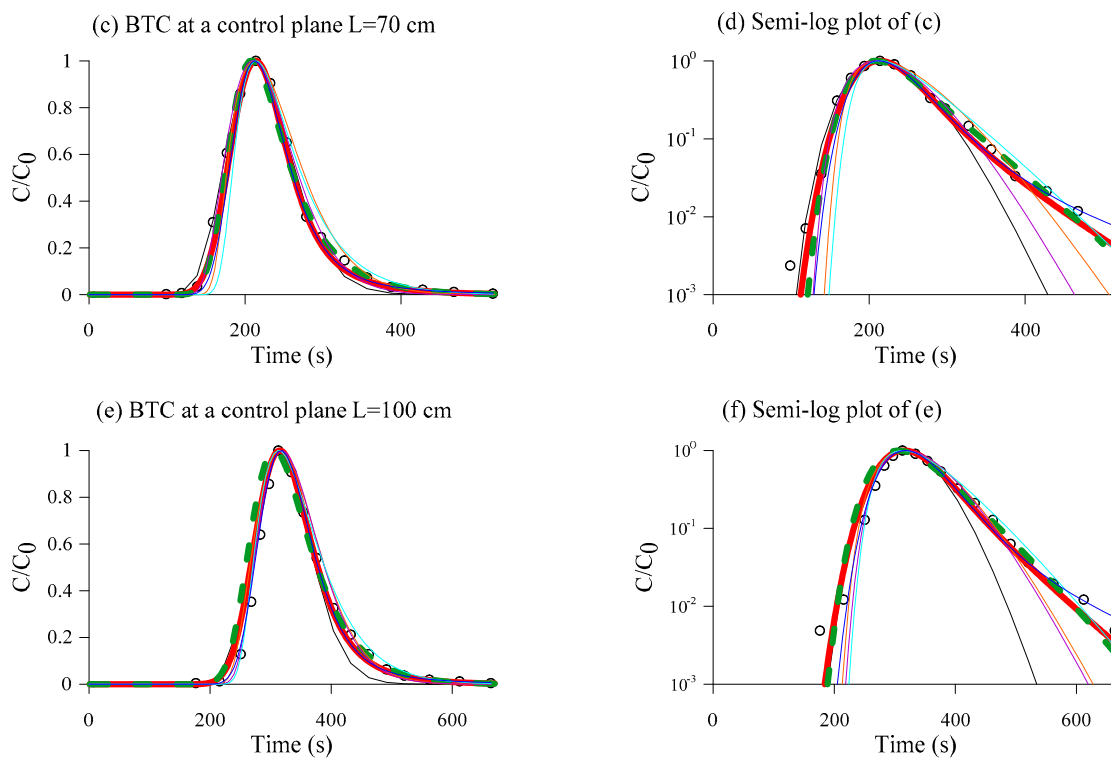


Figure 2. Laboratory tracer test: Comparison between the measured (symbols) and the modeled (lines) breakthrough curves (BTCs) using the advection–dispersion equation (ADE), the tempered time fractional advection–dispersion equation (tt-fADE) (red thick lines), the continuous time random walk (CTRW) (green dashed lines), and the multi-rate mass transfer (MRMT) model with the water flow rate $Q = 1.4$ mL/s.

4. Model Fit and Comparison

Comparisons of the measured and best-fit BTCs using the above three time-nonlocal transport models and the classical ADE are shown in Figure 2 and Figures S6–S10 in the Supplementary File. In this section, we briefly introduce the model fitting process, and then compare the model results.

4.1. The ADE Model with Equilibrium Sorption

For comparison, we first use the classic ADE with equilibrium sorption to quantify the measured BTCs. The governing equation for one-dimensional chemical transport in groundwater with advection, dispersion, and retardation is [34]:

$$R \frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} - v \frac{\partial C}{\partial x}, \quad (29)$$

which has the following solution with an instantaneous point source at the origin:

$$C(x, t) = \frac{M}{2A\sqrt{\pi D^* t}} \exp \left[-\frac{(x - v^* t)^2}{4D^* t} \right], \quad (30)$$

where M [M] represents the initial mass injected into the column; A [L^2] is the cross-section area; and R [dimensionless] is the retardation coefficient. From a mathematical perspective, R acts as a rescaling factor in time. Hence, we reduce the three parameters v , D , and R in the ADE model to two parameters

v^* ($=v/R$) and D^* ($=D/R$). The fitted values for parameters v^* and D^* are listed in Table 1. These two parameters were calibrated manually based on visual inspection.

In our Leca-column transport experiments, the relatively large flow velocity led to a large Peclet number ($Pe \gg 1$) and the dispersion coefficient D relates to v via $D = \alpha v$, where α [L] denotes the dispersivity. The relationship between dispersivity α and the travel distance in the laboratory experiments is shown in Figure 3. For a fixed flow rate, α varies with the travel distance without any fixed trend, perhaps due to varying dispersion over the relatively short travel distance.

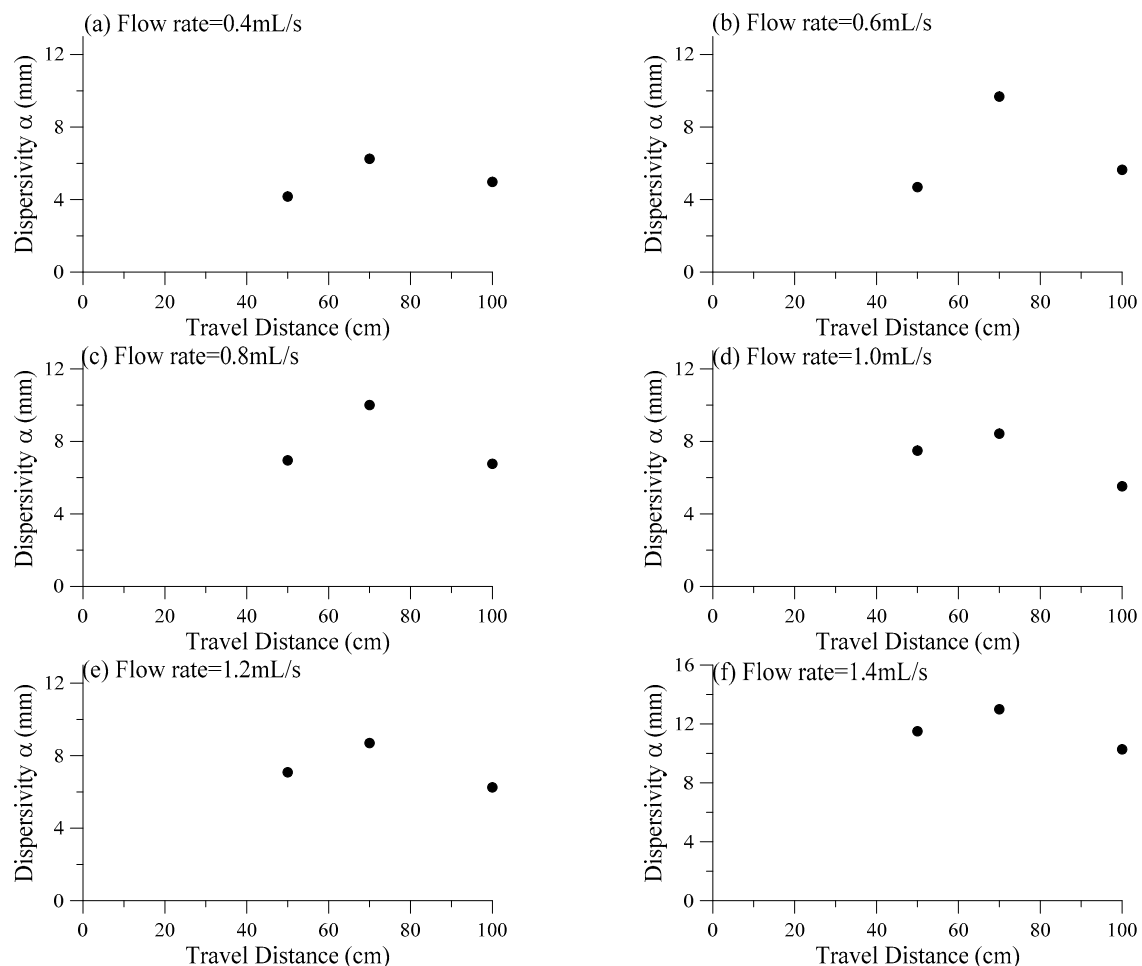


Figure 3. The best-fit dispersivity in the ADE model (29) versus the travel distance.

The relationship between dispersivity and flow rate in the laboratory experiments is shown in Figure 4. Although the best-fit dispersivity fluctuates with the flow rate, it generally increases with an increasing flow velocity. A larger flow velocity causes a wider spatial distribution of contaminant plume and increases the apparent dispersivity in the ADE model. This phenomenon is consistent with known non-Fickian solute transport behavior and indicates that the ADE model is performing as expected [33].

In the BTCs measured at three travel distances with six different flow rates in the laboratory experiments and field tracer tests, the ADE model (29) captures the rapid increase of early-time BTCs, but overestimates the decline of the late-time BTC tails although the delay of transport due to retardation was considered in the ADE (29). This implies that the delayed transport observed in our laboratory experiments and field tests is more complex than equilibrium adsorption described in (29), where the sorbed or immobile concentration is a simple linear function of the dissolved concentration.

