Adjoint subordination to calculate backward travel time probability of pollutants in water with various velocity resolutions

Yong Zhang[1], Graham E. Fogg[2], Hongguang Sun[3], Donald M. Reeves[4], Roseanna M. Neupauer[5], Wei Wei[6]

[1]Department of Geological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA
[2]Hydrologic Sciences, University of California, Davis, CA 95616, USA
[3]College of Mechanics and Materials, Hohai University, Nanjing 210098, China
[4]Department of Geological and Environmental Sciences, Western Michigan University, Kalamazoo, MI 49008, USA
[5]Department of Civil, Environmental, and Architectural Engineering, University of Colorado Boulder, Boulder, CO 80309, USA
[6]School of Environment, Nanjing Normal University, Nanjing 210023, China

Correspondence to: Yong Zhang (yzhang264@ua.edu) and HongGuang Sun (shg@hhu.edu.cn)

Abstract. Backward probabilities such as the backward travel time probability density function for pollutants in natural aquifers/rivers haved been used by hydrologists for decades in water-quality related applications. Reliable calculation ofCalculating these backward probabilities, however, has been challenged byfaces challenges is challenging due to non-Fickian pollutant transport dynamics and variability in thevelocity resolution variability of velocity at study sites. To address these two issues, we built an adjoint model by deriving a backward-in-time fractional-derivative transport equation subordinated to regional flow, developed a Lagrangian solver, and applied the model/solver to backtrack trace pollutant transport in various diverse flow systems. The adjoint model applies subordinatesion to a reversed regional flow field, converts transforms forward-in-time boundaries into either absorbing or reflective boundaries, and reverses the tempered stable density to define backward mechanical dispersion. The corresponding Lagrangian solver is computationally efficiently in projectsing backward super-diffusive mechanical dispersion along streamlines. Field applications demonstrate that the adjoint subordination model's can successfully in recovering release history, dated groundwater age, and pollutant spatial source location(s) of pollutant source(s) for various flow systems. These include systems with with either upscaled constant velocity, non-uniform divergent flow field, or fine-resolution velocities in a non-stationary, regional-scale aquifer, where non-Fickian transport significantly affects pollutant dynamics and backward probabilitiesy characteristics. Caution is needed when identifying the phase-sensitive (aqueous versus absorbed) pollutant source in natural media. The study also explores Possible possible extensions of the adjoint subordination model are also discussed and tested for quantifying backward probabilities of pollutants in more complex media, such as discrete fracture networks.

1 Introduction

Backward probabilities of pollutants in natural aquifers/rivers, such as the backward travel time probability density function (BTTP), has been used by hydrologists for decades in water quality ~~related~~ applications. For example, BTTP ~~defines the possible length of~~estimates the time ~~for~~ contaminants take to reach a sampling location (e.g., a monitoring well screen or stream sampling location) from their source(s) ~~location(s)~~ (Neupauer and Wilson, 2001; Ponprasit et al., 2023). It ~~contains~~ provides useful ~~information on contaminant fate and transport which can help~~insights for water management, remediation, and assessment. For instance, a common application of BTTP is to recover contamination history and identify responsible parties, where the BTTP's peak captures the most likely release time of contaminants from the source (Skaggs and Kabala, 1994; Woodbury and Ulrych, 1996; Woodbury et al., 1998; Sun et al., 2006a, 2006b; Jha and Datta, 2015; Yeh et al., 2015; Jamshidi et al., 2020; Chen et al., 2023). BTTP can also be used to date groundwater since BTTP characterizes the age distribution of groundwater due to borehole mixture and/or hydrodynamic dispersion in regional-scale aquifers (Weissmann et al., 2002; Cornaton and Perrochet, 2006; LaBolle et al., 2006; Zinn and Konikow, 2007a, 2007b; Janssen et al., 2008; McMahon et al., 2008; Maxwell et al., 2016; Ponprasit et al., 2022; Mao et al., 2023). In addition, BTTP provides a more comprehensive method to assess aquifer vulnerability than classical statistics-based approaches through the generation of three-dimensional (3-*d*), transient vulnerability maps for groundwater to non-point source contamination (Fogg et al., 1999; Zhang et al., 2018). BTTP can also be used to estimate solute concentration trends (Green et al., 2014), and rates of oxygen and nitrate reduction in regional groundwater settings (Green et al., 2016). These diverse applications ~~demonstrate that mathematical models can reliably be applied to quantify backward probabilities including BTTP, and most importantly,~~underscore the need for a general BTTP model, which ~~is still needed due to the challenges mentioned below, motivating~~is the focus of this study.

There are two main challenges in numerically quantifying~~ication of~~ backward probabilities, including BTTP, for contaminant transport in surface water and groundwater. Firstly, a novel model is required ~~needed~~ to ~~account for~~address the impact of complex transport dynamics of contaminants on BTTP. Previous BTTP models, ~~which are~~ usually based on ~~the~~ inverse or backward advection-dispersion equations (ADEs), assumed Fickian diffusion of contaminants, where ~~(meaning that~~ the plume variance grows linearly over~~in~~ time~~)~~; see the extensive review by Moghaddam et al. (2021). Real-world contaminant transport, however, is ~~usually~~often non-Fickian at ~~almost all relevant~~various scales, ~~where~~ exhibiting either ~~the temporal evolution of the plume variance can be either~~ slower~~-~~than~~-~~linear temporal plume variance growth (~~which is called~~known as "sub-diffusion") or faster-~~-~~than-~~-~~linear growth ("super-diffusion"), as recently reviewed by Guo et al. (2021). Particularly, super-diffusion can be driven by factors like turbulence or flooding events in streams (Phillips et al., 2013; Boano et al., 2014), preferential flow pathways consisting of fractures in fractured porous media (Reeves et al., 2008), or high-permeability paleochannels within alluvial deposits (Bianchi et al., 2016). Sub-diffusion is more common in natural water systems due to pervasive solute retention or storage mechanisms such as physical/chemical sorption-desorption, heterogeneous advection (meaning a broad range of advective velocities), and~~/or~~ multi-rate mass exchange between mobile and relative immobile flow zones (Haggerty et al, 2000; Zhou et al., 2021). Classical Fickian-diffusion ~~based classical~~ models cannot effectively capture super/sub-diffusive non-Fickian transport ~~if~~when the velocity field ~~is not~~lacks sufficient~~ly~~ resolution~~ved~~ (e.g., coarser than the centimetre scale; see Zheng et al. (2011)) or when~~if~~ the model underestimates the spatial interconnectivity of high-permeability deposits (Yin et al., 2020). To address this issue, various nonlocal transport models, which are typically non-Markovian models considering the spatiotemporal memory during solute transport, have been developed to efficiently simulate forward-in-time non-Fickian transport (Neuman and Tartakovsky, 2009).~~,~~ ~~but~~ However, their corresponding BTTP models have remained less ~~remained obscure~~explored (Zhang et al., 2022; Zhang, 2022).

The second challenge is how to ~~incorporate~~ integrate the observed velocity field, ~~the resolution of which typically~~which often varies significantly in resolution ~~between~~ across field sites, into ~~the~~ backward probabilit~~ies~~y ~~(including BTTP)~~ calculations, including BTTP. Many field sites ~~have only limited~~lack extensive hydrologic ~~information~~data, ~~requiring a fully~~necessitating an upscaled BTTP model ~~which can function using~~capable of operating with ~~a~~ coarsely resolved velocity ~~field~~ or ~~even a~~ uniform velocity fields. Contrarily, ~~some~~ well-studied ~~field~~ sites ~~may have~~with abundant geologic and ~~/~~hydrologic data should ~~, providing a~~incorporate detailed spatiotemporal ~~distribution of~~ velocity distributions to~~ies that should be incorporated into the numerical model to improve the reliability of~~ enhance BTTP calculation reliabilitys. Ideally, an efficient BTTP model should ~~be able to~~seamlessly incorporate ~~the~~ velocity fields without ~~any~~ resolution constraints.

To fill these two knowledge gaps, this study proposes an adjoint subordination approach by deriving a backward-in-time model (~~which is~~ also known as an~~called~~ "adjoint") for~~of~~ the 3-*d*~~,~~ time fractional-derivative equation (FDE) subordinated to water flow with or without a highly resolved velocity field. Such a forward-in-time FDE was proposed by Zhang et al. (2015) as a general forward model for pollutant transport in various geological media. Notably, two other vector nonlocal transport models ~~can also incorporate the local variation of velocity in non-Fickian diffusion~~, ~~which are~~ the well-known continuous-time random walk (CTRW) framework (Hansen and Berkowitz, 2020) and the multi-scaling FDE model (Zhang, 2022), can also incorporate local velocity variations into non-Fickian diffusion.~~.~~ The CTRW framework allows for various memory functions to define solute transition times, but does not separate sub-diffusion (due to solute retention) and super-diffusion (e.g., due to ~~for example~~ preferential flow paths) (Lu et al., 2018). This study selects ~~The~~ the subordinated time-FDE, as ~~shown~~ explained in **section ~~1~~2**, ~~is selected for this study because~~for two key reasons: (i) it can capture both sub-diffusion (using the time fractional derivative) and sub-grid super-diffusion (via subordination, distinct from the space fractional derivative), and (ii) it offers ~~is~~ computational~~ly~~ ~~more~~ efficienc~~t~~y ~~than~~ compared to the multi-scaling FDE (introduced in **section 4**).

The ~~rest~~ remainder of this work is ~~organized~~ structured as follows. **Section 2** applies a sensitivity analysis approach to build the adjoint of the subordinated time-FDE, and then develops and validates a Lagrangian solver of the ~~resultant~~ resulting BTTP model. **Section 3** checks the feasibility of the adjoint model and its solver by quantifying BTTP, identifying the release history of contaminants in an alluvial aquifer and a river with ~~a~~ uniform velocity, and calculating groundwater ages dated ~~by~~ using environmental tracers in a regional-scale alluvial aquifer with ~~a~~ fine velocity resolution ~~of velocities~~. **Section 4** discusses the identification of contaminant source locations based on the backward location probability density function (BLP) and extends the backward probability model ~~extension~~. **Section 5** draws the main conclusions.

2 Methodology development

This section derives the model and solver for backward-in-time subordination to water flow in heterogeneous media. The concept of subordination to regional flow was initially ~~was first~~ proposed by Baeumer et al. (2001) and later ~~then~~ extended to multi-dimensional flow by Zhang et al. (2015). ~~Subordination is~~ Subordination is a statistical method that ~~can~~ randomizes the operational time experienced by ~~each~~ individual particles in a random process (Feller, 1971). When ~~applying~~ applied ~~subordination~~ to regional flow, this process captures fast displacement of pollutant particles along streamlines ~~is captured~~ during the randomized operational time, as shown and explained in the following model (1a).

2.1 Forward and backward models

2.1.1 Three-dimensional transport and adjoint models

We propose the following 3-*d* subordinated time-FDE to track pollutants in streams and aquifers with ~~a~~ vector velocity, after adding source/sink and reaction terms and initial/boundary conditions in the vector model proposed by Zhang et al. (2015):

$$b\frac{\partial(\theta C)}{\partial t} + \beta\frac{\partial^{\gamma,\lambda}(\theta C)}{\partial t^{\gamma,\lambda}} = -\nabla_{\vec{V}}(\theta C) + \sigma^*\left(\nabla_{\vec{V}}\right)^{\alpha,\kappa}(\theta C) + q_I C_I - q_o C - \theta r C \tag{1a}$$

$$C(\vec{x}, t=0) = \frac{M_0}{\theta}\,\delta(\vec{x} - \vec{x}_0) \tag{1b}$$

$$C(\vec{x}, t)|_{\xi_1} = g_1(t) \tag{1c}$$

$$\left[\sigma^*\left(\nabla_{\vec{V}}\right)^{\alpha-1,\kappa}(\theta C)\right]\cdot n_2\Big|_{\xi_2} = g_2(t) \tag{1d}$$

$$\left[V(\vec{x}, t)\,\theta C - \sigma^*\left(\nabla_{\vec{V}}\right)^{\alpha-1,\kappa}(\theta C)\right]\cdot n_3\Big|_{\xi_3} = g_3(t) \tag{1e}$$

where C [ML^{-3}] denotes the solute concentration, b (= 0 or 1) [dimensionless] is a factor controlling the type of the time FDE, θ [dimensionless] is the effective porosity, β [$T^{\gamma-1}$] is the fractional capacity coefficient, σ^* [L] is a scaling factor for subordination, $\vec{V}$ [LT^{-1}] is the velocity vector, $\nabla_{\vec{V}}$ is an advection operator defined via $\nabla_{\vec{V}} = \nabla(\vec{V}C)$, q_I [T^{-1}] is the source inflow rate, C_I is the inflow concentration, q_O is the sink outflow rate, r [T^{-1}] is the first-order decay constant, M_0 is the initial source mass, g_i ($i = 1, 2, 3$) is a known function at the type-i boundary (to define the constant concentration or pollutant flux at the boundary), ξ_i ($i = 1, 2, 3$) is the domain of the type-i boundary, $\vec{x}$ [L] denotes the spatial coordinate, t [T] is the (forward) time, and n_2 and n_3 are ~~the~~ outward unit normal vectors on the type-2 and type-3 boundaries, respectively. We ~~name~~ refer to Eq. (1a) as the subordinated fractional-~~dispersion~~ derivative equation (S-FDE).

The S-FDE (1a) captures the concurrent sub-diffusion and super-diffusion, ~~(~~driven by different mechanisms~~) using~~ represented by different terms. ~~i~~In Eq. (1a), ~~The~~ the symbol $\frac{\partial^{\gamma,\lambda}}{\partial t^{\gamma,\lambda}}$ ~~in Eq. (1a)~~, which is the mixed Caputo fractional derivative with an index γ [dimensionless] ($0 < \gamma \leq 1$) and a temporal truncation parameter λ [T^{-1}] (Baeumer et al., 2018), defines sub-diffusion due to solute retention. The operator $\left(\nabla_{\vec{V}}\right)^{\alpha,\kappa}$, ~~which denotes~~representing subordination to the flow field with an index α [dimensionless] ($1 < \alpha \leq 2$) for the tempered stable density (with the maximumly positive skewness $\beta^* = +1$) and a spatial truncation parameter κ [L^{-1}], describes fast downstream displacements~~ to downstream motion~~. It is worth noting that pollutant particles undergo advective displacement controlled by local mean velocity, with individual particles migrating along various flow paths in a heterogeneous medium, leading to random mechanical dispersion due to local speeds deviating from the mean velocity. Eq. (1a) assumes a (tempered) α-stable density distribution for random mechanical dispersive jumps, rescaled by the mean local velocity. This (tempered) α-stable density encompasses both Gaussian and power-law densities as two end members. ~~The~~ Therefore, ~~method of~~ subordination to regional flow ~~expands~~ extends ~~the~~ standard symmetric mechanical dispersion to non-symmetric, super-diffusive mechanical dispersion along streamlines, ~~caused~~ driven by ~~the~~ local velocity variation~~s of velocities (such as~~, like super-diffusion along preferential flow paths~~)~~. Notably, if molecular diffusion is not negligible, it can be included in Eq. (1), combining with the subordination term responsible for mechanical dispersion to define hydrodynamic dispersion~~the molecular diffusion term can be added to Eq. (1) to define the full range of hydrodynamic dispersion, if the molecular diffusive strength is not negligible~~.