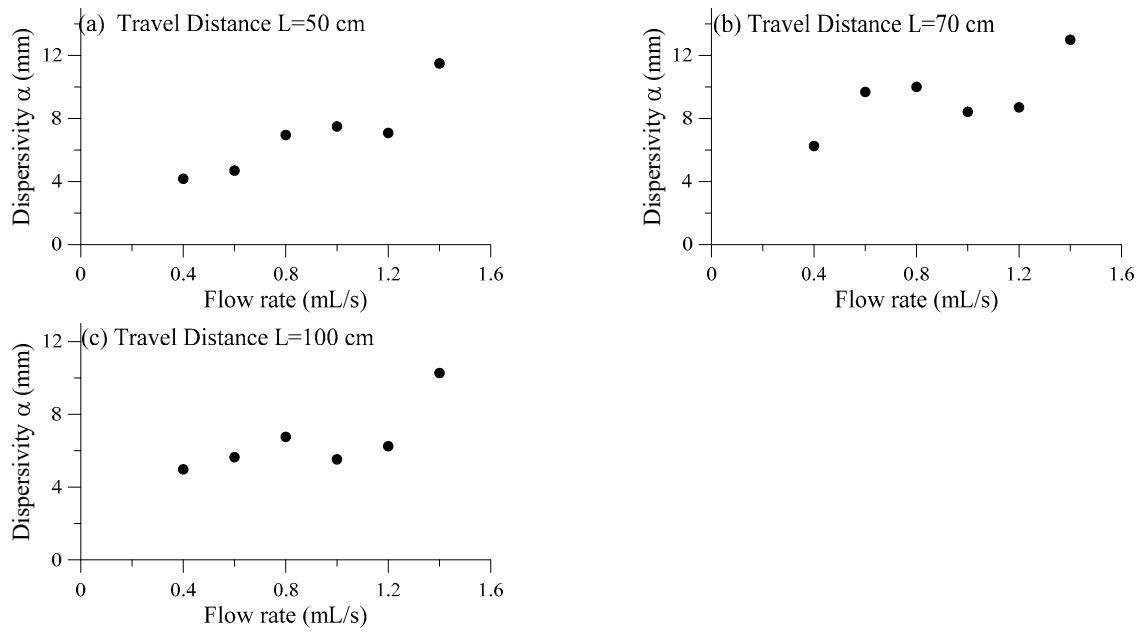


Figure 4. The best-fit dispersivity in the ADE model versus the water flow rate.

4.2. The MRMT Model

The MRMT model (1) has various specific forms, which might be useful for applications and therefore require further evaluation. For example, the immobile domain can be characterized as multiple layers with a distribution of diffusion rate coefficients or first-order mass-transfer rates, or be simplified as a layer with a single diffusion rate which can be attractive in applications due to its simplicity in manipulation. There is, however, no solid physical justification for the selection of any of these forms, given simply the limited information for a sand column like the one used in our laboratory. For systematic analyses of all potential mass transfer models, we selected the following four MRMT subsets with different mass-transfer formulations:

- (1) MRMT model 1 (single mass-transfer rate): the immobile zones can be simplified by a single, homogeneous first-order mass-transfer rate (which is also the single-rate double-porosity model);
- (2) MRMT model 2 (single diffusion rate): the immobile zones have a single diffusion rate for all layers;
- (3) MRMT model 3 (two mass-transfer rates): the immobile zones have two sets of rate coefficients;
- (4) MRMT model 4 (multiple mass-transfer rates): the immobile zones have a power-law distribution of (first-order mass-transfer) rate coefficients.

We use the STAMMT-L version 3.0 code [35] to solve the above four MRMT models. Best-fit results, which were calibrated manually, are shown in Figure 2 and Figures S6–S10 and discussed below.

4.2.1. MRMT Model 1 with a Single Mass-Transfer Rate

The general mobile-immobile (MIM) model with diffusion in the immobile zone can be written as [20]:

$$\frac{\partial C_{m,t}}{\partial t} + T(x,t) = -v\nabla C_{m,t} + D\nabla^2 C_{m,t}, \quad (31)$$

$$\frac{\partial C_{im,t}}{\partial t} = \frac{Da}{r^{n-1}} \nabla \left[r^{n-1} \nabla C_{im,t} \right], \quad (32)$$

where $T(x,t)$ [$ML^{-3}T^{-1}$] is a transient term accounting for rate-limited mass transfer between the mobile and immobile domains, and Da is the apparent diffusion coefficient. Here n denotes the

dimensionality of the problem, and $n = 1, 2, 3$ denotes diffusion into layers, cylinders, and spheres, respectively. For a single rate first-order mass-transfer approximation, $T(x, t)$ can be expressed as [35]:

$$T(x, t) = \beta \frac{\partial C_{im,t}}{\partial t}, \quad (33)$$

$$\frac{\partial C_{im,t}}{\partial t} = \alpha [C_{m,t} - C_{im,t}], \quad (34)$$

which is the single rate version of the MRMT model (1).

The best-fit parameters using the single-rate MIM model (31)–(34) are listed in Table 2. The fitting exercise showed the sensitivity of model parameters to the travel distance and water flow rate in the laboratory experiments. First, the dispersivity α_L does not change with the travel distance (Table 2), since the plume expansion with time is captured by the mass transfer term in the model. This is different from the standard ADE model, where the dispersivity must increase with the travel distance to capture scale-dependent dispersion.

Table 2. The fitted parameters for the single mass-transfer rate mobile-immobile model (i.e., Equations (31)–(34)) at different flow rates and travel distances. In the legend, L denotes the travel distance (i.e., the length of the column); α_L is the dispersivity; v is the flow velocity; β_{tot} is the total capacity coefficient; and α stands for the mass transfer rate.

Experiment	Q (mL/s)	L (cm)	α_L (mm)	v (mm/s)	β_{tot}	α (s ^{−1})
	Measured	Measured	Fitted	Fitted	Fitted	Fitted
Lab	0.4	50	1.5	1.4	0.50	0.05
		70	1.5	1.35	0.50	0.04
		100	1.5	1.35	0.50	0.04
	0.6	50	1.5	1.68	0.50	0.05
		70	1.5	1.71	0.40	0.04
		100	1.5	1.68	0.31	0.03
	0.8	50	1.5	2.38	0.45	0.06
		70	1.5	2.32	0.40	0.04
		100	1.5	2.15	0.33	0.05
	1	50	1.5	2.75	0.45	0.06
		70	1.5	2.93	0.43	0.06
		100	1.5	2.93	0.35	0.07
	1.2	50	1.5	3.15	0.30	0.06
		70	1.5	3.15	0.30	0.06
		100	1.5	3.13	0.28	0.06
	1.4	50	1.5	3.52	0.35	0.06
		70	1.5	3.95	0.35	0.06
		100	1.5	3.98	0.35	0.06
Field	83.33	200	210	0.004	0.50	0.14
		400	230	0.004	0.50	0.10
		600	230	0.004	0.50	0.03

Second, the average velocity used in the model is slightly larger than the measured BTC peak velocity and the ADE velocity (note that the ADE velocity is also larger than the real BTC peak velocity), probably due to the assumption that there might be an immobile domain interacting with the mobile domain. The parameters v and D in the MRMT and ADE models have different meanings and hence may not have the same values. Parameters v and D in the MRMT model refer to the mobile domain [36], and therefore the velocity in the MRMT model is generally larger than the ADE velocity, while the opposite is expected for the dispersion coefficient D . This holds true for all the MRMT formulations.

Third, the capacity coefficient β either remains constant or decreases very slightly with the travel distance, implying that the medium heterogeneity might not significantly change with the medium's length.

Fourth, the rate coefficient α increases slightly and the capacity coefficient β decreases slightly with an increasing flow rate at each control plane. This subtle change is likely due to the assumption that the faster water flows correspond to less volumetric proportion of immobile domains. A faster flow may decompose immobile domains and enhance the mass exchange between mobile and immobile domains, resulting in a larger mass exchange rate α .

4.2.2. MRMT Model 2 with a Single Diffusion Rate

The layered diffusion model is a specific case of the MRMT model [20] with the following rate and capacity coefficients:

$$\alpha_j = \frac{(2j-1)^2 \pi^2}{4} \frac{D_a}{a^2}, \quad j = 1, 2, \dots, N_{im}, \quad (35)$$

$$\beta_j = \frac{8}{(2j-1)^2 \pi^2} \beta, \quad j = 1, 2, \dots, N_{im}, \quad (36)$$

The best-fit parameters for model (35) and (36) are listed in Table 3. There is no apparent correlation between the mass transfer rate and the travel distance or water flow velocity. The same conclusion is found for the capacity coefficient.

Table 3. The fitted parameters for the single diffusion rate multi-rate mass transfer (MRMT) model at different flow rates and travel distances. In the legend, α_d is the diffusion rate coefficient ($\alpha_d = D_a/a^2$, where D_a is the apparent diffusion coefficient and a is the layer half-thickness).

Experiment	Q (mL/s)	L (cm)	α_L (mm)	v (mm/s)	β_{tot}	α_d (s ⁻¹)
	Measured	Measured	Fitted	Fitted	Fitted	Fitted
Lab	0.4	50	1.5	1.55	0.60	0.02
		70	1.5	1.48	0.60	0.02
		100	1.5	1.43	0.60	0.02
	0.6	50	1.5	1.73	0.50	0.03
		70	1.5	2.22	0.80	0.02
		100	1.5	2.12	0.61	0.02
	0.8	50	1.5	2.69	0.61	0.03
		70	1.5	2.85	0.70	0.02
		100	1.5	2.55	0.55	0.03
	1	50	1.5	3.00	0.55	0.03
		70	1.5	3.30	0.55	0.02
		100	1.5	3.25	0.48	0.05
	1.2	50	1.5	3.73	0.48	0.04
		70	1.5	3.69	0.48	0.03
		100	1.5	3.25	0.31	0.03
	1.4	50	1.5	4.25	0.60	0.03
		70	1.5	4.90	0.60	0.03
		100	1.5	4.10	0.40	0.02
Field	83.33	200	210	0.004	0.50	0.05
		400	210	0.004	0.50	0.04
		600	210	0.005	0.50	0.03

4.2.3. MRMT Model 3 with Two Mass-Exchange Rates

MRMT model 3 contains two pairs of coefficients: two rate coefficients (α_1 and α_2) and two capacity coefficients (β_1 and β_2) corresponding to the first and the second immobile domains, respectively.

The best-fit parameters are listed in Table 4. The dispersivity α_L remains constant, since solute plume expansion (due likely to solute retention) is mainly captured by mass exchange between mobile and the two immobile zones. Or in other words, the two pairs of parameters, α_j ($j = 1, 2$) and β_j ($j = 1, 2$), control the mass exchange. Their values fluctuate (without predictable trends) with the travel distance and flow rate in the laboratory experiments, although α_2 decreases with increasing α_1 for the same flow rate. There is no efficient way to directly measure the rate coefficient or capacity coefficient, which creates a challenge for relating MRMT model parameters to observable physical phenomena, and for providing information to MRMT models based on ancillary observations in heterogeneous media.

Table 4. The fitted parameters for the two-set MRMT model at different flow rates and travel distances. In the legend, α_1 and β_1 represent the mass transfer rate and capacity coefficient of the first immobile domain, respectively; and α_2 and β_2 are the mass transfer rate and capacity coefficient of the second immobile domain, respectively.