To derive the backward model for the S-FDE (1) using the adjoint approach (Neupauer and Wilson, 2001), we first convert it to the model governing the state sensitivity $\phi = \frac{\partial C}{\partial f}$, where f is a system parameter and selected as the initial mass M_0 as in Neupauer and Wilson (2001) and Zhang (2022). This can be done by taking the first-order derivative of each term in the S-FDE (1) with respect to M_0, which leads to:

$$\left(b\frac{\partial}{\partial t} + \beta\frac{\partial^{\gamma,\lambda}}{\partial t^{\gamma,\lambda}}\right)(\theta\phi) = -\nabla_{\vec{V}}(\theta\phi) + \sigma^*\left(\nabla_{\vec{V}}\right)^{\alpha,\kappa}(\theta\phi) - (q_o + \theta r)\phi \tag{2a}$$

$$\phi(\vec{x}, t = 0) = \frac{\partial C(\vec{x}_i)}{\partial M_0} = \frac{1}{\theta}\,\delta(\vec{x} - \vec{x}_0) \tag{2b}$$

$$\phi(\vec{x}, t)|_{\xi_1} = 0 \tag{2c}$$

$$\left[\sigma^*\left(\nabla_{\vec{V}}\right)^{\alpha-1,\kappa}(\theta\phi)\right]\cdot n_2\Big|_{\xi_2} = 0 \tag{2d}$$

$$\left[V\theta\phi - \sigma^*\left(\nabla_{\vec{V}}\right)^{\alpha-1,\kappa}(\theta\phi)\right]\cdot n_3\Big|_{\xi_3} = 0 \tag{2e}$$

where the time fractional derivative operator commutes.

We then ~~add~~ incorporate the adjoint state of the concentration in the S-FDE (2a) by taking the inner product of each term of Eq. (2a) with an arbitrary function A, which represents the adjoint state:

$$\int_0^T \int_\Omega \left[Ab\frac{\partial(\theta\phi)}{\partial t} + A\beta\frac{\partial^{\gamma,\lambda}}{\partial t^{\gamma,\lambda}}(\theta\phi) + A\nabla_{\vec{V}}(\theta\phi) - A\sigma^*\left(\nabla_{\vec{V}}\right)^{\alpha,\kappa}(\theta\phi) + A(q_o + \theta r)\phi\right] d\Omega\, dt = 0 \tag{3}$$

where Ω denotes the whole model domain. ~~Next, we change the position of the state sensitivity ϕ and the adjoint sate A in the first four terms of Eq. (3).~~ Afterward, through sensitivity analysis, we derive the backward model (please refer to Appendix A for details):

~~For example, the 1st term in Eq. (3), denoted as I_1, can be re-arranged using integration by parts:~~

~~$$I_1 = \int_\Omega \left[\int_0^T Ab\frac{\partial(\theta\phi)}{\partial t}dt\right] d\Omega = \int_\Omega \left\{[Ab\theta\phi]|_{t=0}^{t=T} - \int_0^T \theta\phi b\frac{\partial A}{\partial t}dt\right\} d\Omega\,. \tag{4}$$~~

The 2nd term in Eq. (3) contains the time fractional derivative and can be re-arranged using the fractional-order integration by parts, as shown in Zhang (2022):

$$I_2 = \int_\Omega \left[\int_0^T A\beta \frac{\partial^{\gamma,\lambda}(\theta\phi)}{\partial t^{\gamma,\lambda}} dt\right] d\Omega = \int_\Omega \left\{A|_{t=T}\, \beta\, I_+^{1-\gamma,\lambda}(\theta\phi)|_{t=T} - [\theta\phi]|_{t=0}\, \beta\, I_-^{1-\gamma,\lambda}(A)|_{t=0} + \int_0^T \theta\phi\beta \frac{\partial^{\gamma,\lambda} A}{\partial(-t)^{\gamma,\lambda}} dt\right\} d\Omega, \quad (5)$$

where the symbol $I_+^{1-\gamma,\lambda}(f)$ denotes the positive fractional integral of order $1-\gamma$: $I_+^{1-\gamma,\lambda}(f) = e^{-\lambda T}\int_0^T f e^{\lambda t} \frac{(T-t)^{-\gamma}}{\Gamma(1-\gamma)} dt$, the symbol $I_-^{1-\gamma,\lambda}(f) = e^{\lambda T}\int_0^T f e^{-\lambda t} \frac{t^{-\gamma}}{\Gamma(1-\gamma)} dt$ denotes the negative fractional integral of order $1-\gamma$, and $\Gamma(\cdot)$ is the gamma function.

The 3rd term in Eq. (3), which describes the net advective flux, can be re-arranged using the integer-order integration by parts:

$$I_3 = \int_0^T \left\{\int_\Omega \nabla\cdot[A\theta V\phi]\, d\Omega - \int_\Omega \theta V\phi\, \nabla A\, d\Omega\right\} dt = \int_0^T \left\{\oint_\xi [A\theta V\phi]\cdot n\, d\xi - \int_\Omega \theta V\phi\, \nabla A\, d\Omega\right\} dt, \quad (6)$$

where the Gauss' divergence theorem is used: $\int_\Omega \nabla\cdot f\, d\Omega = \oint_\xi f\cdot n\, d\xi$, and n is the outward normal direction on the boundary ξ. Eqs. (4)-(6) are the same as those shown in Zhang (2022), which is expected since the same time fractional derivative term was used in these FDEs.

The 4th term in Eq. (3) contains the subordination operator and can be re-arranged using the integration by parts for twice, as shown in Zhang (2022):

$$I_4 = \int_0^T \left[\int_\Omega A\sigma^*(\nabla_V)^{\alpha,k}(\theta\phi)\, d\Omega\right] dt = \int_0^T \left\{\oint_\xi \sigma^*\left[A\, I_+^{2-\alpha,k}\left(\nabla_V(\theta\phi)\right)\right]\cdot n\, d\xi + \oint_\xi \sigma^*\left[\nabla_V(\theta\phi)\, I_-^{2-\alpha,k}(A)\right]\cdot n\, d\xi\right\} dt$$

$$+ \int_0^T \left\{\oint_\xi \sigma^*\left[\theta\phi\, (\nabla_{\overleftarrow{V}})^{\overline{\alpha-1},k}(A)\right]\cdot n\, d\xi\right\} dt - \int_0^T \left\{\int_\Omega \sigma^*\theta\phi\, (\nabla_{\overleftarrow{V}})^{\overline{\alpha},k}(A)\, d\Omega\right\} dt\, . \quad (7)$$

Here the operator $(\nabla_{\overleftarrow{V}})^{\overline{\alpha},k}$ denotes subordination to the reversed flow field ($\overleftarrow{V}$) where the tempered stable density (with order α) has the maximumly negative skewness $\beta^* = -1$, meaning that fast displacements are from downstream to upstream (for backward tracking).

Neupauer and Wilson (2001) showed that the adjoint state A is a measure of the change in concentration for a unit change in source mass M_0. In sensitivity analysis, the marginal sensitivity of a performance measure A with respect to M_0 is (Neupauer and Wilson, 2001):

$$\frac{dP}{dM_0} = \int_0^T \int_\Omega \left[\frac{\partial h(M_0, C)}{\partial C}\phi\right] d\Omega\, \mathrm{d}t\, , \quad (8)$$

where $h(M_0, C)$ is a functional of the state of the system. Inserting I_1~I_4 expressed by Eqs. (4)-(7) into the inner product equation (3), and then subtracting this updated Eq. (3) from the marginal sensitivity equation (8), we obtain:

$$\frac{dP}{dM_0} = \int_\Omega \int_0^T \phi\left\{\frac{\partial h}{\partial C} + b\theta\frac{\partial A}{\partial t} - \beta\theta\frac{\partial^{\gamma,\lambda} A}{\partial(-t)^{\gamma,\lambda}} + \theta V\nabla A - \sigma^*\theta(\nabla_{\overleftarrow{V}})^{\overline{\alpha},k}(A) - (q_o + \theta r)A\right\} d\Omega\, dt$$

$$- \int_\Omega \left\{[Ab\theta\phi]|_{t=T} - [Ab\theta]|_{t=0}\frac{\partial C_i}{\partial M_0} + A|_{t=T}\beta I_+^{1-\gamma,\lambda}(\theta\phi)|_{t=T} - [\theta\phi]|_{t=0}\,\beta I_-^{1-\gamma,\lambda}(A)|_{t=0}\right\} d\Omega$$

$$- \int_0^T \oint_\xi \left[A\theta V\phi - A\, I_+^{2-\alpha,k}\left(\nabla_V(\theta\phi)\right) - \nabla_V(\theta\phi)\, I_-^{2-\alpha,k}(A) - \theta\phi\,(\nabla_{\overleftarrow{V}})^{\overline{\alpha-1},k}(A)\right]\cdot n\, d\xi dt. \quad (9)$$

~~To eliminate ϕ from Eq. (9), we define A such that the terms containing ϕ vanish. Since the double integral in Eq. (9) (shown by the first line in Eq. (9)) can be eliminated when the summation of all the terms inside the bracket is zero, this produces the adjoint equation of the S-FDE (1a):~~

~~$$b\theta\frac{\partial A}{\partial t}-\beta\theta\frac{\partial^{\gamma,\lambda}A}{\partial(-t)^{\gamma,\lambda}}=-\theta V\nabla A+\sigma^{*}\theta\left(\nabla_{\overleftarrow{V}}\right)^{\overline{\alpha},k}(A)+(q_{O}+\theta r)A-\frac{\partial h}{\partial C}. \quad (10)$$~~

~~Assuming (i) the backward time $s=T-t$ where T is the detection time, (ii) steady-state groundwater flow (so that $\theta V\,\nabla A-q_{O}A=\nabla(\theta VA)-q_{I}A$), and (iii) un-compressible aquifer skeleton (so that $\partial\theta/\partial t=0$), we can re-write Eq. (10) as:~~

$$b\frac{\partial(\theta A)}{\partial s}+\beta\frac{\partial^{\gamma,\lambda}(\theta A)}{\partial s^{\gamma,\lambda}}=\nabla_{\overleftarrow{V}}(\theta A)-\theta\sigma^{*}\left(\nabla_{\overleftarrow{V}}\right)^{\overline{\alpha},\kappa}A-(q_{I}+\theta r)A+\frac{\partial h}{\partial C} \quad (4~~11~~a)$$

$$A(\vec{x},s)|_{s=0}=0 \quad (4~~11~~b)$$

$$A(\vec{x},s)|_{\xi_{1}}=0 \quad (4~~11~~c)$$

$$\left[-A\theta V+\sigma^{*}\theta\left(\nabla_{\overleftarrow{V}}\right)^{\overline{\alpha-1},\kappa}(A)\right]\cdot n_{2}\Big|_{\xi_{2}}=0 \quad (4~~11~~d)$$

$$\left[\sigma^{*}\theta\left(\nabla_{\overleftarrow{V}}\right)^{\overline{\alpha-1},\kappa}(A)\right]\cdot n_{3}\Big|_{\xi_{3}}=0 \quad (4~~11~~e)$$

where s $(=T-t)$ represents backward time (with T as the detection time), and the operator $\left(\nabla_{\overleftarrow{V}}\right)^{\overline{\alpha},k}$ denotes subordination to the reversed flow field ($\overleftarrow{V}$) with a tempered α-stable density characterized by maximum negative skewness ($\beta^{*}=-1$), indicating fast displacements from downstream to upstream during backtracking. ~~where~~ Here, the initial condition (~~11b~~4b) $A(\vec{x},t)|_{t=T}=A(\vec{x},s)|_{s=0}=0$ and the boundary conditions (~~11c~~4c)~(~~11e~~4e) are obtained by making sure that the remaining terms in Eq. (A6~~9~~) in Appendix A defines the following marginal sensitivity:

$$\frac{dP}{dM_{0}}=\int_{\Omega}\left\{\left[(Ab\theta)|_{t=0}+\theta|_{t=0}\,\beta I_{-}^{1-\gamma,\lambda}(A)|_{t=0}\right]\frac{\partial C_{i}}{\partial M_{0}}\right\}d\Omega\ . \quad (5~~12~~)$$

Therefore, to convert the forward-in-time S-FDE (1) to its backward counterpart (4~~11~~), we need to (i) reverse the flow field, (ii) convert the source/sink terms and boundary conditions, and (iii) reverse the skewness in the stable density defining backward mechanical dispersive jumps. The first two changes were identified before by Neupauer and Wilson (2001) for the classical ADE (although the exact forward-backward transition is new here), and the last change is new. In the following we name the backward-in-time model (4~~11~~) as the adjoint S-FDE.

2.1.1 One-dimensional simplifications

The 1-*d* simplification of the vector forward-in-time S-FDE (1) takes the form:

$$b\frac{\partial(\theta C)}{\partial t}+\beta\frac{\partial^{\gamma,\lambda}(\theta C)}{\partial t^{\gamma,\lambda}}=-\frac{\partial(V\theta C)}{\partial x}+\sigma^{*}\left(\frac{\partial}{\partial x}\right)_{V}^{\alpha,\kappa}(\theta C)+q_{I}C_{I}-q_{O}C-\theta rC$$

$$C(x,t=0)=\frac{M_{0}}{\theta}\,\delta(x-x_{0})$$

$$C(x,t)|_{\xi_{1}}=g_{1}(t)$$

$$\left[\sigma^{*}\left(\frac{\partial}{\partial x}\right)_{V}^{\alpha-1,\kappa}(\theta C)\right]\Big|_{\xi_2} = g_2(t)$$

$$\left[V\theta C - \sigma^{*}\left(\frac{\partial}{\partial x}\right)_{V}^{\alpha-1,\kappa}(\theta C)\right]\Big|_{\xi_3} = g_3(t)$$

If the velocity V in the equations listed above is constant, this 1-*d* S-FDE reduces to the following 1-*d* standard FDE:

$$b\frac{\partial(\theta C)}{\partial t} + \beta\frac{\partial^{\gamma,\lambda}(\theta C)}{\partial t^{\gamma,\lambda}} = -V\frac{\partial(\theta C)}{\partial x} + D^{*}\frac{\partial^{\alpha,\kappa}}{\partial x^{\alpha,\kappa}}(\theta C) + q_I C_I - q_o C - \theta r C$$ (6~~13~~a)

$$C(x, t = 0) = \frac{M_0}{\theta}\,\delta(x - x_0)$$ (6~~13~~b)

$$C(x,t)|_{\xi_1} = g_1(t)$$ (6~~13~~c)

$$\left[D^{*}\frac{\partial^{\alpha-1,\kappa}}{\partial x^{\alpha-1,\kappa}}(\theta C)\right]\Big|_{\xi_2} = g_2(t)$$ (6~~13~~d)

$$\left[V\theta C - D^{*}\frac{\partial^{\alpha-1,\kappa}}{\partial x^{\alpha-1,\kappa}}(\theta C)\right]\Big|_{\xi_3} = g_3(t)$$ (6~~13~~e)

where $D^{*} = \sigma^{*}V$. Therefore, ~~for~~ in 1-*d* transport with a constant velocity, the scaling factor σ^{*} in the S-FDE is analogous to dispersivity, a parameter ~~commonly~~ often used to scale mechanical dispersion (~~and~~ typically fitted by observed plume data), and the subordination index α is equal to the index of the (tempered) space fractional derivative.

The 1-*d* adjoint of FDE (6~~13~~) is a simplified version of the 3-*d* adjoint S-FDE (4~~11~~):

$$b\frac{\partial(\theta A)}{\partial s} + \beta\frac{\partial^{\gamma,\lambda}(\theta A)}{\partial s^{\gamma,\lambda}} = V\frac{\partial(\theta A)}{\partial x} + D^{*}\theta\frac{\partial^{\alpha,\kappa}}{\partial(-x)^{\alpha,\kappa}}A - (q_I + \theta r)A + \frac{\partial h}{\partial C}$$ (7~~14~~a)

$$A(x,s)|_{s=0} = 0$$ (7~~14~~b)

$$A(x,s)|_{\xi_1} = 0$$ (7~~14~~c)

$$\left[A\theta V - D^{*}\theta\frac{\partial^{\alpha-1,\kappa}}{\partial(-x)^{\alpha-1,\kappa}}A\right]\Big|_{\xi_2} = 0$$ (7~~14~~d)

$$\left[D^{*}\theta\frac{\partial^{\alpha-1,\kappa}}{\partial(-x)^{\alpha-1,\kappa}}A\right]\Big|_{\xi_3} = 0$$ (7~~14~~e)

The backward FDE (7~~14~~) ~~is consistent~~aligns with the one~~at~~ derived by Zhang et al. (2022), validating the 1-*d* simplification of the backward model (4~~11~~).