Experiment	Q (mL/s)	L (cm)	α_L (mm)	v (mm/s)	α_1 (s ⁻¹)	β_1	α_2 (s ⁻¹)	β_2
	Measured	Measured	Fitted	Fitted	Fitted	Fitted	Fitted	Fitted
Lab	0.4	50	1.5	1.12	0.08	0.08	0.015	0.15
		70	1.5	1.10	0.08	0.09	0.015	0.15
		100	1.5	1.03	0.10	0.09	0.008	0.10
	0.6	50	1.5	1.31	0.08	0.07	0.020	0.13
		70	1.5	1.54	0.15	0.08	0.020	0.25
		100	1.5	1.55	0.15	0.08	0.015	0.15
	0.8	50	1.5	1.96	0.05	0.15	0.030	0.08
		70	1.5	2.08	0.05	0.20	0.015	0.10
		100	1.5	1.88	0.07	0.10	0.018	0.10
	1	50	1.5	2.30	0.05	0.12	0.030	0.15
		70	1.5	2.50	0.07	0.13	0.025	0.13
		100	1.5	2.57	0.20	0.11	0.020	0.11
	1.2	50	1.5	2.85	0.07	0.12	0.020	0.13
		70	1.5	2.87	0.06	0.11	0.027	0.13
		100	1.5	2.87	0.07	0.10	0.025	0.01
	1.4	50	1.5	3.33	0.08	0.17	0.030	0.17
		70	1.5	3.80	0.09	0.17	0.030	0.17
		100	1.5	3.75	0.08	0.15	0.030	0.16
Field	83.33	200	210	0.004	0.4	0.2	0.1	0.2
		400	210	0.004	0.4	0.2	0.1	0.2
		600	210	0.004	0.4	0.1	0.1	0.1

4.2.4. MRMT Model 4 with Power-Law Distributed Rate Coefficients

The density of rate coefficient for MRMT model 4 can be defined as [16]:

$$b(\alpha) = b(\alpha_{min}, \alpha_{max}, k) = \begin{cases} \beta_{tot} \frac{(k-2)\alpha^{k-3}}{\alpha_{max}^{k-2} - \alpha_{min}^{k-2}} & k \neq 2 \\ \beta_{tot} \frac{1}{\ln(\alpha_{max}/\alpha_{min})\alpha} & k = 2 \end{cases}, \text{ for } \alpha_{min} \leq \alpha \leq \alpha_{max} \quad (37)$$

where α_{min} [T^{-1}] denotes the minimum rate coefficient, α_{max} [T^{-1}] is the upper boundary of the rate coefficient, and k is the exponent.

In the fitting parameters shown in Table 5, the dispersivity α_L remains stable for the same reason mentioned above for the other MRMT models. Flow velocity used in this model can be approximated by the peak velocity for the measured BTC. The total capacity coefficient (β_{tot}) does not significantly change with the travel distance. The exponent k controls the slope of the late-time BTC in a log-log plot, and hence a larger k denotes faster decline of the late-time BTC. The best-fit k slightly increases with an increasing flow rate, due likely to the relatively faster decline of the late-time solute concentration under a larger water flow rate.

Table 5. The best-fit parameters for the power-law MRMT model 4 at different flow rates and travel distances. In the legend, k stands for the slope of the late-time tail, and α_{min} and α_{max} stand for the lower and upper boundary of the mass transfer rates, respectively.

Experiment	Q (mL/s)	L (cm)	α_L (mm)	v (mm/s)	β_{tot}	k	α_{min} (s ⁻¹)	α_{max} (s ⁻¹)
	Measured	Measured	Fitted	Fitted	Fitted	Fitted	Fitted	Fitted
Lab	0.4	50	1.6	1.40	0.50	2.250	0.010	0.6
		70	1.6	1.35	0.50	2.250	0.010	0.6
		100	1.6	1.35	0.50	2.250	0.010	0.6
	0.6	50	1.6	1.69	0.50	2.292	0.011	0.6
		70	1.6	1.85	0.60	2.292	0.012	0.6
		100	1.6	2.10	0.60	2.292	0.012	0.6
	0.8	50	1.6	2.83	0.70	2.290	0.013	0.6
		70	1.6	2.95	0.80	2.290	0.012	0.6
		100	1.6	3.05	0.80	2.290	0.013	0.6
	1	50	1.6	3.59	0.80	2.300	0.013	0.6
		70	1.6	4.05	0.85	2.300	0.013	0.6
		100	1.6	4.16	0.80	2.300	0.014	0.6
	1.2	50	1.6	4.63	0.80	2.335	0.014	0.6
		70	1.6	4.63	0.80	2.335	0.014	0.6
		100	1.6	4.44	0.75	2.335	0.014	0.6
	1.4	50	1.6	4.95	0.85	2.358	0.014	0.6
		70	1.6	5.48	0.80	2.358	0.015	0.6
		100	1.6	5.50	0.80	2.358	0.015	0.6
Field	83.33	200	210	0.004	0.5	2.25	0.01	0.6
		400	210	0.004	0.5	2.25	0.01	0.6
		600	210	0.005	0.5	2.25	0.01	0.6

It is also noteworthy that, on one hand, the minimum mass transfer rate (α_{min}) in (37) controls the maximum waiting time for solute particles, which is functionally equivalent to the inverse of the cutoff time scale t_2 in the CTRW framework and the truncation parameter λ in the tt-fADE model. In general, α_{min} increases with an increasing flow rate (Table 5). Faster flow may accelerate the mass exchange between mobile and immobile domains, generating shorter mean residence times for solute particles in the immobile domain and leading to a greater mass transfer rate. On the other hand, the maximum mass transfer rate (α_{max}) in (37) defines the shortest waiting time (for solute particles between two displacements), whose impact on the late-time transport dynamics can be overwhelmed by the other smaller rates. Numerical results also show that α_{max} apparently does not change with the travel distance and flow rate. Hence, α_{max} can be kept constant for all cases (Table 5).

As shown in Figure 2 and Figures S6–S10, the MRMT model 4 captures the BTC late-time tail much better than the other mass-transfer formations with fewer rate coefficients. However, compared with the tt-fADE model, the simulated tail of the MRMT model 4 tends to be slightly heavier at the end of the modeling time. In other words, it slightly underestimates the mass transfer rate at the late time. Note that the memory function in the tt-fADE is not exactly the same as that in the MRMT model 4. The tt-fADE has exponentially-truncated power-law rate coefficients, while the MRMT model 4 simply deletes any rate coefficient larger than α_{max} and smaller than α_{min} . This subtle difference in memory functions between the tt-fADE and MRMT model 4 might be the reason for the differences observed here. In addition, the MRMT model 4 (with six model parameters) requires one more parameter than the tt-fADE model.

As shown in Figure 2 and Figures S6–S10, both the MRMT model 1 and model 2 can capture most characteristics of the observed BTCs, except for the late-time tailing. Therefore, we need more than one mass-transfer rate to fit the solute transport in the one-dimensional heterogeneous sand column and field tracer tests. Increasing the number of immobile domains improves the model's performance at late times, but more immobile domains can complicate the model application. To be specific, adding

one more immobile domain adds two unknown parameters (a pair of capacity coefficient and mass exchange rate coefficient for that immobile domain).

4.3. The tt-fADE Model (10) and (11)

The failure of the standard ADE model and the single rate mass transfer model in capturing the late-time BTC tailing motivates the application of the other time nonlocal transport models such as the tt-fADE model (10) and (11). A numerical solver of (10) and (11) can be found in [37]. The best-fit results using this model are shown in Figure 2 and Figures S1–S5, where both the peak and the late time tailing can be captured simultaneously.

The best-fit parameters are shown in Table 6, where the best-fit velocity v is slightly larger than the measured peak velocity v_{peak} , due to the fractional-order capacity coefficient β in the left-hand side of the tt-fADE Equations (10) and (11) when $\gamma \rightarrow 1$:

$$v = (1 + \beta) v_{peak}, \quad (38)$$

Table 6. The best-fit parameters for the tempered time fractional advection–dispersion equation (tt-fADE) model (10) and (11) at different flow rates and travel distances. In the legend, α is the dispersivity; γ is the time/scale index; β is the fractional-order capacity coefficient; and λ is the truncation parameter.

Experiment	Q (mL/s)	L (cm)	v (mm/s)	D (mm ² /s)	α (mm)	γ	β (s $^{\gamma-1}$)	λ (s $^{-1}$)	$RMSE$
	Measured	Measured	Fitted	Fitted	$\alpha = D/v$	Fitted	Fitted	Fitted	Calculated
Lab	0.4	50	1.10	0.2	0.18	0.250	0.011	0.010	0.379
		70	1.05	0.2	0.19	0.250	0.011	0.010	0.171
		100	1.00	0.2	0.20	0.250	0.011	0.010	0.343
	0.6	50	1.30	0.2	0.15	0.292	0.013	0.011	0.461
		70	1.40	0.4	0.29	0.292	0.013	0.012	0.296
		100	1.49	0.4	0.27	0.292	0.013	0.012	0.489
	0.8	50	1.91	0.5	0.26	0.290	0.018	0.013	0.577
		70	1.95	0.6	0.31	0.290	0.018	0.012	0.398
		100	1.84	0.6	0.33	0.290	0.018	0.013	0.443
	1	50	2.25	1.8	0.80	0.300	0.014	0.013	0.529
		70	2.40	1.8	0.75	0.300	0.014	0.013	0.218
		100	2.40	1.8	0.75	0.300	0.014	0.014	0.326
	1.2	50	2.86	1.0	0.35	0.335	0.019	0.014	0.405
		70	2.84	1.0	0.35	0.335	0.019	0.014	0.190
		100	2.81	1.0	0.36	0.335	0.019	0.014	0.260
	1.4	50	3.18	3.0	0.94	0.358	0.026	0.014	0.636
		70	3.55	3.0	0.85	0.358	0.026	0.015	0.380
		100	3.45	3.0	0.87	0.358	0.026	0.015	0.329
Field	83.33	200	0.004	0.69	198	0.358	0.17	0.01	n/a
		400	0.005	0.69	148	0.358	0.17	0.01	n/a
		600	0.005	0.69	148	0.358	0.17	0.01	n/a

Figures 5 and 6 show the fitted parameters versus the flow rate and travel distance in the laboratory experiments. The best-fit dispersivity α does not change significantly with the travel distance (Figure 5), which is different from the scale-dependent dispersion observed by other studies [38].

This discrepancy might be due to the following reasons: (1) the travel distance is too short to observe any stable trend of dispersion (tracer injection at the inlet may cause a boundary effect on transport and random deviation from the local mean velocity), and (2) the system is advection dominated (as a typical laboratory sand-column transport experiment) and therefore the estimates of D might be too uncertain to clearly show a subtle trend. In addition, the best-fit dispersivity increases with an increasing flow rate. The faster the water flows, the wider the plume becomes, requiring a larger dispersivity. This is consistent with the fitting results of the other models mentioned above.

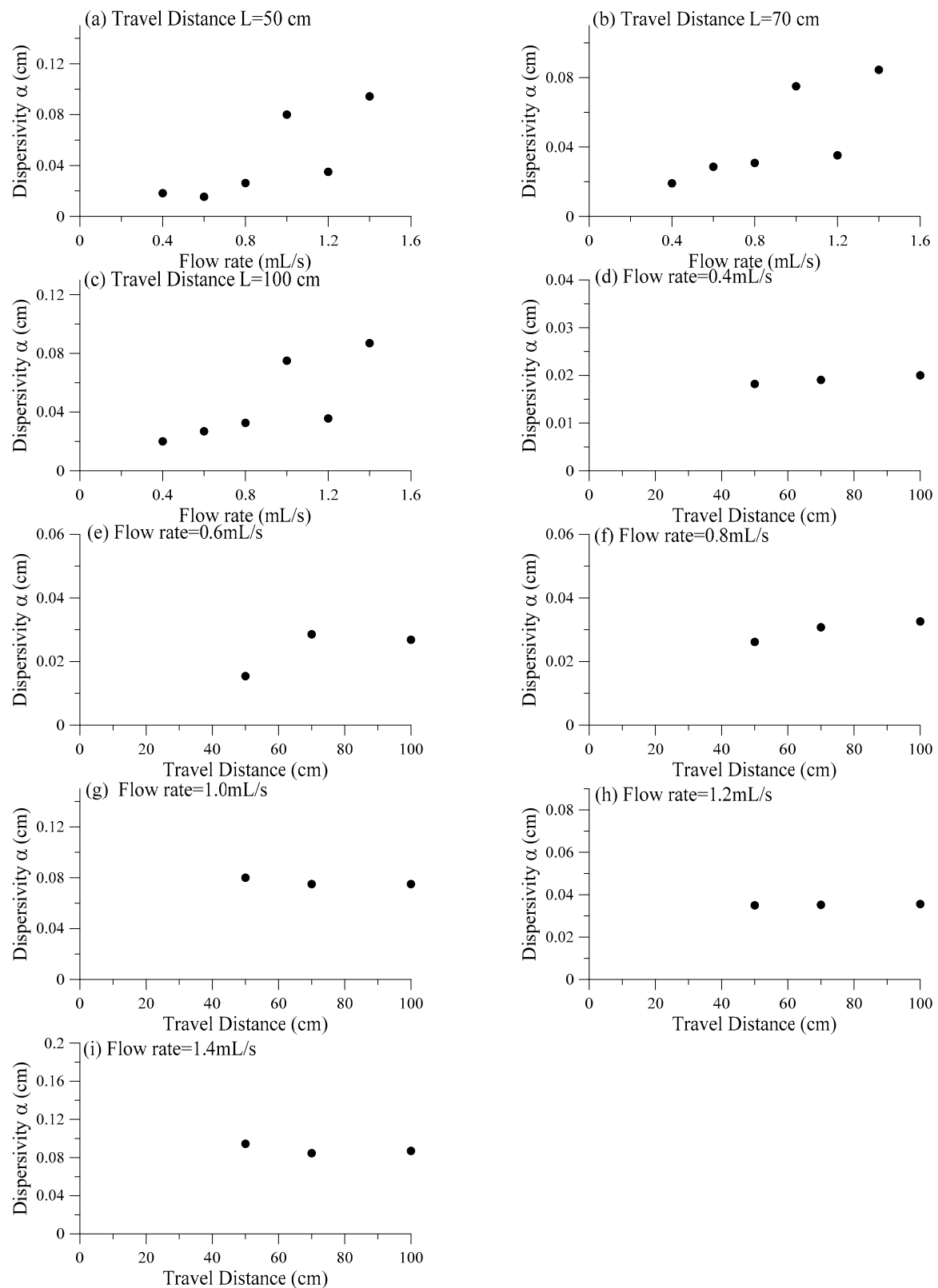


Figure 5. Dispersivity in the tt-fADE model changes with the travel distance and flow rate.

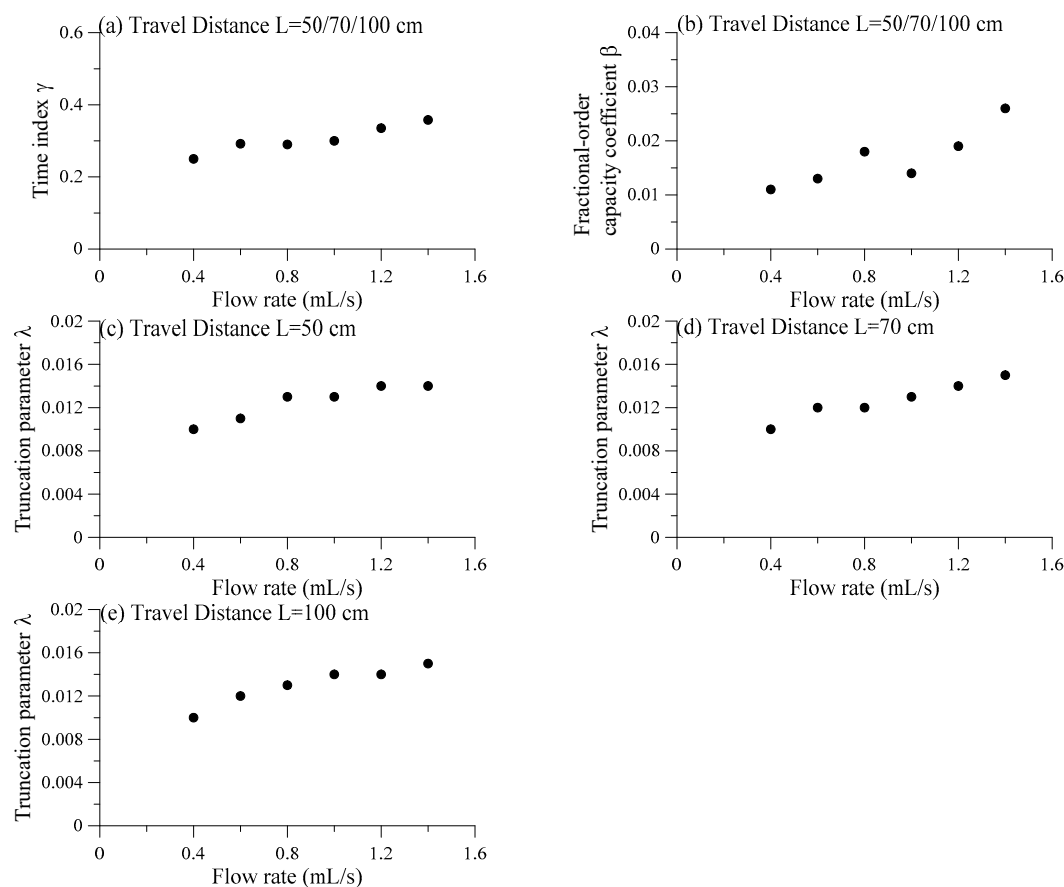


Figure 6. The tt-fADE model parameters change with the travel distance and flow rate.

The other model parameters, including the time index γ (representing residence time in the immobile domain), the capacity coefficient β (the ratio of immobile to mobile volume), and the truncation parameter λ (controlling the transition from power-law to exponential tailing), vary with the flow rate (Figure 6). First, the time index γ increases with an increasing flow rate, since a larger γ represents a shorter mean residence time for solute particles in the immobile domain (here we assume that the mean residence time in the immobile domain decreases with an increasing flow velocity). Second, the capacity coefficient β also increases with the flow rate, showing that the variation of the time index might overshadow the variation of the capacity coefficient. Third, the truncation parameter λ increases with the flow rate, implying that the non-Fickian transport converges to its Fickian asymptote more quickly if the flow rate is larger. This trend might be due to the decreased immobile portion with a higher flow rate in the short sand column. In addition, the truncation parameter λ is relatively small, in order to capture the relatively heavy late-time tails.

The above three parameters (γ , β , and λ), however, generally remain constant at different travel distances if the flow rate remains unchanged (Table 6). The packing procedure for the sand column was designed to yield a macroscopically homogeneous texture, and, indeed, the results imply that the statistics of medium properties (such as the sand size distribution, specific surface area, effective porosity, and/or internal structure) are spatially uniform. For example, a uniform fractional capacity coefficient β suggests that the ratio of mobile to immobile volume is constant in space. This phenomenon holds true for the field tracer tests. The spatial uniformity of these parameters simplifies the fitting procedure. For each flow rate with different sample distances, the time index and the capacity coefficient can be calibrated a single time for one sampling distance, and then kept constant for the other sampling distances.

Therefore, when using the tt-fADE model (10) and (11), we only need to fit three parameters: velocity, dispersion coefficient, and truncation parameter. Fitting is further simplified by using the observed BTC peak velocity as a lower bound when predicting the velocity in the tt-fADE model.

4.4. The CTRW Model

The fitted results using the CTRW model (17) are shown in Figure 2 and Figures S1–S5. We used the CTRW MATLAB Toolbox Version 3.1 developed by [39]. The CTRW framework with the truncated power-law transition time $\phi(t)$ (22) was used, as suggested by [21], since it can efficiently capture the transition from non-Fickian to Fickian transport [32].

This CTRW model contains five unknown parameters, ξ , t_1 , t_2 , v_ψ , and D_ψ . The time t_1 expressed by formula (21) represents the approximated mean transition time. The power law behavior in the BTC begins from t_1 and ends at t_2 . Knowledge of parameters gained in the fitting exercise of the tt-fADE model in Section 4.3 improves the predictability of the CTRW model. In particular, we set the power-law exponent ξ in the CTRW model equal to the time index γ in the tt-fADE model (10) and (11), and approximate the cutoff time scale t_2 in the CTRW framework using the inverse of the truncation parameter λ in the tt-fADE (see Section 2.3). Here a smaller ξ represents more disorder of the host system. The remaining three parameters, including the velocity v_ψ , the dispersion coefficient D_ψ , and the mean waiting time t_1 in (17), can be fitted using the observed BTCs. The fitted parameters using the CTRW model are listed in Table 7.

Table 7. The best-fit parameters for the continuous time random walk (CTRW) model. In the legend, v_ψ is the CTRW transport velocity, which can be different from the average pore velocity v ; D_ψ is the dispersion coefficient with the subscript ψ indicating CTRW interpretation; ξ is the power-law exponent; t_1 is the mean transition time; and t_2 is the truncation time scale. ξ is converted from γ in the tt-fADE model (10) and (11), and t_2 is converted from λ in the tt-fADE model.