When the factor $b = 1$, the capacity coefficient $\beta = 0$ (meaning no immobile phase or solute retention), and the space index $\alpha = 2$ (representing normal diffusion), the forward S-FDE model (6~~13~~) reduces to the classical 2nd-order ADE:

$$\frac{\partial(\theta C)}{\partial t} = -V\frac{\partial(\theta C)}{\partial x} + D^{*}\frac{\partial^{2}}{\partial x^{2}}(\theta C) + q_I C_I - q_o C - \theta r C$$

$$C(x, t = 0) = \frac{M_0}{\theta}\,\delta(x - x_0)$$

$$C(x,t)|_{\xi_1} = g_1(t)$$

$$\left[D^{*}\frac{\partial}{\partial x}(\theta C)\right]\Big|_{\xi_2} = g_2(t)$$

$$\left[V\theta C - D^{*}\frac{\partial}{\partial x}(\theta C)\right]\Big|_{\xi_3} = g_3(t)$$

and the corresponding backward model (7~~14~~) is simplified to:

$$\frac{\partial(\theta A)}{\partial s} = V\frac{\partial(\theta A)}{\partial x} + D^*\theta\frac{\partial^2 A}{\partial x^2} - (q_I + \theta r)A + \frac{\partial h}{\partial C} \quad (8\text{~~15~~a})$$

$$A(x,s)|_{s=0} = 0 \quad (8\text{~~15~~b})$$

$$A(x,s)|_{\xi_1} = 0 \quad (8\text{~~15~~c})$$

$$\left[A\theta V + D^*\theta\frac{\partial A}{\partial x}\right]\Big|_{\xi_2} = 0 \quad (8\text{~~15~~d})$$

$$\left[D^*\theta\frac{\partial A}{\partial x}\right]\Big|_{\xi_3} = 0 \quad (8\text{~~15~~e})$$

which is the same as the 1-*d* backward ADE derived by Neupauer and Wilson (1999).

The applicability of both the 3-*d* backward model (4~~11~~) and its 1-*d* simplification (7~~14~~) is examined using real-world aquifers and streams in **section 3**. The 3-*d* backward model (4~~11~~) is needed since most transport processes in natural aquifers are multi-dimensional. The 1-*d* backward model (7~~14~~) can also be useful since (i) ~~many times we need to first~~ focusing on longitudinal transport is often necessary, and (ii) most successful hydrology applications of ~~the~~ FDEs ~~in hydrology~~ are limited to 1-*d*, as discussed in~~;~~ ~~see~~ the ~~extensive~~ comprehensive review by Zhang et al. (2017). The classical 1-*d* backward ADE model (8~~15~~) will also be applied to reveal the impact of non-Fickian transport on BTTP by comparing with the adjoint S-FDE solutions.

2.2 Lagrangian solver

The adjoint S-FDE (4~~11~~) with complex boundary conditions ~~cannot be solved~~lacks an analytical~~ly~~ solution ~~to obtain the~~for BTTP, and hence a grid-free, fully Lagrangian numerical solver is proposed here. The Lagrangian solver for the forward-in-time S-FDE (1) under various boundary conditions was developed and tested by Zhang et al. (2019a). We briefly introduce it here. This forward-in-time Lagrangian solver contains three main steps. *Step 1* decomposes mobile and immobile phases using the following temporal Langevin equation ~~, which is a stochastic model~~ that separates particle waiting time and operational time~~,~~ w~~hose~~ith a probability density function (PDF) ~~follows~~ following the tempered stable density with index γ (Meerschaert et al., 2008):~~.~~

$$dt_i = dM_i + \left[\cos\left(\frac{\pi\gamma}{2}\right)\beta\, dM_i\right]^{1/\gamma} dL_{\gamma,\lambda}(\beta^* = +1, \varepsilon = 1, \mu = 0)$$

where dt_i denotes the total time for the particle spent in the *i*-th jump, dM_i represents the operational time during this jump (which can be assigned uniformly), and $dL_{\gamma,\lambda}$ is a tempered stable random variable with the maximum positive skewness β^*, unit scale ε, and zero shift μ. *Step 2* applies subordination to regional flow by calculating ~~the~~ streamline-oriented random mechanical displacements for each particle (whose PDF follows the tempered α-stable density), ~~re~~scaled by ~~the~~ local velocity, as described above. *Step 3* then adjusts particle trajectories ~~around~~ near boundaries ~~according to the~~using particle-tracking schemes developed by Zhang et al. (2015).

We convert the above-mentioned forward-in-time Lagrangian solver to its backward counterpart for~~to~~ ~~approximate the~~ adjoint S-FDE (4~~11~~) approximation with ~~, by incorporating~~ three main modifications. ***First***, reverse ~~each~~ vector components of ~~the~~ velocity ~~is reversed to calculate the~~for backward advective displacement of particles during the operational time. ***Second***, change ~~the~~ skewness of the (tempered) α-stable Lévy jumps ~~is changed from the~~from positive ~~maximum~~ (to capture ~~downward~~ downstream mechanical displacement) to ~~the~~ negative maximum (to backtrack pollutants located upstream initially). ***Third***, modify ~~the~~ source/sink terms and boundary conditions ~~are modified~~ according to those defined in the adjoint model (4~~11~~) and **Table 1**. For example, ~~the~~ forward sink term ~~in the forward model, which is~~ ($-q_o C$ in Eq. (1a))~~,~~ ~~is replaced by~~becomes the load term $\frac{\partial h}{\partial c}$ in the adjoint model (4~~11~~a)~~,~~ ~~which describes~~representing the initial probability source in the backward Lagrangian solver. **Table 1** ~~shows the main~~details changes and hydrogeologic interpretations of these boundary conditions (~~including their~~ value and type) converted from the forward S-FDE to ~~its~~ backward counterpart at ~~the~~ upstream (inlet) and downstream (outlet) boundaries. In this 1-*d* ~~For~~ simpli~~fication~~~~eity~~, we ~~**Table 1** uses the 1-*d* simplification and~~ assume~~s~~ ~~that the~~ forward flow ~~direction is from~~ left to right. The Dirichlet, Neumann, Robin, and infinite boundaries in the forward model transform to the absorbing, fully reflective, partially reflective, and free boundaries in the backward model, respectively, to correctly backtrack particle trajectories around boundaries and recover pollutant release history. For example, the non-zero~~,~~ Dirichlet boundary condition in the forward model (~~expressed by~~ Eq. (1c)) converts to an absorbing boundary in the backward model (~~expressed by~~ Eq. (4~~11~~c)), which is expected since the forward source term ~~in the forward model~~ becomes the sink term in the backward model. In addition, a non-zero~~,~~ Neumann boundary condition in the forward model (1-*d*) (representing an immobile diffusive source ~~located~~ at the inlet boundary) transforms into a fully reflective boundary condition in the backward model (4~~11~~d) (meaning that no external sources outside the upstream boundary), ~~which is necessary to~~ ensur~~e~~ing ~~that~~ no particles ~~can~~ exist this upstream boundary (**Table 1**).

This backward-in-time Lagrangian solver is ~~significantly more~~ computationally more efficient than the standard Eulerian solver~~, since~~ because (i) particles in the immobile phase remain motionless and therefore ~~do not~~ require no ~~any~~ calculations, and (ii) the streamlines can be ~~calculated~~ semi-analytically calculated (LaBolle, 2006) for ~~the~~ streamline-projected mechanical dispersion during ~~subordination to~~ regional flow subordination.

2.3 Numerical experiments and validation

Here we check this Lagrangian solver using either simplified~~e~~ cases (~~i.e.,~~ 1-*d*) or qualitative evaluation due to the lack of other numerical solvers for the 3-*d* adjoint S-FDE (4~~11~~a). The number density of particles exiting the source location~~,~~ ~~(~~re-scaled by ~~the~~ velocity~~,)~~ defines the flux-concentration based BTTP. This method~~,~~ ~~which~~ estimates the PDF of each release time (s) for the pollutants identified at the monitoring well at present.

Results of the first numerical experiments are plotted in **Figure 1**. For validation, we developed an implicit Eulerian finite difference solver for~~of~~ the 1-*d* adjoint FDE (~~14a~~7a)~~,~~ ~~was also developed by~~ adopting the Grünwald approximation scheme proposed by Meerschaert and Tadjeran (2004) ~~which can~~for efficient~~ly~~ ~~approximate~~ fractional derivative approximations. The

Lagrangian BTTP solutions ~~of BTTPs~~align with the Eulerian solutions, ~~although containing~~despite some apparent noise at low BTTPs, arising from ~~due to~~ the finite number of particles used in the model~~, generally match the Eulerian solutions~~ (**Figure 1**). In these ~~Here~~experiments, we ~~the~~assumed the backward travel distance ~~, which represents the known source location, is assumed to be~~of 10 (dimensionless)~~,~~ and a~~the~~ model domain dimension ~~is~~ 100 times larger than the backward travel distance. Consequently, we treated ~~T~~the boundaries as~~therefore can be assumed to be~~ effectively infinite~~,~~ and ~~hence,~~ applied the free boundary condition as outlined ~~listed~~ in **Table 1** ~~is applied for the Lagrangian solver~~. Our ~~N~~numerical analysis also reveal~~s~~ed that~~,~~ varying the time truncation parameter λ impacts the BTP peak time and the late-time tail. ~~when all the other parameters remain unchanged, a~~A larger ~~time truncation parameter~~ λ delays the BTTP peak time ~~of the BTTP~~ (because a larger λ leads to a longer peak waiting time in the truncated stable density) and ~~shrinks~~ narrows the late-time tail of the BTTP (because a larger λ significantly narrows the particle's waiting time PDF by truncating extremely long waiting times) (**Figures 1a** and **1b**). ~~In addition, w~~When λ is very small (i.e., $\lambda \leq 10^{-6}\ T^{-1}$, representing an untruncated, standard stable density for the random waiting time), the late-time BTTP tail declines at a rate of $s^{-1-\gamma}$ (**Figure 1d**). In addition, ~~W~~a small and negligible ~~hen the~~ space truncation parameter κ ~~is small and negligible, the~~results in an early-time BTTP tail increasing~~es~~ ~~at~~ with a rate of s^{1}, a characteristic stable across various subordination indexes ~~(this "+1" slope in a log-log plot remains stable for all the subordination index~~ α varying from 1 to 2~~)~~ (**Figures 1b** and **1d**). When all the other parameters remain unchanged, a smaller subordination index α and a larger time index γ accelerates the BTTP peak, because a smaller α engenders a faster-moving plume peak and a larger γ describes weaker retention. Therefore, the BTTP early-time tailing behaviour (representing super-diffusion) is ~~dominated~~ governed by ~~the subordination index~~ α and ~~its truncation parameter~~ κ, while the late-time ~~the BTTP late-time~~ tailing behaviour (representing sub-diffusion) is mainly controlled by ~~the time index~~ γ and ~~its truncation parameter~~ λ.~~, and t~~The BTTP peak is affected by all these four parameters, reflecting the interplay ~~(due to the competition~~ between super- and sub-diffusive transport~~)~~. These BTTP features can be critical signals for real-world applications. For example, the BTTP peak time describes the most likely release time of an instantaneous point source, and the BTTP tails control the backward travel time distribution which also defines the groundwater age distribution (see the application in **section 3.2**) and transient indexes for assessing aquifer vulnerability (Zhang et al., 2018).

The second numerical experiments apply the Lagrangian solver to backtrack particles in non-uniform flow fields (**Figure 2**). Two 2-*d* Brownian random hydraulic conductivity (K) fields were first generated using the method developed by Zhang et al. (2019a) (**Figures 2a** and **2c**). Particularly, log-normal random K values were distributed in space using the Fourier filter function. The Hurst parameter in the filter function defines the spatial correlation of K values: a relatively "homogeneous" K-field exhibits weak correlation of K (e.g., **Fig. 2a**), while a "heterogeneous" K-field displays strong correlation (e.g., **Fig. 2c**). Steady-state groundwater flow was then calculated by the United States Geological Survey (USGS) software MODFLOW (Harbaugh, 2005) (shown by the black lines in **Figures 2b** and **2d**). Backward particle tracking plumes were finally obtained by the Lagrangian solver proposed above (shown by the contour maps in **Figures 2b** and **2d**). ~~For~~ In K field #1 characterized by~~with~~ a relatively "homogeneous" distribution of K, particles ~~starting~~ originating from different wells move backward at a similar rate, and are eventually removed fromexiting the system when upon reaching the upstream boundary (which is the left boundary locatedlocated as at $x = 0$ and is assumed to be an absorbing boundary in the backward model) (**Figure 2b**). These All plumes follow the general path of local streamline paths, as expected forin accordance with the streamline projection method proposed aboveoutlined earlier. The transverse expansion of the plume is due attributed to molecular diffusion added toincorporated into particle dynamics (to capture hydrodynamic dispersion). For In K field #2, representing a more heterogeneous K field with layereding deposits, particles starting from in the high-K zone move quickly rapidly and then exit the model domain (**Figure 2d**). These backward dynamics follow our logical expectations, but cannot be independently validated, as far as since, to the best of our knowledge extends, due to the absence of there are no otheralternative solvers available for the vector model (411).

3 Field applications

The adjoint S-FDE model is applied in this section to recover the release history of pollutants in aquifers and rivers and calculate groundwater ages dated by environmental tracers. As shown below, tThese surface and subsurface flow systems, characterized by exhibit different degrees levels of medium heterogeneity, various diverse flow velocity resolutions, and boundary conditions, and different spatiotemporal scales, serve as a comprehensive testbed to evaluate which provide an ideal set of natural variability to test the real-world applicability of the physical model and numerical solver developed in this study.

3.1 BTTP application case 1: Recover release history of pollutants at the MADE site

Natural-gradient tracer tests were conducted at the Macrodispersion Experiment (MADE) site in Columbus, Mississippi, U.S. (Adams and Gelhar, 1992; Boggs et al., 1992), identifying mixed sub- and super-diffusive pollutant transport in an alluvial aquifer measuring approximatelya ~11 m in thickness and ~300 m leongth alluvial aquifer (Bianchi et al., 2016; Yin et al., 2020). Non-Fickian transport at the MADE site motivated the development of various numerical and /stochastic transport models in the last three decades (see the review by Zheng et al. (2011)), but the BTTP dominated by mixed sub/super-diffusion remained unknownuncharted. Here, we calculate its BTTP using the adjoint S-FDE (14a7a), (which is an upscaled model,) with a uniform velocity. The 1-d backward model is selected since the MADE site transport can be simplified by a 1-d process projected into the longitudinal direction, as demonstrateda convention upheld by many previous models (Zheng et al., 2011).

The seven parameters in the backward model (714a) can be conveniently estimated using mainly literature data. The strong sub- and super-diffusion observed at the MADE site implies that the two truncation parameters (λ and κ) can be simply neglected, reducing the leaving 5 unknown parameters to 5in model (14a). The subordination index α is analogous to the spatial index (=1.1) estimated by Benson et al. (2001) using the distribution of measured permeability. The time index γ (=0.39) and capacity coefficient β (=0.082 day$^{\gamma-1}$) were estimated by Zhang et al. (2010) using the decline rate of the observed mobile tracer mass. The velocity V (=0.24 m/day) can be approximated by the mean field velocity measured in the field, and the scaling factor σ^* is assumed to be 1 m since dispersion at the MADE site was found to be a similar the same order as V (Benson et al., 2001).

The predicted BTTPs are plotted in **Figure 3**. Here we choose the monitoring well located at the bromide plume's peak (obtained in from the MADE-1 bromide tracer test) as the detection location, denoted as x_w (which is defined as the location of the monitoring well detecting the maximum concentration), since this location represents the mass center of the tracer plume. The known contaminant source was knownis situated at the origin ($x_0 = 0$). The plume peak at during the first (Day 49) and second (Day 126) sampling cycles is located at $x_w = 3.0$ m and 7.0 m, respectively, providing two possible detection locations. These two detection locations lead to the two predicted BTTPs depicted in **Figure 3**, after applying the adjoint S-FDE (714a) with the seven parameters estimated above.