Experiment	Q	L	v_ψ	D_ψ	ξ	$\log_{10} t_1$ (s)	$\log_{10} t_2$ (s)	RMSE
	(mL/s)	(cm)	(mm/s)	(mm ² /s)				
Lab	0.4	Measured	Fitted	Fitted	Fixed	Fitted	Fixed	Calculated
		50	0.465	0.38	0.250	1.70	2.00	0.239
		70	0.637	0.69	0.250	1.70	2.00	0.447
	0.6	100	0.900	14.00	0.250	1.70	2.00	0.862
		50	0.505	0.50	0.292	1.70	1.94	0.932
		70	0.735	1.47	0.292	1.70	1.93	0.859
	0.8	100	1.130	2.00	0.292	1.70	1.93	1.201
		50	0.575	0.75	0.290	1.80	1.90	1.317
		70	0.840	3.43	0.290	1.80	1.92	0.491
	1	100	1.150	4.00	0.290	1.80	1.90	0.976
		50	0.660	1.00	0.300	1.80	1.90	2.039
		70	0.994	2.94	0.300	1.81	1.89	0.780
	1.2	100	1.410	4.00	0.300	1.81	1.86	0.265
		50	0.760	1.75	0.335	1.81	1.84	0.972
		70	1.050	3.43	0.335	1.81	1.84	0.457
	1.4	100	1.500	4.00	0.335	1.81	1.84	0.913
		50	0.820	2.75	0.358	1.80	1.84	1.237
		70	1.288	5.88	0.358	1.80	1.83	0.713
		100	1.800	12.00	0.358	1.80	1.82	0.439
Field	83.33	200	0.30	0.0300	0.358	1	2	n/a
		400	0.18	0.0090	0.358	1	2	n/a
		600	0.12	0.0025	0.358	1	2	n/a

5. Discussion

5.1. Comparison of Transport Models

The transport models discussed in Section 4 can be classified into two main groups. The first group includes the ADE with equilibrium adsorption, the single mass-transfer rate MIM model, and the single diffusion rate MRMT model, which capture the observed early-time BTC and its peak,

but underestimate the persistent late-time tail of each BTC. The second group includes the MRMT model with two or a power-law distributed rate coefficients, the tt-fADE, and the CTRW model with a truncated power-law memory function, which can capture the overall trend and positive skewness of the BTCs. The two rate coefficients in the MRMT model, however, might not be adequate to capture heavy late-time tailing for solute transport observed in the field. There is no efficient way to predict the capacity coefficient and the mass transfer rate in the MRMT model, especially when there are multiple sets of mass transfer rates due to various retention capabilities in natural soils with spatial variable hydraulic properties. In contrast, the tt-fADE and the CTRW model parameters are more directly relatable to physical features [29].

MRMT model: As articulated by [24], the MRMT model relates closely to the CTRW framework in capturing solute retention times. Particularly, the mean transition time t_1 in the CTRW framework is related to the inverse of the mean rate coefficient in the MRMT model with power-law distributed rate coefficients. Our fitting exercise in Section 4.4 shows that the BTC is not sensitive to α_{max} . The mean transition time t_1 in the CTRW framework and the maximum boundary α_{max} used in the MRMT model might not be needed when capturing the late-time tailing of solute transport (note that the cutoff time t_2 in the CTRW model or the minimum rate coefficient α_{min} in the MRMT model play a more important role than t_1 or α_{max} in affecting the late-time BTC), and therefore they may be removed from the fitting parameters to simplify the model applications. Note that the tt-fADE model does not need the lower-bound of retention times. In addition, the MRMT model with power-law distributed rate coefficients tends to slightly overestimate the late-time tailing in BTCs (Figure 2 and Figures S6–S10), implying that the actual mass transfer rates may decline faster than a power-law function at late times.

CTRW model: The CTRW framework has a complex relationship to the tt-fADE model. On one hand, as discussed in Section 2.3 and checked in Section 4.4, the power-law exponent ζ in the CTRW framework is functionally equivalent to the scale index γ in the tt-fADE, and the cutoff time scale t_2 in CTRW is equivalent to the inverse of the truncation parameter λ in the tt-fADE. On the other hand, the velocity and dispersion coefficient in the CTRW framework significantly differ from those in the tt-fADE model. Similarly, they are not directly related to the solution of the traditional ADE. Applications in Section 4.4 show that the average CTRW transport velocity v_ψ (1.303 mm/s) is ~59% less than the average real peak velocity (3.18 mm/s) of the observed BTC and ~62% less than the average best-fit velocity (3.39 mm/s) in the tt-fADE model at the flow rate $Q = 1.4$ mL/s. For the field tracer tests, the CTRW transport velocity v_ψ (0.3 mm/s) is two orders of magnitude larger than the real peak velocity (0.003 mm/s) of the observed BTC, which is similar to the best-fit velocity (0.004 mm/s) in the tt-fADE model. The velocity used in the tt-fADE model can be calculated by using Equation (38), instead of fitting. This procedure further eliminates the number of parameters in the tt-fADE model and makes the fitting more convenient. In addition, the dispersion coefficient in the CTRW framework D_ψ cannot be kept constant under the same flow rate for different travel distances like the dispersion coefficient in the tt-fADE model for both laboratory experiments and field tracer tests. Therefore, the spatially averaged velocity defined by Formula (19) for the CTRW framework may differ from the actual pore-scale velocity. In the CTRW model, different from the traditional ADE, the advective, dispersive, and diffusive transport mechanisms are combined in the random walk formalism. The advective component and the dispersive component are calculated by spatial moments of the same joint probability density function (PDF) for particle transitions and hence cannot be disconnected [21]. In particular, according to Equation (19), velocity in the CTRW model can be estimated by determining the characteristic time and mean distance, which is the first moment of the PDF of transition displacement. It is, however, difficult to predict the effective velocity v_ψ without detailed knowledge of the porous medium. It is also noteworthy that the generalized master Equation (21) does not separate the effects of the spatially varying velocity field on solute particle displacement into an advective part and a dispersive part. The concept of CTRW therefore does not build an explicit relationship between real velocity and model velocity, and the same is true for the dispersion coefficient. In addition, the power-law exponent ζ can also affect the overall magnitude of

solute plume expansion, an effect that intermingles with the dispersion coefficient D_ψ and the effective velocity v_ψ . The above analysis is consistent with the result in [40] that many parameters, particularly v_ψ and ξ , in the CTRW model are correlated with each other. Their fitting exercises showed that different parameter combinations can lead to the same mean Eulerian velocity predicted by the model, implying that the estimated model parameters were not unique.

tt-fADE model: In the tt-fADE model, the best-fit velocity v can be larger than the plume peak's velocity v_{peak} because the effective velocity used in the tt-fADE is adjusted by the elapsed time for solutes spent in retention, as represented by the fractional-order capacity coefficient β on the right-hand side of Equation (38). In this study, laboratory experiments with six flow rates and field tracer tests show that the best-fit velocity in the tt-fADE model is close to the measured BTC's peak velocity, since the capacity coefficient shown in Equation (38) is relatively small. Because the transport velocity used in the tt-fADE model (10) and (11) is not significantly different from the BTC's peak velocity, the tt-fADE model uses an independent parameter, the time index γ , to control the power-law distribution of the late-time BTC. This parameterization of the tailing is relatively simple as compared to the CTRW model, with three parameters, v_ψ , D_ψ , and ξ that contribute to time-nonlocal non-Fickian dispersion. We calculate the root mean square error (RMSE) for the laboratory experiments. Comparison of the RMSE (see Tables 1, 6 and 7) of the two models and the ADE model demonstrates that they both perform better than the standard ADE, especially in describing tailing behavior in a heterogeneous medium. When the Formula (38) is used, the tt-fADE model is more convenient to apply in practice than the CTRW framework, since the former contains fewer parameters.

Why the fitted four-parameter tt-fADE model may provide the best performance? This might be due to two reasons. First, according to the physical derivation of the tt-fADE in Section 2.2, the tt-fADE separates motion in the mobile zone (using the basic transport parameters v and D to define advection and Gaussian diffusive displacement) and retention in the immobile domains. When the tracer particle moves in the mobile domain, its average speed is v , with a finite time required to finish the jump. This physical separation might be reasonable in hydrogeological media. In the CTRW framework, the memory function defines the random waiting time between two subsequent jumps, while the motion can finish instantaneously (in other words, no physical time is required for the tracer particle to move). This physical discrepancy may also cause the discrepancy of the two basic transport parameters v and D between the two models, in addition to the spatial average parameters required by the CTRW framework. Second, the tt-fADE does not specify the lower-bound of the waiting time (using an additional parameter such as the lower-limit t_1 in the CTRW framework), since the slow advection can also affect the late-time BTC tail. Mass exchange can apparently affect the late-time BTC tail only when the diffusive time scale is much longer than the advective time scale, as pointed out by [16]. This may explain why t_1 might not be needed in the CTRW framework. Further real-world tests are needed to check the above hypotheses.

5.2. Parameter Sensitivity

One example of parameter sensitivity is tested here for the tt-fADE (10) and (11). Sensitivity of the BTCs to variations of the four main parameters (D , γ , β , λ) in the tt-fADE model is shown in Figure 7.

First, the dispersion coefficient D has a subtle impact on the overall shape of BTCs. Decreasing D from $0.005 \text{ cm}^2/\text{s}$ to $0.0005 \text{ cm}^2/\text{s}$ results in similar BTCs after normalization (i.e., re-scaling), implying that trapping due to the immobile zones may account at least partially for the spatial expansion of solute plumes. When D increases from $0.005 \text{ cm}^2/\text{s}$ to $0.05 \text{ cm}^2/\text{s}$, the BTC becomes wider and its shape slightly changes (Figure 7a,b).

Second, the time index γ controls the power-law slope of the late-time BTC. For example, when γ increases from 0.29 to 0.50 (representing the decrease of probability for long retention times), the late-time BTC becomes steeper, approaching relatively fast to its Gaussian asymptote (Figure 7c,d). When γ decreases from 0.29 to 0.05, the BTC's late-time tail becomes heavier (i.e., with a gentler slope),

although the overall BTC looks narrower (likely due to the normalization of BTCs). The simulated BTC with the time index γ equal to 0.29 gives the best fit.