The model results show that, on the one hand, the peak of the flux-concentration based BTTP captures well the true release time (**Figures 3a** and **3b**). On the other hand, the peak of the BTTP based on the concentration profile for "immobile" particles (which were located at the source location and remained nearly stationary at the source location motionless during each unit time interval in calculatingfor BTTP calculation), has a higher value and corresponds to a much later time (twice that of the flux-concentration based BTTP peak), which significantly overestimates the true release time. This discrepancy is explained by the slower movement of the immobile phase source moving slower (than the mobile phase source) due to strong solute retention, resulting in an oldera more aged release time. For an aqueous phase observation, the flux-concentration based BTTP describes the PDF of release times for aqueous (or mobile) phase sources, while the immobile particles' concentration based BTTP describes the PDF of release times for absorbed (or immobile) phase sources. In the MADE-1 tracer test, the bromide tracer was initially injected into the upstream well as an initially mobile source, necessitating the and henceuse of the flux-concentration based BTTP is needed,. This demonstrates meaning that the adjoint S-FDE (714a) successfully recovers the tracer's release history. In addition, as shown in **Figure 3c**, the slope of the late-time BTTP for the immobile phase sources in a log-log plot (which is $-\gamma$) is -1 smaller (i.e., heavier) than that for the mobile phase sources (which is $-\gamma - 1$), describing the persistent sustained release of immobile pollutant mass at the source location and implying a high degree of uncertainty in the BTTP for the immobile phase source.

The adjoint ADE is also applied here for comparison. When the same velocity V (=0.24 m/day) and dispersion coefficient D^* ($= \sigma^* V = 0.24$ m^2/day) are used, the adjoint ADE significantly underestimates the true release time (not shown displayed here), as it. This is expected because the adjoint ADE cannot capture account for solute retention. Subsequently, We we attempted then calibration by adjustinge V (=0.068 m/day) and D^* (=0.68 m^2/day) by fittingto match the mean and variance of the observed bromide plumes. However, but the resultingant BTTP peak still underestimates underestimated the true release time by more thanover one order of magnitude (shown by the solid black line in **Figure 3**). Finally, we directly fitted V (=0.026 m/day, which is one order of magnitude smaller than the mean groundwater velocity) and D^* (=0.031 m^2/day) using the true release time for the detection well located at $x_w = 3.0$ m (shown by the dashed black line in **Figure 3a**)., Nevertheless, but thise best-fit adjoint ADE then overestimates overestimated the true release time by > 50% for the detection well at $x_w = 7.0$

m (shown by the dashed black line in **Figure 3b**). Therefore, the adjoint ADE with a constant velocity cannot reliably recover the release history of pollutants ~~undergoing~~ experiencing strong non-Fickian transport in the MADE aquifer, reaffirming ~~. The same~~ conclusions ~~was~~ drawn ~~by~~ in previous studies ~~for fitting~~regarding tracer transport at the MADE site using ~~the~~ ADE based models (Zheng et al., 2011).

3.2 BTTP application case 2: Groundwater age dating in Kings River alluvial aquifer, California

The vector backward S-FDE (4~~11~~a) is then used to calculate groundwater age~~the~~ distributions ~~of groundwater ages~~ ~~at~~ for the Kings River alluvial aquifer (KRAA) ~~located~~ in Fresno County, California, U.S. (**Figure 4**). The flux-concentration based BTTP also represents the groundwater age distribution and ~~provides core information~~serves as crucial data for groundwater sustainability assessments (Fogg et al., 1999; Weissmann et al., 2002; Fogg and LaBolle, 2006).

The KRAA system ~~consists of~~comprises five paleosol-bounded stratigraphic sequences recognized by Weissmann and Fogg (1999). One realization of the 3-*d* hydrofacies model built upon the Markov Chain model developed by Weissmann et al. (2004) is shown in **Figure 4**, where the hydrofacies model ~~contains~~ incorporates both the large-scale stratigraphic sequences and the intermediate-scale hydrofacies within each sequence. This 3-*d* Markov Chain model was built using hydrofacies distribution data ~~obtained~~ from 11 cores, 132 drillers' logs, and soil survey data. ~~The hydrofacies model contains a~~All ~~the~~ cores and drillers' logs were integrated as hard conditional data, ~~to~~ maximizing~~ally~~ the incorporation~~e~~ of observed information into the numerical model. This regional-scale model contains ~1 million cells, each with ~~the cell's~~ dimensions of 200 m, 200 m, and 0.5 m in the depositional strike, depositional dip, and vertical directions, respectively, with a total model domain size of 12,600 m × 15,000 m ×100.5 m along these three directions. We calculated ~~The~~ steady-state groundwater flow ~~was then calculated by~~using MODFLOW, applying ~~using~~ parameters and ~~/~~boundary conditions described by Weissmann et al. (2004) and Zhang et al. (2018b). ~~For example~~Specifically, we assigned ~~the~~ measured K values ~~was assigned~~ to each facies (gravel, sand, muddy sand, mud, and paleosol). The top of the model accounted for a recharge boundary, and the lateral and basal boundaries of the model were general head boundaries to allow inflow and ~~/~~outflow. The modeled hydraulic heads ~~were~~ closely matched ~~to~~ the measured data (Zhang et al., 2018b). We ~~utilized~~used ~~The~~ the resulting~~ant~~ fine-resolution velocity field ~~was used~~ to calculate BTTP using the adjoint S-FDE (~~11a~~4a).

We ~~f~~begin with~~irst~~ ~~conduct~~ a parameter sensitivity test using the adjoint S-FDE (4~~11~~). In these backward particle tracking models, ~~both~~ the water table (representing an internal boundary) and the lateral~~,~~ upstream boundary of the model are both set as absorbing boundaries, ~~(because they~~ representing the source locations.~~),~~ ~~and t~~The ~~other~~ remaining model boundaries are ~~simply~~ treated as fully reflective boundaries. An effective porosity of 0.33, ~~which was the best-fit~~a value previously determined as the best-fit in ~~in~~ Weissmann et al. (2004) and Zhang et al. (2018b), is applied for these simulations. We consider ~~T~~three cases to explore~~with~~ decreasing super-diffusion and increasing sub-diffusion ~~are considered here~~. Case 1 ~~captures~~ exhibits strong super-diffusion, ~~(~~characterized by ~~with the~~a time index $\gamma = 0.80$, a~~the~~ capacity coefficient $\beta = 0.1$ yr$^{\gamma-1}$, a~~the~~ subordination index $\alpha = 1.40$, and a~~the~~ scaling factor $\sigma^* = 0.4$ m.~~),~~ Case 2 represents ~~the~~ an intermediate scenario ~~(~~with $\gamma =$

0.72, $\beta = 0.2$ yr$^{\gamma-1}$, $\alpha = 1.45$, and $\sigma^* = 0.3$ m.), and Case 3 describes strong sub-diffusion, featuring (with $\gamma = 0.65$, $\beta = 0.3$ yr$^{\gamma-1}$, $\alpha = 1.50$, and $\sigma^* = 0.2$ m). The subordination truncation parameter (κ) remains the same for all three cases ($\kappa = 1.0 \times 10^{-5}$ m^{-1}). The resultant backward particle tracking snapshots at the backward time s=50 yrs is are plotted in **Figures 5a~5c** for these three cases. Driven by subordination to regional flow, particles move alongfollow streamlines and expand, especially particularly within high-permeability deposits (due also to molecular diffusion simultaneously along all three axis directions). Case 1 captures fast rapid backward (i.e., toward upstream) movement of particles due to strong super-diffusion, resuch thatsulting in -most particles arrive atreaching the water table within 50 yrs and are then leaving removed from the system, leaving only a few particles behind (**Figure 5a**). Contrarily, Case 3 captures the most delayed backward movement due to strong sub-diffusion, and resulting in the majority ofmost particles remaining in the aquifer with a limited spatial expansion, as shown depicted in **Figure 5c**. This parameter sensitivity test, therefore, shows demonstrates that the capability of the adjoint S-FDE (411) can to reasonably interpret non-Fickian dynamics in multi-dimensional aquifers. In addition, the corresponding BTTP for each case, which representings the age distribution for groundwater sampled at the well screen shown indicated in **Figure 5a** (the green rectangle), is plotted in **Figure 5d**. Notably, as the adjoint S-FDE transitions from Case 1 to Case 3, characterized by With a larger subordination index α and a smaller time index γ in the adjoint S-FDE (i.e., from Case 1 to Case 3), the BTTP shifts apparently towards older ages, with a decreasing peak and an expanding distribution. This illustrates , characterizing the impact of decreasing super-diffusion and increasing sub-diffusion on groundwater age distributions. This test shows underscores that keymain properties of the BTTP, including the mean, peak, and variance of groundwater ages, are sensitive to the two indexes α and γ. In Ffurther comparisons, it becomes evident that show that the classical adjoint ADE misses fails to capture the early arrivals of in the BTTP, because itprimarily due to its inability to account for cannot capture super-diffusion (figures not shown).

Finally, we compared the adjoint S-FDE solutions were compared towith chlorofluorocarbon-11 (CFC-11) ages measured by Burow et al. (1999) from USGS data for KRAA in 1994. The S-FDE model parameters cannot be predicted using the hydrofacies property-based method proposed by Zhang et al. (2014) for stationary hydrofacies models, due to the nonstationary distribution of hydrofacies inat KRAA. Instead, An an alternative approach was employed (of parameter fitting)by fitting is the age distribution for groundwater, especially forparticularly shallow groundwater, which can be calibrated using the groundwater age dated by environmental tracers such as CFCs. **Figures 6a~6d** show present the calculated BTTP for the USGS wells sampled by Burow et al. (1999) (listed in **Figure 4**). Both the adjoint S-FDE (411a) and the adjoint ADE (815a) were first calibrated to fit the measured CFC-11 age of Well B41, (the modeling of CFC ages followinged the methodology proposed by Weissmann et al. (2002)). Preliminary tests showed revealed that the simulated CCF-11 age is insensitive to the two truncation parameters, since these parameters primarily -subordination truncation parameter κ (or the temporal truncation parameter λ) mainly affects the very early time, (i.e., < 1 day) (or very late time, (i.e., > 50 yrs) times in the BTTP. The velocity field was directly resolved directly from the MODLFOW solutions of hydraulic head, and therefore, velocity was not a fitted parameter. Hence, the adjoint S-FDE (411a) now has 4 unknown parameters: the subordination index

α and ~~the~~ scaling factor σ^*, which control the climbing limb of the BTTP, and the time index γ and ~~the~~ capacity coefficient β, which ~~control~~govern the declining limb of the BTTP. The ~~competition~~ interplay between these two groups of parameters, particularly ~~(mainly~~ the two indexes,~~)~~ affects the BTTP peak, as discussed in **Section 2.3**. Here the primary ~~goal~~ objective is to ~~select~~ determine the best-fit ~~set of~~ parameters for the two indexes ~~while remaining within the known range for these two indexes which~~ defining~~e~~ super- and sub-diffusion while staying withing their established range. To represent~~capture~~ strong super-diffusion within a very coarse velocity field, such as a uniform velocity, the subordination index α ($1 < \alpha \leq 2$) should ~~be close to~~approach the lower limit~~end~~. ~~(f~~For example, the MADE-1 site ~~has~~utilized a best-fit $\alpha = 1.1$ with~~hen~~ a uniform, upscaled velocity ~~is used);~~. ~~similarly~~Conversely, ~~for~~ when modeling strong sub-diffusion with a uniform velocity, the time index γ ($0 < \gamma \leq 1$) ~~needs to be close to~~should approach the lower end. ~~end (f~~For example, the MADE-1 site ~~has~~ had a best-fit $\gamma = 0.39$~~)~~. With the availability of a fine-~~an increase in the~~ resolution ~~of~~ velocity field, values of α (or γ) increase and may approach the upper limit of 2 (or 1) if velocity is resolved at the pore-scale. The fine-resolution velocity field available for KRAA allowed for the selection of α and γ close to their upper ends in trial-and-error calibrations, leading to the following best-fit results: the subordination index $\alpha = 1.90$, the scaling factor $\sigma^* = 0.2$ m^{-1}, the time index $\gamma = 0.80$, and the capacity coefficient β=0.2 day$^{\gamma-1}$. For the adjoint ADE, the ~~only~~ sole fitting parameter is dispersivity, ~~and~~ with the best-fit isotropic dispersivity (longitudinal and transverse dispersivities α_L and α_T) ~~is~~ of 0.04 m. This same value of isotropic dispersity was also applied ~~by~~ in previous studies ~~for~~ modeling KRAA transport processes using ADE based models by Weissmann et al. (2002, 2004) and Zhang ~~(~~et al.~~,~~ (2018b). These studies~~,~~ ~~who~~ found that (i) simulation results were ~~not~~ insensitive to the value of α_L, as ~~(because~~ plume spreading is mainly controlled by the hydrofacies-scale heterogeneity captured by the geostatistical model~~)~~, and (ii) the Lagrangian solver ~~ran faster~~operated more efficiently with ~~for~~ isotopic dispersivity.

The best-fit parameters were then applied to predict the CFC-11 age for the other wells. The CFC-11 age calculated by the adjoint S-FDE matched~~s~~ the observed age better than the adjoint ADE for all wells under consideration~~ed here~~. ~~The BTTP simulated by the~~ The adjoint ADE produced BTTPs with ~~exhibits~~ multiple or secondary peaks, often ~~whose locations can differ~~deviating significantly from the measured CFC-11 ages. In contrast, ~~The~~ the adjoint S-FDE~~, however, usually~~ typically ~~shows~~ generated a single BTTP peak ~~in the BTTP which is~~ closer to the true CFC-11 age, ~~which may provide a convenient~~simplifying the interpretation of ~~the~~ environmental tracer dating:~~:~~ th~~eis tracer-dated~~ apparent age determined from the tracer data ~~is~~ usually fell within ~~located around (i.e., in~~ the range of the 25th to 75th percentiles of~~)~~ the BTTP peak. In addition, **Figure 6e** shows the joint BTTP for all wells, ~~which~~ representing~~s~~ ~~the~~ groundwater recharge times for all four wells simultaneously. The joint BTTP, depicted in a log-log plot (**Figure 6j**), ~~is~~ exhibited narrower uncertainty compared to ~~than each~~individual marginal BTTPs. This reduction in uncertainty results from~~,~~ ~~because~~ the availability of concentration data ~~the uncertainty (in recovering the pollutant release history) decreases when concentration data~~ from multiple observation wells ~~are available~~. ~~Notably~~Importantly, this represents ~~is~~ the first validated large-scale transport model that combines non-local super/sub-diffusion and local velocities. This application ~~proves~~ confirms the suitability ~~applicability~~ of the adjoint S-FDE

(4~~11~~a) and its Lagrangian solver ~~in~~ for capturing BTTP in a 3-*d*, regional-scale, nonstationary alluvial aquifer with a fine-resolution velocity field.

3.3 BTTP application case 3: Recover the release time for tracers in Red Cedar river, Michigan

Phanikumar et al. (2007) conducted a study involving the release~~d~~ of fluorescein dye into the Red Cedar River (RCR), a fourth-order stream in Michigan, US. They~~,~~ ~~and~~ then measured ~~the~~ breakthrough curves (BTCs) at three locations with travel distances of 1.4 km, 3.1 km, and 5.08 km, respectively, to explore the impact of river system retention on dissolved chemicals. The resulting~~ant~~ BTCs were fitted by Chakraborty et al. (2009) using ~~the~~a standard, 1-*d* space FDE with a constant velocity. The choice of a 1-*d* model was appropriate ~~applicable because of~~due to the relatively straight nature of the river reach. However, ~~S~~since sub-diffusion was found in this stream (Phanikumar et al., 2007) (likely due to open channel retention and/or hyporheic exchange) and the space FDE cannot account for~~model~~ sub-diffusion, we applied~~y~~ the more versatile backward FDE (7~~14~~a). This model ~~, which is a more general model (~~encompasses ~~containing~~ both space and time fractional derivatives~~)~~ and ~~than the adjoint of the standard space FDE,~~offers a solution to predict the tracer release time.

We first estimated the seven parameters in the 1-*d* adjoint S-FDE (7~~14~~a) using ~~from~~ the tracer data. The tracer BTCs measured by Phanikumar et al. (2007) displayed characteristic behaviours, including ~~all exhibited~~ an exponential mass increase in the BTC's ~~climbing~~ ascending limb and ~~fast~~ rapid mass decrease in the descending~~clining~~ limb. These behaviours~~,~~ suggest~~implying~~ Fickian diffusion in the operational time (meaning that the subordination index α is close to 2 and the spatial truncation parameter κ is negligible) and weak solute retention (so that the time index γ should be large, and we initially tried~~selected~~ γ =0.9~~ as the initial try~~). The capacity coefficient β should be small, considering the high ~~~90% of the~~ mass recovery rate in the field (approximately 90%) (Phanikumar et al., 2007), and hence we approximated $\beta = 0.08$ minute$^{1-\gamma}$ (representing 90% of mobile mass recovery). The temporal truncation parameter λ (=0.034 minute^{-1}) was approximated by the reverse of the time interval from the BTC peak to the inflection point of the BTC slope, as shown by Zhang et al. (2022). The mean velocity V (=0.0317 km/minute) was estimated by the speed of the BTC peak moving from the 1st sampling location (L=1.4 km) to the 2nd one (L=3.1 km). The last parameter, dispersion coefficient D^* ($= \sigma^* V$), was estimated to be 0.00317 km^2/minute by assuming that dispersion is one order of magnitude smaller than advection, ~~(~~since solute transport in rivers is usually dominated by advection~~)~~. These ~~rough~~ estimations, while inherently ~~contain high~~ uncertain~~t,y,~~ ~~but they~~served to simplify ~~significantly simplify~~ the ~~field~~ application of a complex model ~~containing 7~~with seven unknown parameters in the field.

The peak of the predicted flux-concentration-~~-~~based BTTPs using the 1-*d* adjoint S-FDE (7~~14~~a) ~~can~~ captures the true release time for ~~the~~ stream gauges located at L=3.1 km (gauge #2) and 5.08 km (gauge #3) (shown by the red solid line in **Figures 7b** and **7c**). However, ~~, although~~ it slightly underestimates the true release time for gauge #1 located at L=1.4 km (**Figure 7a**). ~~(t~~This ~~deviation~~ discrepancy arises~~is~~ because the velocity was estimated ~~using the~~ based on transport data for tracers passing gauge #1~~)~~. For comparison~~ purposes~~, we also employed the adjoint ADE model~~ is also used here:~~. ~~when~~ When using the same values of V (=0.0317 km/minute) and D^*(=0.00317 km^2/minute)~~, as those in the S-FDE are used,~~ the adjoint ADE model consistently underestimates ~~all~~ the true release times for all gauges (illustrate by ~~see~~ the black solid line in **Figure 7**). Attempts to ~~We then~~ fit V and D^* for the first gauge ~~by~~ to match~~ing~~ the true release time for tracers captured at gauge #1 still result in ~~, but~~ the adjoint ADE model ~~then~~ underestimating~~es~~ the true release time for tracers captured at gauges #2 and #3. Therefore, the adjoint S-FDE (7~~14~~a) proves to be a ~~is~~ more suitable ~~appropriate~~choice than the classical adjoint ADE for recovering pollutant release history in this river with a constant velocity.

It is also noteworthy that the BTTP for the immobile phase sources ~~has~~ exhibits ~~the~~ a similar peak time and tailing behaviour ~~as~~ to that~~ose~~ ~~in~~ of the BTTP for the mobile phase sources (**Figure 7**). This similarity ~~is due to~~arises from the weak solute retention, as ~~captured~~ indicated by the large time index γ (~~meaning~~ resulting in a relatively narrow distribution of the waiting time PDF), the small capacity coefficient β (~~meaning~~ indicating a smaller ~~portion~~ fraction of immobile pollutants at equilibrium), and the relatively large time truncation parameter λ (~~meaning~~ indicating that pollutant transport approaches Fickian scaling once time exceeds $\frac{1}{\lambda} \approx 32$ minutes). This ~~result differs from~~contrasts with ~~that~~ the findings for ~~found for~~ the MADE aquifer discussed in **section 3.2**, ~~implying~~ suggesting a more pronounced ~~stronger~~ sub-diffusion in regional-scale alluvial aquifer/aquitard systems compared to~~than~~ rivers.

4. Discussion: Extension of field applications and model capabilities

The adjoint subordination approach developed and applied above can also ~~be used to~~help identify the pollutant source location, ~~which plays~~ a ~~crucial~~ critical factor ~~role~~ in pollution source control and water resource management. Furthermore, ~~The~~ the backward-in-time vector model (4~~11~~a) ~~may also~~has the potential for extension to address ~~be extended to a more general form for~~ more complex transport scenarios. These ~~possible~~ potential extensions are discussed in the following two subsections.

4.1 Identify pollutant source location using backward location probability density function (BLP)

Pollutant source location identification has remained ~~a hot~~an important topic in hydrology for more than two decades, ~~as~~ extensively reviewed by Atmadja and Bagtzoglou (2001), Chadalavada et al. (2011), and Moghaddam et al. (2021). Process-based and statistical models had also been developed in the last two years to successfully identify pollutant source in groundwater and rivers. These models~~,~~ including~~e~~ genetic algorithms combined with groundwater models (Han et al., 2020; Habiyakare et al., 2022) or optimization models (Ayaz et al., 2022), ~~the~~ modified export coefficient models ~~combined~~ integrated with SWAT (Guo et al., 2022), physical/stochastic inverse models (Moghaddam et al., 2021), isotope mixing models (Wiegner et al., 2021; Ren et al., 2021), deep learning models (Kontos et al., 2021; Pan et al., 2021), the model-based backward probability method (Khoshgou and Neyshabouri, 2022), and the Null space Monte Carlo stochastic model (Pollicino et al., 2021), among ~~many~~ others~~ models~~.

The adjoint S-FDE (411) provides introduces a new process-based modeling approach in to pollutant source location identification by. It calculateomputings a backward location probability density function (BLP), (which is analogous to the normalized resident concentration at a previous time.), where tThe peak of this BLP defines the most probablelikely point source location. The term "BLP" represents a standard backtracking scheme, adhering to the established standard procedure for calculating particle number density-based PDFs in space. As shown in **section 3**-, where wein recoverededing pollutant release history, the adjoint S-FDE (411) may offers potential improvementse over the classical process-based pollutant source identification models. It can by (i) identifying the source location for pollutants undergoing non-Fickian diffusion, (including super-diffusion, sub-diffusion, their mixturecombination, and any intermediate transitions from between non-Fickian to and Fickian diffusion), (ii) distinguishing the initial source phase, and (iii) incorporating accommodate flow fields with varying ious resolutions. We will check validate this hypothesis using real-world data below.

4.1.1 BLP application case 1: SHOAL test site

The adjoint S-FADE (411a) was first applied to pinpointidentify the point tracer source at the SHOAL test site in, Churchill County, central Nevada, US. At this site, A radial tracer test was conducted by Reimus et al. (2003) conducted a radial tracer test in a saturated, fractured granite located at the SHOAL siteformation. Although The the detailed fracture configuration was not available for the granite aquifer was unavailable, although researchers classified categorized the discrete fracture networks (DFNs) into three groups based on fracture aperture (small, medium, and large, according to the fracture aperture) using a stochastic approach (Pohll et al., 1999). The slow, ambient groundwater velocity in this setting was estimated to be 0.3 to 3 m/yr (Pohll et al., 1999), which was considered negligible compared to the radial flow generated by the pumping test. During the test, A total of 20.81 kg of bromide with an average concentration of 3.6 g/L was injected into an injection well located 30 m from the extraction well. The measured tracer BTC exhibited both power law tails at both early time and late times power law tails, although the late time BTC data was insufficienttoo short to reveal determine the full extent of exhaustive mass decline (depicted by see symbols in **Figure 8**).

We applied MODFLOW to calculate the steady-state flow, by simplifyingapproximating the intricatecomplex velocity field as a radial flow with an average pumping rate of $Q = 12.4$ m^3/day, consistent with (the same value used for the SHOAL field test). For the sake of upscaling, we The simplified the aquifer as "homogeneous," aquifer (selected here for the purpose of upscaling) hasfeaturing an average K of 5.78×10^{-6} m/s, falling (inwithin the range of the bulk hydraulic conductivity, which was 1.48×10^{-6} ~ 4.7× 10^{-5} cm/s, measured by Pohll et al. (1999)). We then applied Tthe vector S-FDE (1a) with a convergent flow field was then applied to match fit the observed bromide BTC. **Figure 8** compares the measured and fitted bromide BTCs. The best-fit parameters in the S-FDE model (1a) are as follows: the time index $\gamma = 0.44$ (without truncation), the capacity coefficient $\beta = 0.48$ d$^{\gamma-1}$, the subordination index $\alpha = 1.95$, the scalar factor $\sigma^* = 1.0$, the truncation parameter $\kappa = 1.3 \times 10^{-3}$ m^{-1}, and the molecular diffusion coefficient $D^* = 1.0 \times 10^{-5}$ m^2/d. In **Figure 9**, we display The the resultingant 2-*d* forward-in-time plume snapshots (along in the horizontal plane) using model (1a) are plotted in **Figure 9** at both early

~~time~~ ($t = 2$ d) and late times ($t = 200$ d) for all phases (~~the~~ mobile, immobile, and total phases). The simulated fractional mass recovery for tracer bromide at the final ~~last~~ sampling cycle ($t = 322$ d) ~~for the tracer bromide was~~reached 20.2%, which is close to the recovery ratio (18.0%) estimated by Reimus et al. (2003).

The resulting~~ant~~ backward streamlines, computed using the adjoint S-FDE (4~~11~~a), are perpendicular to the groundwater head contour (**Figure 10a**), ~~validating~~ confirming the validity of the concept of subordination to regional flow and our Lagrangian solver.~~:~~ This demonstrates that particles ~~should~~ move backward along streamlines, ~~to~~ effectively describing~~e~~ ~~the~~ backward mechanical dispersion. The simulated BLP is plotted in **Figures 10b~10d**, where the peak BLP for the mobile phase source captures the true point source location, considering ~~(note~~ that the initial point source was within the mobile phase~~)~~.~~,~~ ~~while~~ In contrast, the peak BLP for the immobile phase source ~~stays~~ lags behind and is closer to the pumping well ~~(~~due to strong retention.~~)~~ Notably, the divergence in backward flow can disperse particles to different locations, leading to multiple potential sources. Therefore, ~~The~~ the adjoint S-FDE (4~~11~~a) and its Lagrangian solver, as developed in **section 2**, ~~therefore, can calculat~~can calculate~~e~~ ~~the~~ BLP for~~for~~ a divergent flow field in a 2-*d* fractured aquifer.

4.1.2 BLP application case 2: KRAA

We then applied~~y~~ the adjoint S-FDE (4~~11~~a) to calculate BLP for non-point pollutant sources ~~for~~ within the KRAA aquifer. **Figure 11a** shows the ~~resultant~~ resulting BLP for W~~w~~ell B51, representing ~~. Here the BLP captures properties (i.e.,~~the locations and weights~~)~~ of non-point source pollutants reaching ~~from the water table that can reach~~ Well 51 ~~during the last~~over the past 200 yrs. This BLP~~It~~ can also ~~be adopted~~serve as the well-head protection zone ~~(~~under ~~the~~ ambient flow conditions, i.e., without pumping~~)~~. To ~~explore~~ assess ~~the~~ BLP sensitivity ~~of BLP~~ to the well depth, we modeled a deeper well named "5b", located ~~(which is~~ 14.0 m deeper~~, right~~ below Well 51, and ~~) was also modeled, with~~ the resulting~~ant~~ BLP is shown ~~plotted~~ in **Figure 11b**. The BLP for Well 5b ~~exhibits~~ indicates a ~~relatively closer~~ source center relatively closer to the well than that for Well 51, suggesting ~~implying~~ the presence~~existence~~ of preferential flow paths within~~at~~ the deeper aquifer ~~which~~ that ~~can be captured by~~ the adjoint S-FDE (~~11a~~4a) can capture. **Figure 11c** presents~~shows~~ the joint BLP for both wells 51 and 5b, identifying ~~the~~ locations ~~for~~ where non-point source pollutants ~~that~~ can potentially contaminate both wells. For comparison ~~purposes~~, we also calculated ~~the~~ BLP ~~was also calculated by~~using the adjoint ADE, which covers a larger area, ~~especially~~ particularly near the monitoring well (**Figure 11d**). This expansion is~~, which is most~~ likely due to the ~~strong~~ substantial transverse (vertical) dispersivity ($\alpha_T = 0.04$ m) mentioned in~~(see~~ **section 3.2**~~)~~. As ~~With the increase of~~ well depth increases, the center of ~~the~~ related pollutant sources shifts~~moved~~ further upstream (**Figure 11e**). Overall, most of the ~~the majority of~~ BLP calculated by the adjoint S-FDE (4~~11~~a) ~~is located inside of~~falls within the BLP ~~calculated~~ determined by the adjoint ADE (**Figure 11g**). This suggests~~, implying~~ that the adjoint S-FDE (4~~11~~a) tends to reduce the uncertainty in pollutant source identification by emphasizing the impact of dominant flow paths, ~~(~~including ~~the~~ preferential flow paths,~~)~~ on regional-scale pollutant transport. Furthermore, ~~T~~this ~~also~~ explains why the BLP calculated by the adjoint S-FDE extends~~ed~~ slightly further upstream than that of the adjoint ADE, as ~~(because~~ the adjoint S-FDE captures super-diffusive, large-scale jumps~~)~~.

4.2 Extension to multi-scaling subordinated model

The backward-in-time vector model (411a) has two main limitations. Firstly, it requires up torelies on up to seven parameters, the predictability of whichose predictability remains a challenge. This study provided conducted preliminary tests for model parameter estimation (in **sections 3** and **4**), and further research on parameter predictability for fractional-derivative models can be found in Zhang et al. (2022). More Additional efforts are still needednecessary in future studies to improve enhance the predictability of FDEs.

Secondly, the subordination index α and scaling factor σ^* in model (411a) are limited to constant values, while whereas pollutant plumes in natural geological media may exhibit non-uniform, super-diffusive spreading rates. As a preliminary test, here we propose the following multi-scaling subordination model as a possible extension of (411a), by adoptingincorporating the multi-scaling fractional derivative concept proposed by Meerschaert et al. (2001):

$$b\frac{\partial(\theta A)}{\partial s}+\beta\frac{\partial^{\gamma,\lambda}(\theta A)}{\partial s^{\gamma,\lambda}}=\nabla_{\overleftarrow{V}}(\theta A)-\theta\left(\nabla_{\overleftarrow{V}}\right)^{\mathbf{H}(\overleftarrow{V})^{-1}}_{M(\overleftarrow{V})}A-(q_I+\theta r)A+\frac{\partial h}{\partial C}\ , \qquad (916)$$

where $M(\overleftarrow{V})$ denotes the mixing measure which defines the (rescaled) probability offor particles movementing along in each direction of the vector velocity $\overleftarrow{V}$, and $\mathbf{H}(\overleftarrow{V})^{-1}$ denotes represents the inverse of the scaling matrix which defines the subordination index (with tempering) along the water flow direction of $\overleftarrow{V}$. When $M(\overleftarrow{V})$ is remains constant (i.e., reduces to the constant σ^*) and the matrix $\mathbf{H}(\overleftarrow{V})^{-1}$ also reduces to a constant $\bar{\alpha}$ (with the truncation parameter κ) along in all directions, the multi-scaling adjoint S-FDE (916) reduces to the unique-scaling model (411a).