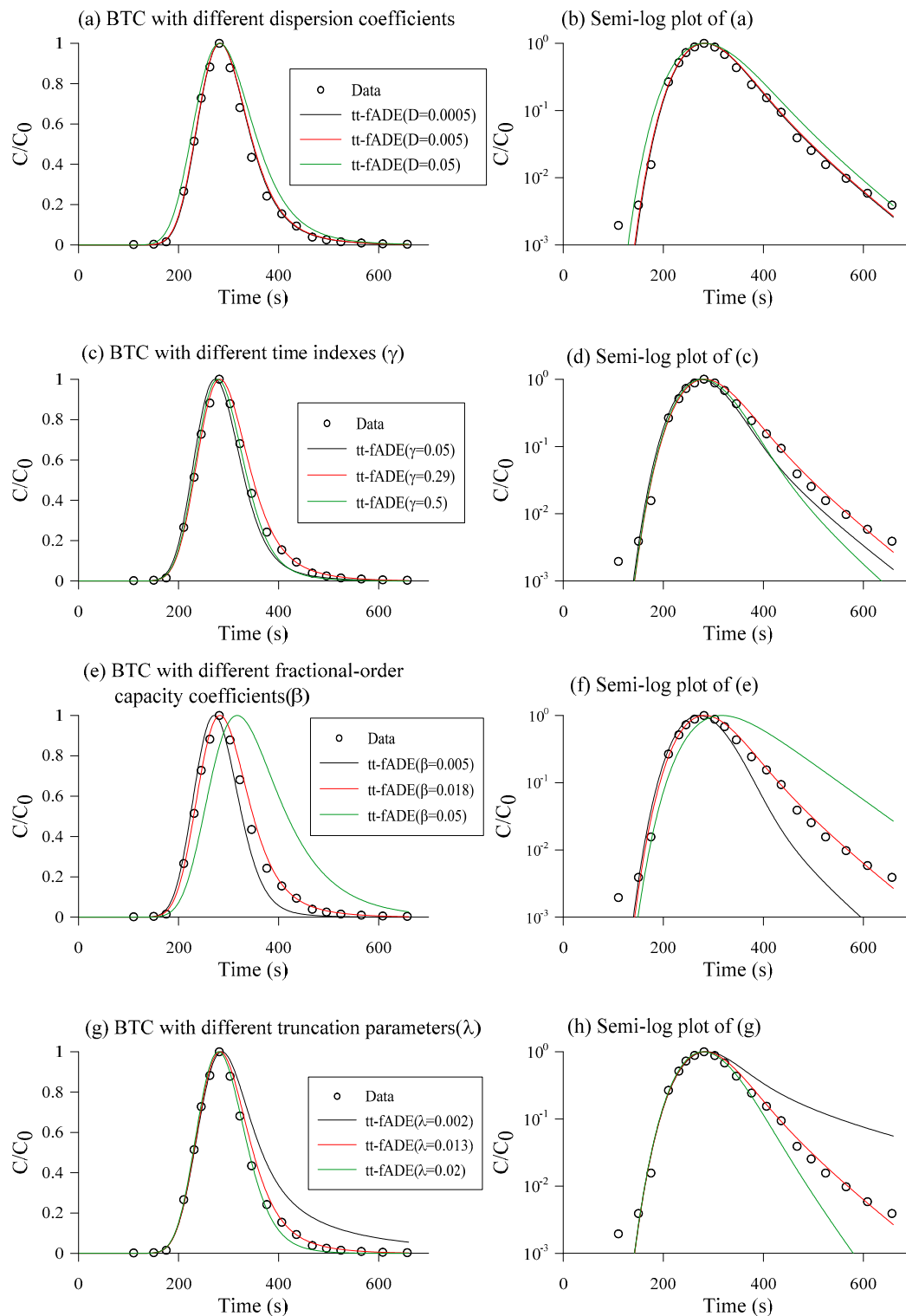


Figure 7. Sensitivity analysis for coefficients in the tt-fADE model.

Third, the fractional-order capacity coefficient β shifts the BTC and expands the BTC's late-time tail. When β decreases (representing a decrease of the immobile zone volume or the immobile solute

mass at equilibrium), the BTC becomes narrower and shifts to the left (representing a larger effective velocity). A faster drop is apparent in the late-time BTC tail with a smaller β . An opposite change of the BTC can be seen for an increasing β (Figure 7e,f).

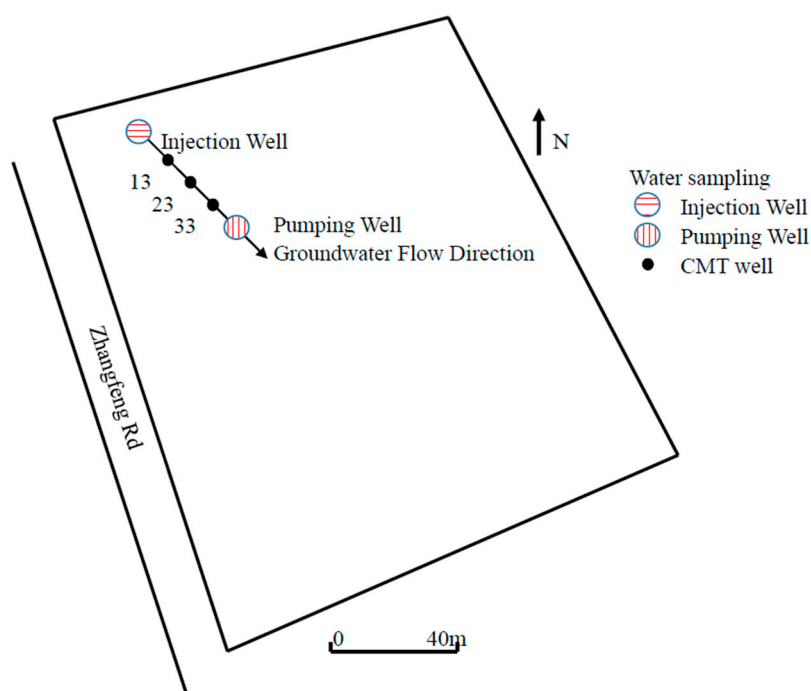
Fourth, the truncation parameter λ affects the speed for the late-time BTC to transfer from a heavy tailed, power-law slope to an exponential tail. A larger truncation parameter means an earlier transform from non-Fickian to Fickian transport (Figure 7g,h). For the largest λ tested, the resultant BTC is the closest to the solution of the ADE model, as expected because the tt-fADE reduces to the ADE model for $\lambda \rightarrow \infty$.

5.3. Application to Field Transport

A field trace transport test was conducted recently by Zheng [41], which provides field data to evaluate further the time nonlocal transport models and compare with the laboratory column experiments. The test site is in the Zhangjiawan Village, Zhangjiawan Town, southeast of the Tongzhou District, Beijing, China, with the longitude of 116.72° and latitude 39.848° . This experimental site is in the Chaobai River alluvial plain. The average annual precipitation in the vicinity is about 533 mm, and the evaporation is 1822 mm. It has a multi-layer aquifer structure. The aquifer is composed of gravelly coarse sand, coarse sand, fine sand, silty clay and clay layer in the test field and surrounding area. The hydraulic conductivity coefficient shows obvious heterogeneity, indicating that the aquifer develops a small-scale preferential flow channel network.

The subsurface network consisted of one injection well, one pumping well, and three monitoring wells with continuous multi-tubing (denoted as well 13, 23, and 33, respectively) (see Figure 8 for the study site).

All the five wells are along the same line of the general groundwater flow direction, and therefore a one-dimensional model may be used to approximate the overall transport. The injection well and the pumping well are separated by 8 m, and the three observation wells have a uniform interval of 2 m. Groundwater flows from the injection well to well 13, 23, 33, and to the pumping well.



(a) Zhangjiawan test site location and distribution of wells

Figure 8. Cont.

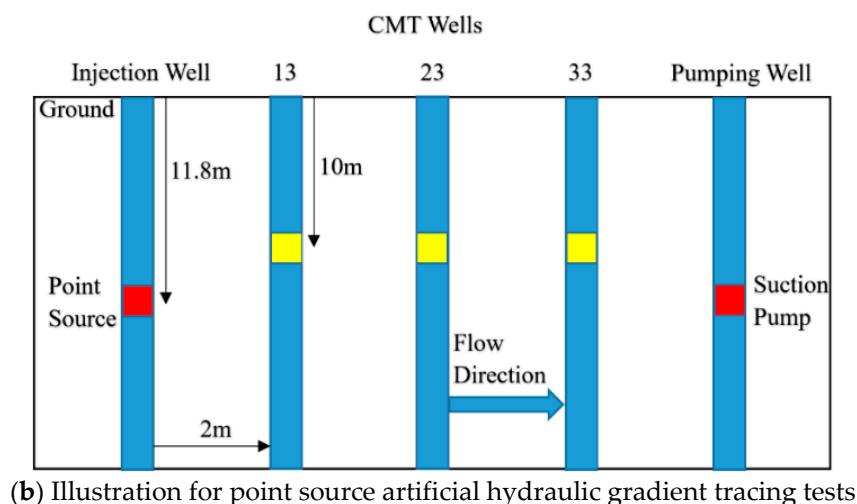


Figure 8. Study site map showing injection well, pumping well, and continuous multichannel tubing (CMT) wells.

Sodium bromide, a commonly used conservative tracer, was injected into the injection well at a depth of 11.8 m, and groundwater samples were taken from the three observation wells (Well 13, 23, 33) located downstream at a depth of ~10 m. The concentration of the injected solution was 1288 mg/L. The injection rate was about 0.3 m³/h, and the injection duration was 6 h. A peristaltic pump was used to collect the samples in a chronological order, and the sample concentration was measured using a MP523-06 bromide ion concentration meter.

The measured BTCs exhibit apparent late-time tailings (see Figure 9), similar to that observed in the laboratory column transport. Applications show that the time nonlocal models can capture part of the late-time BTC tail, but not the whole tail of BTC containing apparent noise. Due to the noise, it is also impossible to obtain a reliable RMSE. Although the apparent noise causes high uncertainty in model fitting, both the CTRW framework and the tt-fADE can capture a heavier late-time tail than the MRMT model with two sets of rate coefficients, and the measured BTCs do contain a high concentration at the very end of the sampling period (Figure 9).

In addition, all the measured BTCs show apparent early time tail, which cannot be captured by the tt-fADE or the CTRW framework with the time index between 0 and 1. It is, however, not a surprise, since the early arrivals are most likely due to fast motion of tracer particles along preferential flow paths, while the delayed arrivals are caused by solute retention due to mass exchange between the mobile and relatively immobile zones. At the field site, high-permeable sand constitutes the layer connecting the injection well and the three monitoring wells, likely forming the preferential channels. The time nonlocal transport models considered in this study were developed to capture solute retention, hence missing the early tail of the BTC. The spatiotemporal fADE may capture both the early and late time tails in the BTC [28,42], which will be explored in a future study.