The general model (916) allows accommodates direction-dependent scaling rates, enabling thefor captureing of multi-dimensional transport in complex media such aslike regional-scale fractured mediasystems. This function is similar toresembles the multi-scaling adjoint fractional-derivative model derived by Zhang (2022):

$$b\frac{\partial(\theta A)}{\partial s}+\beta\frac{\partial^{\gamma,\lambda}(\theta A)}{\partial s^{\gamma,\lambda}}=\nabla\cdot\left(\theta\vec{V}A\right)-\theta D\,\nabla^{\bar{\mathbf{H}}^{-1}}_{\bar{M}(d\theta)}A-(q_I+\theta r)A+\frac{\partial h}{\partial C}\ , \qquad (1017)$$

where the mixing measure $\bar{M}(d\theta)=M(d\theta+\pi)$ is reversed for each discrete angle $d\theta$ for backward particle jumps, and the corresponding scaling matrix $\bar{\mathbf{H}}$ is also reversed by π along each eigenvector direction. The multi-scaling adjoint FDE (1017) is applicablees for to a space-dependent velocity vector $\vec{V}$, where the spreading angles and weights in the mixing measure $\bar{M}(d\theta)$ can change with velocity. The computational burden of model (1017), however, increases with an increasinghigher flow resolution of the flow field,. This is because the particle displacement during each jump event needs tomust be dividedseparated into multiple sections and then projected into anthe adjacent streamline deviating with the angle of $d\theta+\pi$ from the starting velocity vector. This process, known (which can be calledas the streamline projection method with non-zero projection angles), as was demonstrated by Zhang (2022). This It can lead toresult in prohibitive computational burden for a regional-scale aquifer with complex flow, such as the KRAA site. To overcome this challenge, The the multi-scaling adjoint S-FDE (916) solves this challenge usingemploys the streamline-orientation approach, meaning that there is noeliminating the need forto a deviatione by an angle of $d\theta+\pi$ because mechanical dispersion follows the streamlines.

Here we first check validate the Lagrangian solution of model (916) for using a astraightforward scenario simple case where the other solution is availablewith an existing alternative solution. **Figure 12c** shows the Lagrangian solution of the multi-scaling S-FDE, based on given the mixing measure (with divergent flow) and the scaling matrix (with a constant index) shown depicted in **Figure 12b**. This scenario may characterizes define pollutant transport in a discrete fracture network (DFN) with multiple orientations (**Figure 12a**). The Lagrangian solution matches well Nolan's (1998) multivariate stable distribution (**Figure 12d**).

Next, we apply model (916) to track pollutant transport in a 2-*d* DFN. **Figure 13a** shows the ensemble average of plume snapshots at time t=4.6 yrs obtained from Monte Carlos simulations of pollutant transport in 100 DFNs generated by Reeves et al. (2008). These , where the DFNs exhibithas multiple orientations, and theleading to plume therefore movements along in various directions. The best-fit solution using the forward-in-time, multi-scaling S-FDE is shown in **Figure 13c**, which caneffectively capturinge theplume fingering of the plume dueattributed to super-diffusion along fractures. For comparison purposes, we also apply the multi-scaling FDE proposed by Zhang (2022) to capture the plume snapshot (**Figure 13b**), which is similar toclosely resembles that of the multi-scaling S-FDE results. These best-fit parameters are then applied to predict plume snapshots at two subsequentlater –time points. It is noteworthy that Tthe multi-scaling S-FDE slightly can captureoutperforms the multi-scaling FDE in capturing the plume's center density and rear edge, slightly better than the multi-scaling FDE (as evidenced by see for example, **Figures 13f** vs. **13g** and **Figures 13j** vs. **13k**, respectively). The peak of the corresponding BLP calculated by the multi-scaling adjoint S-FDE (916) (where the reflective boundary conditions is are used for alleach boundaries due to y, since nothe absence of pollutants recharge from the outside) can capture the true point source location. (nNotably,e that the plume center appears to remain did not move relatively stationary apparently downstream, due to strong matrix diffusion effects). Additional dDetails regarding of model parameter estimation for the DFNs can be found in Zhang (2022). This application shows that the multi-scaling adjoint S-FDE (169) can conveniently identify the pollutant source location in DFNs characterized by with a uniform, upscaling velocity vector.

5. Conclusion

To reliably track pollutants in natural water flow systems, this study derived the adjoint of the time-fractional nonlocal transport model subordinated to regional flow, developed a completethe fully Lagrangian solver, and then applied thise new approach to track trace pollutants undergoing experiencing non-Fickian transport in surface water and groundwater with various differing velocity resolutions. Through mMathematical analysis and real-worldpractical hydrologic applications, four key conclusions have emerged revealed the following four main conclusions.

First, the adjoint subordination approach led tyieldedo an adjoint S-FDE model for quantifying backward probabilities, which takes subordination to the reversed regional flow, converts the forward-in-time boundary conditions, and reverses inverts the tempered α-stable density for mechanical dispersion. The resultingant backward-in-time boundary conditions can either capture the outsideexternal pollutant sources using the absorbing/free boundary or exclude themany out-of-domain

~~pollutant sources using~~ with the fully reflective boundary, both of which were tested ~~(all of these boundary conditions were tested~~ in applications~~)~~. The adjoint α-stable density, ~~(~~with tempering,~~)~~ reverses ~~its~~ skewness to describe backward, super-diffusive particle ~~large~~ displacements ~~of particles~~ along preferential flow paths,~~;~~ which is combined with the self-adjoint time fractional derivative term in the model ~~(for describing sub-diffusion)~~ to capture a ~~wide range~~broad spectrum of non-Fickian transport dynamics. In addition, the corresponding Lagrangian solver is computationally efficient as it can simply reverse ~~because~~ streamlines to track backward super-diffusive mechanical dispersion of particles ~~can be tracked by simply reversing streamlines~~.

Second, in real-world applications, ~~showed that~~ the adjoint S-FDE reliably tracked pollutants ~~moving~~ in surface water and groundwater ~~with~~ across various velocity resolutions ~~of velocity~~. The ~~new~~ model successfully recovered pollutant~~the~~ release history and identified pollutant source~~the~~ location(s) ~~of pollutant source(s) for water~~in systems characterized by~~with a~~ uniform velocity, ~~a~~ non-uniform flow fields (i.e., divergent/convergent flow), and fine-resolution velocities in a non-stationary, regional-scale alluvial aquifer. ~~In t~~These scenarios often ~~simplified or well-characterized~~exhibited ~~flow fields,~~ non-Fickian dynamics, especially sub-diffusion, ~~(~~influenced by~~due to for example~~ solute retention, hyporheic exchange, or matrix diffusion. In such cases, the adjoint S-FDE outperformed the classical ADE based backward models~~)~~ ~~were ubiquitous and affecting pollutant transport processes, and the adjoint S-FDE performed better than the classical ADE based backward models~~ in calculating BTTP and BLP.

Third, caution regarding the pollutant source phase is needed when backtracking pollutants in natural geologic media. For example, in alluvial aquifers characterized by strong sub-diffusion due to typically abundant aquitard materials, the mobile phase pollutant source ~~can~~ may exhibit a ~~much~~ significantly shorter release time and appear an apparently further source location ~~than~~ compared to the immobile phase source ~~in alluvial aquifers where sub-diffusion is typically very strong due to the usually abundant aquitard materials~~. However, for large-scale transport in rivers with weak solute retention, ~~The~~ the distinction between mobile and ~~-~~immobile pollutant source phases ~~distinction, however,~~ may be less significant. ~~neglected for large-scale transport in rivers with weak solute retention.~~ While many ~~Field~~ field tracer tests (including those revisited in this study) usually ~~had~~ involve a mobile initial phase, ~~but~~ real-world applications may also encompass ~~involve~~ immobile pollutant sources (such as DNAPL), where the method proposed in this study ~~may~~ can be applied.

Fourth, field applications of the adjoint S-FDE face~~are~~ challenges~~d~~ ~~by~~ related to the ~~poor~~ predictability of model parameters, and the model itself may ~~be~~ require extensions~~ded~~ ~~for~~ to handle more complex transport dynamics. This study offered~~provided~~ ~~simple~~ basic parameter estimations ~~for model parameters given~~based on field measurements, ~~while~~ but further~~ture~~ ~~efforts~~ research ~~are~~ is necessary ~~still needed~~ to establish a ~~link quantitatively~~ quantitative connection between model parameters ~~to~~ and media/pollutant properties. In addition, the multi-scaling adjoint S-FDE ~~may~~ presents an opportunity to ~~extend~~ expand upon the unique-scaling adjoint S-FDE and ~~simplify~~ streamline the multi-scaling adjoint FDE for~~in~~ backtracking pollutants in fractured media.

Appendix A. Derivation of the Backward Model (4)

This appendix derives the backward model for the S-FDE (1). Here we first change the position of the state sensitivity ϕ and the adjoint sate A in the first four terms of Eq. (3) shown in Sect. 2.1.1. For example, the 1st term in Eq. (3), denoted as I_1, can be re-arranged using integration by parts:

$$I_1 = \int_\Omega \left[\int_0^T Ab \frac{\partial(\theta\phi)}{\partial t} dt\right] d\Omega = \int_\Omega \left\{[Ab\theta\phi]|_{t=0}^{t=T} - \int_0^T \theta\phi b \frac{\partial A}{\partial t} dt\right\} d\Omega\,. \tag{A1}$$

The 2nd term in Eq. (3) contains the time fractional derivative and can be re-arranged using the fractional-order integration by parts (which doesn't involve vector field flux through a closed surface), as shown in Zhang (2022):

$$I_2 = \int_\Omega \left[\int_0^T A\beta \frac{\partial^{\gamma,\lambda}(\theta\phi)}{\partial t^{\gamma,\lambda}} dt\right] d\Omega = \int_\Omega \left\{A|_{t=T}\, \beta\, I_+^{1-\gamma,\lambda}(\theta\phi)|_{t=T} - [\theta\phi]|_{t=0}\, \beta\, I_-^{1-\gamma,\lambda}(A)|_{t=0} + \int_0^T \theta\phi\beta \frac{\partial^{\gamma,\lambda} A}{\partial(-t)^{\gamma,\lambda}} dt\right\} d\Omega, \tag{A2}$$

where the symbol $I_+^{1-\gamma,\lambda}(f)$ denotes the positive fractional integral of order $1-\gamma$: $I_+^{1-\gamma,\lambda}(f) = e^{-\lambda T}\int_0^T f e^{\lambda t} \frac{(T-t)^{-\gamma}}{\Gamma(1-\gamma)} dt$, the symbol $I_-^{1-\gamma,\lambda}(f) = e^{\lambda T}\int_0^T f e^{-\lambda t} \frac{t^{-\gamma}}{\Gamma(1-\gamma)} dt$ denotes the negative fractional integral of order $1-\gamma$, and $\Gamma(\cdot)$ is the gamma function.

The 3rd term in Eq. (3), which describes the net advective flux, can be re-arranged using the integer-order integration by parts:

$$I_3 = \int_0^T \left\{\int_\Omega \nabla\cdot[A\theta V\phi]\, d\Omega - \int_\Omega \theta V\phi\, \nabla A\, d\Omega\right\} dt = \int_0^T \left\{\oint_\xi [A\theta V\phi]\cdot n\, d\xi - \int_\Omega \theta V\phi\, \nabla A\, d\Omega\right\} dt, \tag{A3}$$

where in the second equality, the Gauss' divergence theorem is used: $\int_\Omega \nabla\cdot f\, d\Omega = \oint_\xi f\cdot n\, d\xi$, and n is the outward normal direction on the boundary ξ. Eqs. (A1)~(A3) are the same as those shown in Zhang (2022), which is expected since the same time fractional derivative term was used in these FDEs.

The 4th term in Eq. (3) contains the subordination operator and can be re-arranged using the integration by parts for twice, as shown in Zhang (2022):

$$I_4 = \int_0^T \left[\int_\Omega A\sigma^* (\nabla_{\vec{V}})^{\alpha,\kappa}(\theta\phi)\, d\Omega\right] dt = \int_0^T \left\{\oint_\xi \sigma^* \left[A\, I_+^{2-\alpha,\kappa}\left(\nabla_{\vec{V}}(\theta\phi)\right)\right]\cdot n\, d\xi + \oint_\xi \sigma^* \left[\nabla_{\vec{V}}(\theta\phi)\, I_-^{2-\alpha,\kappa}(A)\right]\cdot n\, d\xi\right\} dt$$
$$+ \int_0^T \left\{\oint_\xi \sigma^* \left[\theta\phi\, (\nabla_{\overleftarrow{V}})^{\overline{\alpha-1},k}(A)\right]\cdot n\, d\xi\right\} dt - \int_0^T \left\{\int_\Omega \sigma^* \theta\phi\, (\nabla_{\overleftarrow{V}})^{\overline{\alpha},k}(A)\, d\Omega\right\} dt\,. \tag{A4}$$

Here the operator $(\nabla_{\overleftarrow{V}})^{\overline{\alpha},k}$ denotes subordination to the reversed flow field ($\overleftarrow{V}$) where the tempered stable density (with order α) has the maximumly negative skewness $\beta^* = -1$, meaning that fast displacements are from downstream to upstream (for backward tracking).

Neupauer and Wilson (2001) showed that the adjoint state A is a measure of the change in concentration for a unit change in source mass M_0. In sensitivity analysis, the marginal sensitivity of a performance measure A with respect to M_0 is (Neupauer and Wilson, 2001):

$$\frac{dP}{dM_0} = \int_0^T \int_\Omega \left[\frac{\partial h(M_0, C)}{\partial C}\, \phi\right] d\Omega\, \mathrm{dt}\,, \tag{A5}$$

where $h(M_0, C)$ is a functional of the state of the system. Inserting I_1~I_4 expressed by Eqs. (A1)~(A4) into the inner product equation (3), and then subtracting this updated Eq. (3) from the marginal sensitivity equation (A5), we obtain:

$$\frac{dP}{dM_0} = \int_\Omega \int_0^T \phi \left\{ \frac{\partial h}{\partial C} + b\theta \frac{\partial A}{\partial t} - \beta\theta \frac{\partial^{\gamma,\lambda} A}{\partial(-t)^{\gamma,\lambda}} + \theta V \nabla A - \sigma^* \theta \left(\nabla_{\bar{V}}\right)^{\overline{\alpha},k}(A) - (q_o + \theta r) A \right\} d\Omega\, dt$$

$$- \int_\Omega \left\{ [Ab\theta\phi]|_{t=T} - [Ab\theta]|_{t=0} \frac{\partial C_i}{\partial M_0} + A|_{t=T} \beta I_+^{1-\gamma,\lambda}(\theta\phi)|_{t=T} - [\theta\phi]|_{t=0}\, \beta I_-^{1-\gamma,\lambda}(A)|_{t=0} \right\} d\Omega$$

$$- \int_0^T \oint_\xi \left[A\theta V\phi - A\, I_+^{2-\alpha,\kappa}\left(\nabla_{\bar{V}}(\theta\phi)\right) - \nabla_{\bar{V}}(\theta\phi)\, I_-^{2-\alpha,\kappa}(A) - \theta\phi \left(\nabla_{\bar{V}}\right)^{\overline{\alpha-1},k}(A) \right] \cdot n\, d\xi dt. \quad \text{(A6)}$$

To eliminate ϕ from Eq. (A6), we define A such that the terms containing ϕ vanish. Since the double integral in Eq. (A6) (shown by the first line in Eq. (A6)) can be eliminated when the summation of all the terms inside the bracket is zero, this produces the adjoint equation of the S-FDE (1a):

$$b\theta \frac{\partial A}{\partial t} - \beta\theta \frac{\partial^{\gamma,\lambda} A}{\partial(-t)^{\gamma,\lambda}} = -\theta V \nabla A + \sigma^* \theta \left(\nabla_{\bar{V}}\right)^{\overline{\alpha},k}(A) + (q_o + \theta r) A - \frac{\partial h}{\partial C}. \quad \text{(A7)}$$

Assuming (i) the backward time $s = T - t$ where T is the detection time, (ii) steady-state groundwater flow (so that $\theta V\, \nabla A - q_o A = \nabla(\theta V A) - q_I A$), and (iii) un-compressible aquifer skeleton (so that $\partial\theta/\partial t = 0$), we can re-write Eq. (A7) as Eq. (4) listed in Sect. 2.1.1, which is the adjoint of the S-FDE (1) listed in Sect. 2.1.1.