It is also noteworthy that the flow velocity in real aquifers is several orders of magnitude smaller than that used in the laboratory experiments. Hence, the field transport is diffusion dominated, while the laboratory transport is advection dominated. This discrepancy might imply that the late-time BTC tail persists for groundwater flow with a broad range of Peclet numbers, which can be characterized by the time nonlocal models. However, for groundwater flow with a small Peclet number and potential preferential flow paths, the early time BTC tail may occur, which cannot be efficiently captured by the time nonlocal transport models such as the tt-fADE or the CTRW framework with an index less than 1.

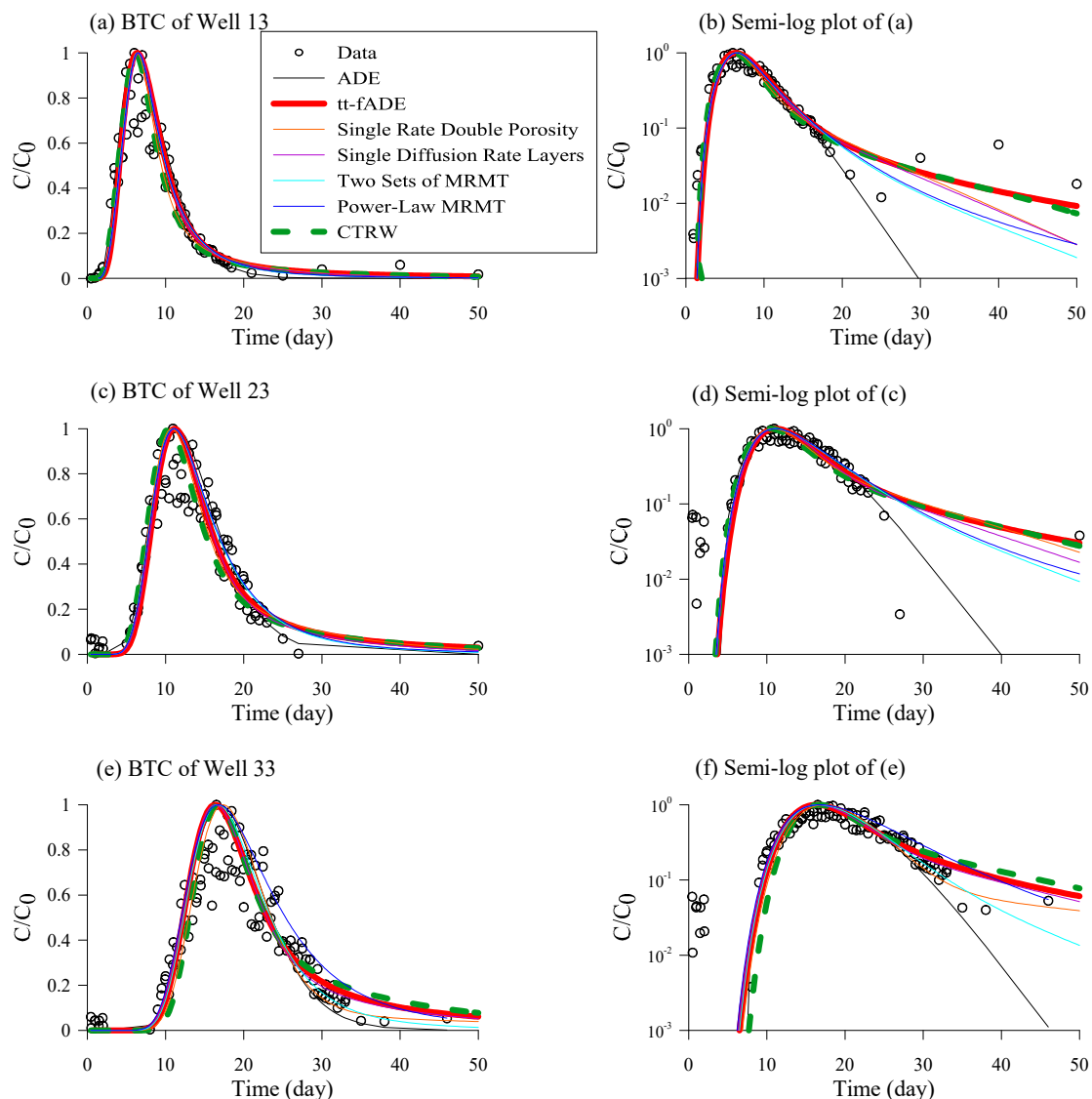


Figure 9. Field tracer test: comparison between the measured field data (symbols) and the modeled (lines) breakthrough curves using the ADE, the tt-fADE (red thick lines), the CTRW (green dashed lines), and the MRMT models.

6. Conclusions

This study compared three time-nonlocal transport models by combining theoretical analyses and applications for laboratory sand column transport experiments and field tracer tests. The models revisited by this study include the MRMT model with various specific forms, the tt-fADE model, and the CTRW framework. Four major conclusions were obtained for these models, which can be used by practitioners to select the appropriate model and improve practical applications, and can improve our understanding of the nature of non-Fickian transport in heterogeneous media.

First, the sand column packed in the laboratory and soil in the field may contain multiple immobile domains with different mass transfer capabilities. This assumption and the laboratory measurements challenge the applicability of the classical ADE model with equilibrium adsorption (i.e., instantaneous sorption/desorption) and the single-rate mobile-immobile model in capturing the tracer BTCs with a late-time tail, which declines at a rate slower than exponential. The MRMT models with multiple rates do capture the BTC's late time tail typical for non-Fickian transport, as revealed before, but the MRMT solutions with power-law distributed mass-exchange rates cannot capture the nuance of observed

transition from power-law to exponential decline of the late-time concentration. The increase of the number of unpredictable parameters (both the rate coefficient and the capacity coefficient) may create a challenge for the applicability of the MRMT model.

Second, the tt-fADE model and the CTRW framework are similar in functionality, but differ in detailed parameters. (1) Both the CTRW framework and the tt-fADE model can capture a complex BTC with late time tailing. Both the CTRW framework and the tt-fADE model assume an exponentially-truncated power-law memory function, to capture the gradual transition from power-law to exponential decline of late-time concentration in the observed BTCs. (2) The tt-fADE parameters can be linked to the CTRW framework parameters. For example, the power-law exponent ξ in the CTRW framework is functionally equivalent to the scale index γ in the tt-fADE model (as revealed before), and the cutoff time scale t_2 in the CTRW framework is also equivalent to the inverse of the truncation parameter λ in the tt-fADE (not shown specifically before). Hence the predictability obtained by the tt-fADE model can also improve the predictability of the CTRW framework, and vice-versa. (3) Compared to the tt-fADE model, the CTRW framework defines one additional parameter t_1 , which represents the mean diffusive time, corresponding to the mean of the inverse of rate coefficients in the MRMT model. Model applications, however, showed that t_1 in the CTRW framework is insensitive to model results, and may be neglected to alleviate model fitting burdens.

Third, in the tt-fADE model, the real BTC's peak velocity can be used to estimate the lower-end of the model velocity, increasing the predictability of the tt-fADE for real-world applications. Hence, the tt-fADE model with less parameters may conveniently and accurately estimate the BTC late-time tailing under the conditions of the column experiments and field tracer tests.

Fourth, for tracer transport in the field, early arrivals are likely due to preferential flow paths. Super-diffusive jumps along preferential flow paths cannot be efficiently captured by a typical time-nonlocal transport model focusing on solute retention with a time index less than one. Fast motion (which can exhibit direction-dependent scaling rates) and delayed transport (which is dimensionless), although co-existing in some field sites, are driven by different mechanisms, and hence we recommend different physical components to capture these processes. This motivates the application of the spatiotemporal fADE [28,43], which will be re-visited in a future study.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4441/10/6/778/s1>, Figure S1: Comparison between the measured (symbols) and modeled (lines) breakthrough curves (BTCs) using the ADE model (black line), the tt-fADE model (red line) and the CTRW model (green line) with the experimental water flow rate $Q = 0.4$ mL/s. Figure S2: Comparison between the measured (symbols) and modeled (lines) breakthrough curves (BTCs) using the ADE model (black line), the tt-fADE model (red line) and the CTRW model (green line) with the experimental water flow rate $Q = 0.6$ mL/s. Figure S3: Comparison between the measured (symbols) and modeled (lines) breakthrough curves (BTCs) using the ADE model (black line), the tt-fADE model (red line) and the CTRW model (green line) with the experimental water flow rate $Q = 0.8$ mL/s. Figure S4: Comparison between the measured (symbols) and modeled (lines) breakthrough curves (BTCs) using the ADE model (black line), the tt-fADE model (red line) and the CTRW model (green line) with the experimental water flow rate $Q = 1.0$ mL/s. Figure S5: Comparison between the measured (symbols) and modeled (lines) breakthrough curves (BTCs) using the ADE model (black line), the tt-fADE model (red line) and the CTRW model (green line) with the experimental water flow rate $Q = 1.2$ mL/s. Figure S6: Comparison between the measured (symbols) and the modeled (lines) breakthrough curves (BTCs) using the ADE, the tt-fADE, the MRMT, and the CTRW models with the water flow rate $Q = 0.4$ mL/s. Figure S7: Comparison between the measured (symbols) and the modeled (lines) breakthrough curves (BTCs) using the ADE, the tt-fADE, the MRMT, and the CTRW models with the water flow rate $Q = 0.6$ mL/s. Figure S8: Comparison between the measured (symbols) and the modeled (lines) breakthrough curves (BTCs) using the ADE, the tt-fADE, the MRMT, and the CTRW models with the water flow rate $Q = 0.8$ mL/s. Figure S9: Comparison between the measured (symbols) and the modeled (lines) breakthrough curves (BTCs) using the ADE, the tt-fADE, the MRMT, and the CTRW models with the water flow rate $Q = 1.0$ mL/s. Figure S10: Comparison between the measured (symbols) and the modeled (lines) breakthrough curves (BTCs) using the ADE, the tt-fADE, the MRMT, and the CTRW models with the water flow rate $Q = 1.2$ mL/s.