Data availability

Data for BTTP Application 1 are available from the published paper Benson et al., Transport in Porous Media, 2001 at https://link.springer.com/article/10.1023/A:1006733002131. Groundwater age data using CFC-11 are available online from the reference Burow et al., U.S. Geol. Surv. Water Resour. Invest., 1999. SHOAL test site data are available from the published paper Reimus et al., Water Resour. Res. (2003) at https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2002WR001597. The discrete fracture network data are available from the published paper Reeves et al., Water Resour. Res. (2008) at https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008WR006858. All the numerical data are available from the Zendo repository (Yong Zhang, 2022).

Author contributions

YZ led the investigation, conceptualized the research, did the formal analysis, supervised the project, and wrote the initial draft. HGS acquired the funding and the resources. All co-authors reviewed and edited the paper.

Competing interests

The contact author has declared that neither they nor their co-authors have any competing interests.

Acknowledgments

WW was partially funded by the National Natural Science Foundation of China (Grant number 41931292). HGS was partially funded by the National Natural Science Foundation of China (Grant numbers U2267218 and 11972148). YZ was partially funded by the Department of the Treasury under the Resources and Ecosystems Sustainability, Tourist Opportunities, and Revived Economies of the Gulf Coast States Act of 2012 (RESTORE Act). The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the Department of the Treasury or ADCNR. This paper does not necessary reflect the view of the funding agencies.

References

Abelson, P. H.: Groundwater contamination, Science, 224(4650), 673–673, https://doi.org/doi:10.1126/science.224.4650.673, 1984.

Adams, E. E. and Gelhar, L. W.: Field study of dispersion in a heterogeneous aquifer: 2. Spatial moment analysis, Water Resour. Res., 28(12), 3293–3307, https://doi.org/doi:10.1029/92WR01757, 1992.

Alijani, Z., Baleanu, D., Shiri, B., and Wu, G. C.: Spline collocation methods for systems of fuzzy fractional differential equations, Chaos Soliton. Fract., 131, 109510, https://doi.org/doi:10.1016/j.chaos.2019.109510, 2020.

Al-Qurashi, M., Rashid, S., Jarad, F., Tahir, M., and Alsharif, A. M.: New computations for the two-mode version of the fractional Zakharov-Kuznetsov model in plasma fluid by means of the Shehu decomposition method, AIMS Math., 7(2), 2044–2060, https://doi.org/doi:10.3934/math.2022117, 2022.

Atmadja, J. and Bagtzoglou, A. C.: State of the art report on mathematical methods for groundwater pollution source identification, Environ. Forensics, 2, 205–214, https://doi.org/doi:10.1006/enfo.2001.0055, 2001.

Ayaz, M., Ansari, S. A., and Singh, O. K.: Detection of pollutant source in groundwater using hybrid optimization model, Int. J. Energy Water Resour., 6(1), 81–93, https://doi.org/doi:10.1007/s42108-021-00118-4, 2022.

Baeumer, B., Benson, D. A., Meerschaert, M. M., and Wheatcraft, S. W.: Subordinated advection-dispersion equation for contaminant transport, Water Resour. Res., 37(6), 1543–1550, https://doi.org/doi:10.1029/2000WR900409, 2001.

Benson, D. A., Schumer, R., Meerschaert, M. M., and Wheatcraft, S. W.: Fractional dispersion, Levy motion, and the MADE tracer tests, Transport Porous Med., 42, 211–240, https://doi.org/doi:10.1023/A:1006733002131, 2001.

Bianchi, M. and Zheng, M. C.: A lithofacies approach for modeling non-Fickian solute transport in a heterogeneous alluvial aquifer, Water Resour. Res., 52 (1), 552–565, https://doi.org/doi:10.1002/2015WR018186, 2016.

Boano, F., Harvey, J. W., Marion, A., Packman, A. I., Revelli, R., Ridolfi, L., and Wörman, A.: Hyporheic flow and transport processes: Mechanisms, models, and biogeochemical implications, Rev. Geophys., 52(4), 603–679, https://doi.org/doi:10.1002/2012RG000417, 2014.

Boggs, J. M., Young, S. C., and Beard, L. M.: Field study of dispersion in a heterogeneous aquifer: 1. Overview and site description, Water Resour. Res. 28(12), 3281–3291, https://doi.org/doi:10.1029/92WR01756, 1992.

Burow, K. R., Panshin, S. Y., Dubrovsky, N. M., VanBrocklin, D., and Fogg, G. E.: Evaluation of processes affecting 1,2-dibromo-3-chloropropane (DBCP) concentrations in ground water in the eastern San Joaquin Valley, California: Analysis of chemical data and ground-water flow and transport simulations. U.S. Geol. Surv. Water Resour. Invest., 99–4059, 1999.

Cardone, A., Conte, D., D'Ambrosio, R., and Paternoster, B.: Multivalue Collocation Methods for Ordinary and Fractional Differential Equations, Mathematics, 10(2), 185, https://doi.org/doi:10.3390/math10020185, 2022.

Chadalavada, S., Datta, B., and Naidu, R.: Optimisation approach for pollution source identification in groundwater: An overview, Int. J. Environ. Waste Manage., 8(1-2), 40–61, https://doi.org/doi:10.1504/IJEWM.2011.040964, 2011.

Chakraborty, P., Meerschaert, M. M., and Lim, C. Y.: Parameter estimation for fractional transport: A particle-tracking approach, Water Resour. Res., 45, W10415, https://doi.org/doi:10.1029/2008WR007577, 2009.

Chen, Z., Xu, T., Gómez-Hernández, J. J., Zanini, A., and Zhou, Q.: Reconstructing the release history of a contaminant source with different precision via the ensemble smoother with multiple data assimilation, J. Contam. Hydrol., 252, 104115, https://doi.org/doi:10.1016/j.jconhyd.2022.104115, 2023.

Cornaton, F. and Perrochet, P.: Groundwater age, life expectancy and transit time distributions in advective-dispersive systems: 1. Generalized reservoir theory, Adv. Water Resour., 29(9), 1267–1291, https://doi.org/doi:10.1016/j.advwatres.2005.10.009, 2006.

Diethelm, K., Ford, N. J., and Freed, A. D.: A predictor-corrector approach for the numerical solution of fractional differential equations, Nonlinear Dynam., 29(1), 3–22, https://doi.org/doi:10.1023/A:1016592219341, 2002.

Duan, J. S., Rach, R., Baleanu, D., and Wazwaz, A. M.: A review of the Adomian decomposition method and its applications to fractional differential equations, Commun. Frac. Calc., 3(2), 73–99, https://doi.org/doi:d1wqtxts1xzle7.cloudfront.net/46811581, 2012.

Feller, W.: An Introduction to Probability Theory and Its Applications, vol. II, 2nd ed., John Wiley, N. Y., 1971.

Fogg, G. E., LaBolle, E. M., and Weissmann, G. S.: Groundwater vulnerability assessment: Hydrogeologic perspective and example from Salinas Valley, California, In: Assessment of Non-Point Source Pollutant in the Vadose Zone, edited by Dennis L. Corwin, Keith Loague, and Timothy R. Ellsworth. Geophysical Monography-American Geophysical Union, 108, pp. 45–61, 1999.

Ford, N. J. and Simpson, A. C.: The numerical solution of fractional differential equations: speed versus accuracy, Numer. Algorithms, 26(4), 333–346, https://doi.org/doi:10.1023/A:1016601312158, 2001.

Garrappa, R.: Numerical solution of fractional differential equations: A survey and a software tutorial, Mathematics, 6(2), 16, https://doi.org/doi:10.3390/math6020016, 2018.

Green, C. T., Zhang, Y., Jurgens, B. C., Starn, J. J., and Landon, M. K.: Accuracy of travel time distribution (TTD) models as affected by TTD complexity, observation errors, and model and tracer selection, Water Resour. Res., 50(7), 6191–6213, https://doi.org/doi:10.1002/2014WR015625, 2014.

Green, C. T., Jurgens, B. C., Zhang, Y., Starn, J. J., Singleton, M. J., and Esser, B. K.: Regional oxygen reduction and denitrification rates in groundwater from multi-model residence time distributions, San Joaquin Valley, USA, J. Hydrol., 543, 155–166, https://doi.org/doi:10.1016/j.jhydrol.2016.05.018, 2016.

Gorelick, S. M., Evans, B. E., and Remson, I.: Identifying sources of groundwater pollution: an optimization approach, Water Resour. Res., 19(3), 779–790, https://doi.org/doi:10.1029/WR019i003p00779, 1983.

Guo, Y., Wang, X., Melching, C., and Nan, Z.: Identification method and application of critical load contribution areas based on river retention effect, J. Environ. Manage., 305, 114314,. https://doi.org/doi:10.1016/j.jenvman.2021.114314, 2022.

Guo, Z., Ma, R., Zhang, Y., and Zheng, C. M.: Contaminant transport in heterogeneous aquifers: A critical review of mechanisms and numerical methods of non-Fickian dispersion, Sci. China Earth Sci., 64(8), 1224–1241, https://doi.org/doi:10.1007/s11430-020-9755-y, 2021.

Haggerty, R., McKenna, S. A., and Meigs, L. C.: On the late-time behavior of tracer test breakthrough curves, Water Resour. Res., 36(12), 3467–3479, https://doi.org/doi:10.1029/2000WR900214, 2000.

Han, K., Zuo, R., Ni, P., Xue, Z., Xu, D., Wang, J., and Zhang, D.: Application of a genetic algorithm to groundwater pollution source identification, J. Hydrol., 589, 125343, https://doi.org/doi:10.1016/j.jhydrol.2020.125343, 2020.

Hansen, S. K. and Berkowitz, B.: Modeling non-Fickian solute transport due to mass transfer and physical heterogeneity on arbitrary groundwater velocity fields, Water Resour. Res., 56(10), e2019WR026868, https://doi.org/doi:10.1029/2019WR026868, 2020.

Harbaugh, A. W.: MODFLOW-2005, The U.S. Geological Survey modular groundwater model–The Ground-water flow process. U.S. Geol. Surv. Tech. Methods, 6-A16, 253 pp, 2005.

Huang, C., An, N., and Chen, H.: Local H1-norm error analysis of a mixed finite element method for a time-fractional biharmonic equation, Appl. Numer. Math., 173, 211–221, https://doi.org/doi:10.1016/j.apnum.2021.12.004, 2022.

Isaacson, E. and Keller, H. B.: Analysis of Numerical Methods. Wiley, New York, pp. 153–156, 1966.

Jamshidi, A., Samani, J. M. V., Samani, H. M. V., Zanini, A., Tanda, M. G., and Mazaheri, M.: Solving inverse problems of unknown contaminant source in groundwater-river integrated systems using a surrogate transport model based optimization, Water 12(9), 2415, https://doi.org/doi:10.3390/w12092415, 2020.

Janssen, G. M. C. M., Valstar, J. R., and van der Zee, S. E. A. T. M.: Measurement network design including travel time determinations to minimize model prediction uncertainty, Water Resour. Res., 44(2), W02405, https://doi.org/doi:10.1029/2006WR005462, 2008.

Khoshgou, H. and Neyshabouri, S. A. A. S.: Using the backward probability method in contaminant source identification with a finite-duration source loading in a river, Environ. Sci. Pollut. Res., 29(4), 6306–6316, https://doi.org/doi:10.1007/s11356-021-15372-6, 2022.

Kontos, Y. N., Kassandros, T., Katsifarakis, K. L., and Karatzas, K.: Deep Learning Modeling of Groundwater Pollution Sources. In: International Conference on Engineering Applications of Neural Networks (pp. 165-177), Springer, Cham., 2021.

LaBolle, E. M.: RWHet: Random Walk Particle Model for Simulating Transport in Heterogeneous Permeable Media, Version 3.2, User's Manual and Program Documentation. Univ. of Calif., Davis, 2006.

LaBolle, E. M., Fogg, G. E., and Eweis, J. B.: Diffusive fractionation of ^{3}H and ^{3}He in groundwater and its impact on groundwater age estimates, Water Resour. Res. 42(7), W07202, https://doi.org/~~doi:~~10.1029/2005WR004756, 2006.

Lapworth, D. J., Boving, T. B., Kreamer, D. K., Kebede, S., and Smedley, P. L.: Groundwater quality: Global threats, opportunities and realising the potential of groundwater, Sci. Total Environ., 811, 152471, https://doi.org/~~doi:~~10.1016/j.scitotenv.2021.152471, 2022.

Li, C. and Zeng, F.: Finite difference methods for fractional differential equations, Int. J. Bifurcat. Chaos, 22(04), 1230014, https://doi.org/~~doi:~~10.1142/S0218127412300145, 2012.

Lu, B. Q., Zhang, Y., Zheng, C. M., Green, C. T., O'Neill, C., Sun, H. G., and Qian, J. Z.: Comparison of time nonlocal transport modes for characterizing non-Fickian transport: From mathematical interpretation to laboratory application, Water, 10(6), 778, https://doi.org/~~doi:~~10.3390/w10060778, 2018.

Mao, R., Luo, X., Jiao, J. J., and Li, H.: Molecular diffusion and pore-scale mechanical dispersion controls on time-variant travel time distribution in hillslope aquifers. J. Hydrol., 616, 128798, https://doi.org/~~doi:~~10.1016/j.jhydrol.2022.128798, 2023.

Maxwell, R. M., Condon, L. E., Kollet, S. J., Maher, K., Haggerty, R., and Forrester, M. M.: The imprint of climate and geology on the residence times of groundwater, Geophys. Res. Lett., 43(2), 701–708, https://doi.org/~~doi:~~10.1002/2015GL066916, 2016.

Meerschaert, M. M., Benson, D. A., and Baeumer, B.: Operator Lévy motion and multiscaling anomalous diffusion, Phys. Rev. E, 63, 021112, https://doi.org/~~doi:~~10.1103/PhysRevE.63.021112, 2001.

Meerschaert, M. M. and Tadjeran, C.: Finite difference approximations for fractional advection-dispersion flow equation, J. Comput. Appl. Math., 172(1), 65–77, https://doi.org/~~doi:~~10.1016/j.cam.2004.01.033, 2004.

Meerschaert, M. M., Zhang, Y., and Baeumer, B.: Tempered anomalous diffusion in heterogeneous systems, Geophys. Res. Lett., 35(17), L17403, https://doi.org/~~doi:~~10.1029/2008GL034899, 2008.

McMahon, P. B., Burow, K. R., Kauffman, L. J., Eberts, S. M., Boehlke, J. K., and Gurdak, J. J.: Simulated response of water quality in public supply wells to land use change, Water Resour. Res., 44, W00A06, https://doi.org/~~doi:~~10.1029/2007WR006731, 2008.

Moghaddam, M. B., Mazaheri, M., and Samani, J. M. V.: Inverse modeling of contaminant transport for pollution source identification in surface and groundwaters: a review, Groundwater Sus. Dev., 15, 100651, https://doi.org/~~doi:~~10.1016/j.gsd.2021.100651, 2021.

Momani, S. and Al-Khaled, K.: Numerical solutions for systems of fractional differential equations by the decomposition method, Appl. Math. Comput., 162(3), 1351-1365, https://doi.org/~~doi:~~10.1016/j.amc.2004.03.014, 2005.

Neuman, S. P. and Tartakovsky, D. M.: Perspective on theories of non-Fickian transport in heterogeneous media, Adv. Water Resour., 32, 670–680, https://doi.org/~~doi:~~10.1016/j.advwatres.2008.08.005, 2009.

Neupauer, R. M. and Wilson, J. L.: Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant, Water Resour. Res., 35, 3389–3398, https://doi.org/~~doi:~~10.1029/1999WR900190, 1999.