Author Contributions: Conceptualization, B.L. and Y.Z.; Methodology, B.L. and Y.Z.; Software, B.L. and Y.Z.; Validation, C.Z., C.T.G. and C.O.; Formal Analysis, B.L.; Investigation, Y.Z.; Resources, H.-G.S.; Data Curation, J.Q.; Writing-Original Draft Preparation, B.L. and Y.Z.; Writing Review & Editing, Y.Z.; Visualization, B.L.; Supervision, Y.Z.; Project Administration, Y.Z.; Funding Acquisition, H.-G.S. and Y.Z.

Funding: This research was partially funded by the National Natural Science Foundation of China (under grants 41330632, 41628202, and 11572112), the University of Alabama, and the US Geological Survey. This paper does not necessarily reflect the views of the National Natural Science Foundation of China or the University of Alabama.

Conflicts of Interest: The authors declare no conflict of interest. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

1. Anderson, M.P. Movement of contaminants in groundwater: Groundwater transport—Advection and dispersion. *Groundw. Contam.* **1984**, *20*, 37–45.
2. Bakshevskaia, V.A.; Pozdniakov, S.P. Simulation of hydraulic heterogeneity and upscaling permeability and dispersivity in sandy-clay formations. *Math. Geosci.* **2016**, *48*, 45–64. [[CrossRef](#)]
3. Chang, C.-M.; Yeh, H.-D. Investigation of flow and solute transport at the field scale through heterogeneous deformable porous media. *J. Hydrol.* **2016**, *540*, 142–147. [[CrossRef](#)]
4. Cvetkovic, V.; Fiori, A.; Dagan, G. Tracer travel and residence time distributions in highly heterogeneous aquifers: Coupled effect of flow variability and mass transfer. *J. Hydrol.* **2016**, *543*, 101–108. [[CrossRef](#)]
5. Bromly, M.; Hinz, C. Non-fickian transport in homogeneous unsaturated repacked sand. *Water Resour. Res.* **2004**, *40*, W07402. [[CrossRef](#)]
6. Huber, F.; Enzmann, F.; Wenka, A.; Bouby, M.; Dentz, M.; Schafer, T. Natural micro-scale heterogeneity induced solute and nanoparticle retardation in fractured crystalline rock. *J. Contam. Hydrol.* **2012**, *133*, 40–52. [[CrossRef](#)] [[PubMed](#)]
7. Naftaly, A.; Dror, I.; Berkowitz, B. Measurement and modeling of engineered nanoparticle transport and aging dynamics in a reactive porous medium. *Water Resour. Res.* **2016**, *52*, 5473–5491. [[CrossRef](#)]
8. Nordin, C.F.; Troutman, B.M. Longitudinal dispersion in rivers: The persistence of skewness in observed data. *Water Resour. Res.* **1980**, *16*, 123–128. [[CrossRef](#)]
9. Zaramella, M.; Marion, A.; Lewandowski, J.; Nützmann, G. Assessment of transient storage exchange and advection–dispersion mechanisms from concentration signatures along breakthrough curves. *J. Hydrol.* **2016**, *538*, 794–801. [[CrossRef](#)]
10. Bianchi, M.; Zheng, C. A lithofacies approach for modeling non-fickian solute transport in a heterogeneous alluvial aquifer. *Water Resour. Res.* **2016**, *52*, 552–565. [[CrossRef](#)]
11. Ziemys, A.; Kojic, M.; Milosevic, M.; Ferrari, M. Interfacial effects on nanoconfined diffusive mass transport regimes. *Phys. Rev. Lett.* **2012**, *108*, 236102. [[CrossRef](#)] [[PubMed](#)]
12. Karadimitriou, N.K.; Joekar-Niasar, V.; Babaei, M.; Shore, C.A. Critical role of the immobile zone in non-fickian two-phase transport: A new paradigm. *Environ. Sci. Technol.* **2016**, *50*, 4384–4392. [[CrossRef](#)] [[PubMed](#)]
13. Van Genuchten, M.T.; Wierenga, P.J. Mass transfer studies in sorbing porous media I. Analytical solutions. *Soil Sci. Soc. Am. J.* **1976**, *40*, 473–480. [[CrossRef](#)]
14. Sardin, M.; Schweich, D.; Leij, F.J.; van Genuchten, M.T. Modeling the nonequilibrium transport of linearly interacting solutes in porous media: A review. *Water Resour. Res.* **1991**, *27*, 2287–2307. [[CrossRef](#)]
15. Cvetkovic, V.; Dagan, G. Transport of kinetically sorbing solute by steady random velocity in heterogeneous porous formations. *J. Fluid Mech.* **1994**, *265*, 189–215. [[CrossRef](#)]
16. Haggerty, R.; McKenna, S.A.; Meigs, L.C. On the late-time behavior of tracer test breakthrough curves. *Water Resour. Res.* **2000**, *36*, 3467–3479. [[CrossRef](#)]
17. Valocchi, A.J. Validity of the local equilibrium assumption for modeling sorbing solute transport through homogeneous soils. *Water Resour. Res.* **1985**, *21*, 808–820. [[CrossRef](#)]
18. Jury, W.A.; Horton, R. *Soil Physics*; John Wiley: Hoboken, NJ, USA, 2004; 370p.
19. Coats, K.K.H.; Smith, B.B.D.; van Genuchten, M.T.; Wierenga, P.J. Dead-end pore volume and dispersion in porous media. *Soil Sci. Soc. Am. J.* **1964**, *4*, 73–84. [[CrossRef](#)]
20. Haggerty, R.; Gorelick, S.M. Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity. *Water Resour. Res.* **1995**, *31*, 2383–2400. [[CrossRef](#)]
21. Berkowitz, B.; Cortis, A.; Dentz, M.; Scher, H. Modeling non-fickian transport in geological formations as a continuous time random walk. *Rev. Geophys.* **2006**, *44*. [[CrossRef](#)]

22. Meerschaert, M.M.; Scheffler, H.-P. *Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice*; Wiley: Hoboken, NJ, USA, 2001; Volume 321.
23. Meerschaert, M.M.; Zhang, Y.; Baeumer, B. Tempered anomalous diffusion in heterogeneous systems. *Geophys. Res. Lett.* **2008**, *35*, L17403. [[CrossRef](#)]
24. Dentz, M.; Berkowitz, B. Transport behavior of a passive solute in continuous time random walks and multirate mass transfer. *Water Resour. Res.* **2003**, *39*, 1111. [[CrossRef](#)]
25. Benson, D.A.; Meerschaert, M.M. A simple and efficient random walk solution of multi-rate mobile/immobile mass transport equations. *Adv. Water Resour.* **2009**, *32*, 532–539. [[CrossRef](#)]
26. Schumer, R.; Benson, D.; Meerschaert, M.; Baeumer, B. Fractal mobile/immobile solute transport. *Water Resour. Res.* **2003**, *39*. [[CrossRef](#)]
27. Zhang, Y.; Benson, D.A.; Baeumer, B. Predicting the tails of breakthrough curves in regional-scale alluvial systems. *Ground Water* **2007**, *45*, 473–484. [[CrossRef](#)] [[PubMed](#)]
28. Schumer, R. Multiscaling fractional advection–dispersion equations and their solutions. *Water Resour. Res.* **2003**, *39*, 1–11. [[CrossRef](#)]
29. Zhang, Y.; Green, C.T.; Baeumer, B. Linking aquifer spatial properties and non-fickian transport in mobile–immobile like alluvial settings. *J. Hydrol.* **2014**, *512*, 315–331. [[CrossRef](#)]
30. Klafter, J.; Silbey, R. Derivation of the continuous-time random-walk equation. *Phys. Rev. Lett.* **1980**, *44*, 55–58. [[CrossRef](#)]
31. Montroll, E.; Weiss, G. Random walks on lattices. II. *J. Math. Phys.* **1965**, *6*, 167. [[CrossRef](#)]
32. Dentz, M.; Cortis, A.; Scher, H.; Berkowitz, B. Time behavior of solute transport in heterogeneous media: Transition from anomalous to normal transport. *Adv. Water Resour.* **2004**, *27*, 155–173. [[CrossRef](#)]
33. Willmann, M.; Carrera, J.; Sánchez-Vila, X. Transport upscaling in heterogeneous aquifers: What physical parameters control memory functions? *Water Resour. Res.* **2008**, *44*. [[CrossRef](#)]
34. Van Genuchten, M.T.; Alves, W. *Analytical Solutions of the One-Dimensional Convective-Dispersive Solute Transport Equation*; Technical Bulletin; United States Department of Agriculture: Washington, DC, USA, 1982.
35. Haggerty, R. STAMMT-L 3.0. *A Solute Transport Code for Multirate Mass Transfer and Reaction along Flowlines*; Sandia National Laboratories: Albuquerque, NM, USA, 2009.
36. Fiori, A.; Becker, M.W. Power law breakthrough curve tailing in a fracture: The role of advection. *J. Hydrol.* **2015**, *525*, 706–710. [[CrossRef](#)]
37. Zhang, Y.; Meerschaert, M.M.; Baeumer, B.; LaBolle, E.M. Modeling mixed retention and early arrivals in multidimensional heterogeneous media using an explicit lagrangian scheme. *Water Resour. Res.* **2015**, *51*, 6311–6337. [[CrossRef](#)]
38. Pickens, J.F.; Grisak, G.E. Scale-dependent dispersion in a stratified granular aquifer. *Water Resour. Res.* **1981**, *17*, 1191–1211. [[CrossRef](#)]
39. Cortis, A.; Berkowitz, B. Computing “anomalous” contaminant transport in porous media: The ctrw matlab toolbox. *Ground Water* **2005**, *43*, 947–950. [[CrossRef](#)] [[PubMed](#)]
40. Fiori, A.; Zarlenga, A.; Gotovac, H.; Jankovic, I.; Volpi, E.; Cvetkovic, V.; Dagan, G. Advective transport in heterogeneous aquifers: Are proxy models predictive? *Water Resour. Res.* **2015**, *51*, 9577–9594. [[CrossRef](#)]
41. Zheng, C.; Hill, M.C.; Cao, G.; Ma, R. Mt3dms: Model use, calibration, and validation. *Trans. ASABE* **2012**, *55*, 1549–1559. [[CrossRef](#)]
42. Benson, D.A.; Wheatcraft, S.W.; Meerschaert, M.M. Application of a fractional advection–dispersion equation. *Water Resour. Res.* **2000**, *36*, 1403–1412. [[CrossRef](#)]
43. Benson, D.A.; Wheatcraft, S.W.; Meerschaert, M.M. The fractional-order governing equation of lévy motion. *Water Resour. Res.* **2000**, *36*, 1413–1423. [[CrossRef](#)]