Neupauer, R. M. and Wilson, J. L.: Adjoint-derived location and travel time probabilities for a multidimensional groundwater system, Water Resour. Res., 37(6), 1657–1668, https://doi.org/~~doi:~~10.1029/2000WR900388, 2001.

Neupauer, R. M., Wilson, J. L., and Bhaskar, A.: Forward and backward temporal probability distributions of sorbing solutes in groundwater, Water Resour. Res., 45, W01420, https://doi.org/~~doi:~~10.1029/2008WR007058, 2009.

Nolan, J. P.: in A Practical Guide to Heavy Tails: Statistical Techniques and Applications, Edited by R. J. Adler, R. Feldman, and M. Taqqu. Birkhauser Boston, Cambridge, MA, 1998.

Pan, Z., Lu, W., and Chang, Z.: Simultaneous identification of groundwater pollution source spatial–temporal characteristics and hydraulic parameters based on deep regularization neural network-hybrid heuristic algorithm, J. Hydrol., 600, 126586, https://doi.org/~~doi:~~10.1016/j.jhydrol.2021.126586, 2021.

Phanikumar, M. S., Aslam, I., Shen, C., Long, D. T., and Voice, T. C.: Separating surface storage from hyporheic retention in natural streams using wavelet decomposition of acoustic Doppler current profiles, Water Resour. Res., 43, W05406, https://doi.org/~~doi:~~10.1029/2006WR005104, 2007.

Phillips, C. B., Martin, R. L. and Jerolmack, D. J.: Impulse framework for unsteady flows reveals superdiffusive bed load transport, Geophys. Res. Lett., 40(7), 1328–1333, https://doi.org/~~doi:~~10.1002/grl.50323, 2013.

Pohll, G., Hassan, A. E., Chapman, J. B., Papelis, C., and Andricevic, R.: Modeling ground water flow and radioactive transport in a fractured aquifer, Ground Water, 37(5), 770–784, https://doi.org/~~doi:~~10.1111/j.1745-6584.1999.tb01170.x, 1999.

Pollicino, L. C., Colombo, L., Formentin, G., and Alberti, L.: Stochastic modelling of solute mass discharge to identify potential source zones of groundwater diffuse pollution, Water Res., 200, 117240, https://doi.org/~~doi:~~10.1016/j.watres.2021.117240, 2021.

Ponprasit, C., Zhang, Y., and Wei, W.: Backward location and travel time probabilities for pollutants moving in three-dimensional aquifers: Governing equations and scale effect, Water, 14(4), 624, https://doi.org/~~doi:~~10.3390/w14040624, 2022.

Ponprasit, C., Zhang, Y., Gu, X., Goodliffe, A. M., and Sun, H.: Assessing vulnerability of regional-scale aquifer-aquitard systems in East Gulf Coastal Plain of Alabama by developing groundwater flow and transport models, Water, 15(10), 1937, https://doi.org/~~doi:~~10.3390/w15101937, 2023.

Reeves, D. M., Benson, D. A., Meerschaert, M. M., and Scheffler, H. P.: Transport of conservative solutes in simulated fracture networks: 2. Ensemble solute transport and the correspondence to operator-stable limit distributions, Water Resour. Res., 44, W05410, https://doi.org/~~doi:~~10.1029/2008WR006858, 2008.

Reimus, P., Pohll, G., Mihevc, T., Chapman, J., Haga, M., Lyles, B., Kosinski, S., Niswonger, R., and Sanders, P.: Testing and parameterizing a conceptual model for solute transport in a fractured granite using multiple tracers in a forced-gradient test, Water Resour. Res., 39(12), 1356, https://doi.org/~~doi:~~10.1029/2002WR001597, 2003.

Ren, L., Cheng, L., Zhang, S., Ding, A., Zhu, Y., Lu, C., Li, Y., Yang, Q., and Li, J.: Quantifying nitrate pollution sources of the drinking water source area using a Bayesian isotope mixing model in the northeastern suburbs of Beijing, China, Isot. Environ. Healt. S., 57(4), 350–367, https://doi.org/~~doi:~~10.1080/10256016.2021.1937149, 2021.

Skaggs, T. H. and Kabala, Z. J.: Recovering the release history of a groundwater contaminant, Water Resour. Res., 30(1), 71–79, https://doi.org/~~doi:~~10.1029/93WR02656, 1994.

Sun, A. Y., Painter, S. L., and Wittmeyer, G. W.: A robust approach for iterative contaminant source location and release history recovery, J. Contam. Hydrol., 88(3-4), 181–196, https://doi.org/~~doi:~~10.1016/j.jconhyd.2006.06.006, 2006a.

Sun, A. Y., Painter, S. L., and Wittmeyer, G. W.: A constrained robust least squares approach for contaminant release history identification, Water Resour. Res., 42(4), W04414, https://doi.org/~~doi:~~10.1029/2005WR004312, 2006b.

Sun, H. G., Chang, A. L., Zhang, Y., and Chen, W.: A review on variable-order fractional differential equations: mathematical foundations, physical models, numerical methods and applications, Fract. Calc. Appl. Anal., 22(1), 27–59, https://doi.org/~~doi:~~10.1515/fca-2019-0003, 2019.

Vargas, A. M.: Finite difference method for solving fractional differential equations at irregular meshes, Math. Comput. Simulat., 193, 204–216, https://doi.org/~~doi:~~10.1016/j.matcom.2021.10.010, 2022.

Weissmann, G. S., Zhang, Y., LaBolle, E. M., and Fogg, G. E.: Dispersion of groundwater age in an alluvial aquifer system, Water Resour. Res., 38(10), 1198, https://doi.org/~~doi:~~10.1029/2001WR000907, 2002.

Weissmann, G. S., Zhang, Y., Fogg, G. E., and Mount, J. F.: Hydrogeologic influence of incised valley fill deposits within a stream-dominated alluvial fan. In J. Bridge, and D. W. Hyndman (ed), Aquifer Characterization, SEPM (Society for Sedimentary Geology) Special Publication, n.80, 15–28, 2004.

Wiegner, T. N., Colbert, S. L., Abaya, L. M., Panelo, J., Remple, K., and Nelson, C. E.: Identifying locations of sewage pollution within a Hawaiian watershed for coastal water quality management actions, J. Hydrol.: Regional Studies, 38, 100947, https://doi.org/~~doi:~~10.1016/j.ejrh.2021.100947, 2021.

Woodbury, A. D. and Ulrych, T. J.: Minimum relative entropy inversion: theory and application to recovering the release history of groundwater contaminant, Water Resour. Res., 32(9), 2671–2681, https://doi.org/~~doi:~~10.1029/95WR03818, 1996.

Woodbury, A., Sudicky, E., Ulrych, T. J., and Ludwig, R.: Three-dimensional plume source reconstruction using minimum relative entropy inversion, J. Contam. Hydrol. 32, 131–158, https://doi.org/~~doi:~~10.1016/S0169-7722(97)00088-0, 1998.

Wu, G. C., Kong, H., Luo, M., Fu, H., and Huang, L. L.: Unified predictor–corrector method for fractional differential equations with general kernel functions, Fract. Calc. Appl. Anal., ~~in press~~25, 648–667,~~.~~ https://doi.org/~~doi:~~10.1007/s13540-022-00029-z, 2022.

Yin, M., Zhang, Y., Ma, R., Tick, G. R., Bianchi, M., Zheng, C., Wei, W., Wei, S., and Liu, X.: Super-diffusion affected by hydrofacies mean length and source geometry in alluvial settings, J. Hydrol., 582, 124515, https://doi.org/~~doi:~~10.1016/j.jhydrol.2019.124515, 2020.

Zhang, Y.: Moments for tempered fractional advection-diffusion equations, J. Stat. Phys., 139, 915–939, https://doi.org/~~doi:~~10.1007/s10955-010-9965-0, 2010.

Zhang, Y., Green, C. T., and Baeumer, B.: Linking aquifer spatial properties and non-Fickian transport in mobile–immobile like alluvial settings, J. Hydrol., 512, 315–331, https://doi.org/~~doi:~~10.1016/j.jhydrol.2014.02.064, 2014.

Zhang, Y., Meerschaert, M. M., Baeumer, B., and LaBolle, E. M.: Modeling mixed retention and early arrivals in multidimensional heterogeneous media using an explicit Lagrangian scheme, Water Resour. Res., 51, 6311–6337, https://doi.org/~~doi:~~10.1002/2015WR016902, 2015.

Zhang, Y., Sun, H. G., Stowell, H. H., Zayernouri, M., and Hansen, S. E.: A review of applications of fractional calculus in Earth system dynamics, Chaos Soliton. Fract., 102, 29–46, https://doi.org/~~doi:~~10.1016/j.chaos.2017.03.051, 2017.

Zhang, Y., Weissmann, G. S., Fogg, G. E., Lu, B. Q., Sun, H. G., and Zheng, C. M.: Assessment of groundwater susceptibility to non-point source contaminants using three-dimensional transient indexes, Int. J. Env. Res. Pub. He., 15, 1177, https://doi.org/~~doi:~~10.3390/ijerph15061177, 2018.

Zhang, Y., Sun, H. G., and Zheng, C. M.: Lagrangian solver for vector fractional diffusion in bounded anisotropic aquifers: Development and application, Fract. Calc. Appl. Anal., 22(6), 1607–1640, https://doi.org/~~doi:~~10.1515/fca-2019-0083, 2019a.

Zhang, Y., Yu, X. N., Li, X. C., Kelly, J. F., Sun, H. G., and Zheng, C. M.: Impact of absorbing and reflective boundaries on fractional derivative models: Quantification, evaluation and application, Adv. Water Resour., 128, 129–144, https://doi.org/~~doi:~~10.1016/j.advwatres.2019.02.011, 2019b.

Zhang, Y., Brusseau, M. L., Neupauer, R. M., and Wei, W.: General backward model to identify the source for contaminants undergoing non-Fickian diffusion in water, Environ. Sci. Technol., 56(15), 10743–10753, https://doi.org/~~doi:~~10.1021/acs.est.2c01873, 2022.

Zhang, Y.: Backward particle tracking of anomalous transport in multi-dimensional aquifers, Water Resour. Res., 58, e2022WR032396, https://doi.org/~~doi:~~10.1029/2022WR032396, 2022.

Zheng, C., Bianchi, M., and Gorelick, S. M.: Lessons learned from 25 years of research at the MADE site, Ground Water, 49, 649–662, https://doi.org/~~doi:~~10.1111/j.1745-6584.2010.00753.x, 2011.

Zhou, D., Han, X., Zhang, Y., Wei, W., Green, C.T., Sun, H., and Zheng, C. M.: Co-transport of biogenic nano-hydroxyapatite and Pb (II) in saturated sand columns: Controlling factors and stochastic modelling, Chemosphere, 275, 130078, https://doi.org/~~doi:~~10.1016/j.chemosphere.2021.130078, 2021.

Zinn, B. A. and Konikow, L. F.: Effects of intraborehole flow on groundwater age distribution, Hydrogeol. J., 15(4), 633–643, https://doi.org/~~doi:~~10.1007/s10040-006-0139-8, 2007a.

Zinn, B. A. and Konikow, L. F.: Potential effects of regional pumpage on groundwater age distribution, Water Resour. Res., 43(6), W06418, https://doi.org/~~doi:~~10.1029/2006WR004865, 2007b.

Table 1. Changes of boundary conditions from the 1-*d* forward FDE (13a6a) to its backward model (14a7a).

Boundary	Forward S-FDE (1a6a)	Backward S-FDE (171a)
Left (upstream)	**Dirichlet boundary**: $C\|_{x=L} = g_1(t)$, representing a stagnant source reservoir at the inlet.	**Absorbing boundary**: $A\|_{x=L} = 0$, which can be used for groundwater age modeling (the foreword source term becomes the backward sink term).
	Neumann boundary: $-\frac{\partial^{\alpha-2}}{\partial x^{\alpha-2}}\left[\theta D \frac{\partial(e^{\kappa x}C)}{\partial x}\right]\Big\|_{x=L} = g_1(t)$, representing an immobile diffusive source located at the inlet (less common).	**Fully reflective boundary**: $\left[-V\theta A + \theta D \frac{\partial^{\alpha-1}(e^{-\kappa x}A)}{\partial(-x)^{\alpha-1}} e^{\kappa x}\right]\Big\|_{x=L} = 0$, where no particles can exist this upstream boundary; so, there are no external sources outside the upstream boundary.
	Robin boundary: $\left\{\theta V C - \frac{\partial^{\alpha-2}}{\partial x^{\alpha-2}}\left[\theta D \frac{\partial(e^{\kappa x}C)}{\partial x}\right]\right\}\Big\|_{x=L} = g_1(t)$, defining the co-existence of an advective source (located outside of the upstream boundary and moving at a constant rate V) and an immobile diffusive source (located at the upstream boundary).	**Partially reflective boundary**: $\theta D \frac{\partial^{\alpha-1}(e^{-\kappa x}A)}{\partial(-x)^{\alpha-1}} e^{\kappa x}\Big\|_{x=L} = 0$, representing a partially free exit boundary. Diffusive particles cannot exit the boundary $x = L$, but are reflected near the boundary (to capture the diffusive source at the upstream boundary); advective particles, however, can exit the boundary $x = L$ freely, to capture the advective source outside $x = L$.
	Infinite boundary: $C\|_{x=-\infty} = 0$, with both advection and dispersion contribution to the mass flux in the domain ($L < x < R$) via the upstream boundary at $x = L$.	**Free boundary**: $A\|_{x=-\infty} = 0$, for infinite domains with advective & dispersive particles freely crossing the upstream boundary at $x = L$ (also called "a fully free exit boundary").
Right (down-stream)	**Dirichlet boundary**: $C\|_{x=R} = g_2(t)$, representing a stagnant source reservoir or a mass sink term (with $g_2(t) = 0$, defining the absorption well or a groundwater barrier) at the downstream boundary.	**Absorbing boundary**: $A\|_{x=R} = 0$. A mass sink term in the forward model at the outlet transforms to a load term (with an initial probability of 1) in the backward model.
	Neumann boundary: $-\frac{\partial^{\alpha-2}}{\partial x^{\alpha-2}}\left[\theta D \frac{\partial(e^{\kappa x}C)}{\partial x}\right]\Big\|_{x=R} = g_2(t)$, representing diffusive flux leaving the system (with zero advective flux), which can define an impermeable layer at the outlet.	**Fully reflective boundary**: $\left[V\theta A - \theta D \frac{\partial^{\alpha-1}(e^{-\kappa x}A)}{\partial(-x)^{\alpha-1}} e^{\kappa x}\right]\Big\|_{x=R} = 0$, to completely close the outlet; so, no particles can exit the outlet from the internal domain and no external sources located downstream of the downstream boundary.
	Robin boundary: $\left\{\theta V C - \frac{\partial^{\alpha-2}}{\partial x^{\alpha-2}}\left[\theta D \frac{\partial(e^{\kappa x}C)}{\partial x}\right]\right\}\Big\|_{x=R} = g_2(t)$, representing both advective and diffusive flux leaving the system, due for example a pumping well.	**Partially reflective boundary**: $-\theta D \frac{\partial^{\alpha-1}(e^{-\kappa x}A)}{\partial(-x)^{\alpha-1}} e^{\kappa x}\Big\|_{x=R} = 0$. This partially reflective boundary is functionally analogous to the fully reflective boundary since the reversed flow direction, to remove any external pollutant sources.
	Infinite boundary: $C\|_{x=+\infty} = 0$, with both advection and dispersion contribution to the mass flux in the domain ($L < x < R$) via the downstream boundary at $x = R$, which is applicable for a site whose dimension is much longer than the pollutant displacement.	**Free boundary**: $A\|_{x=R} = 0$. This can be one of the predominant backward boundary conditions for real-world applications, where no physical boundaries exist or can be identified for forward pollutant transport with a limited scale in a regional-scale aquifer or river corridor.

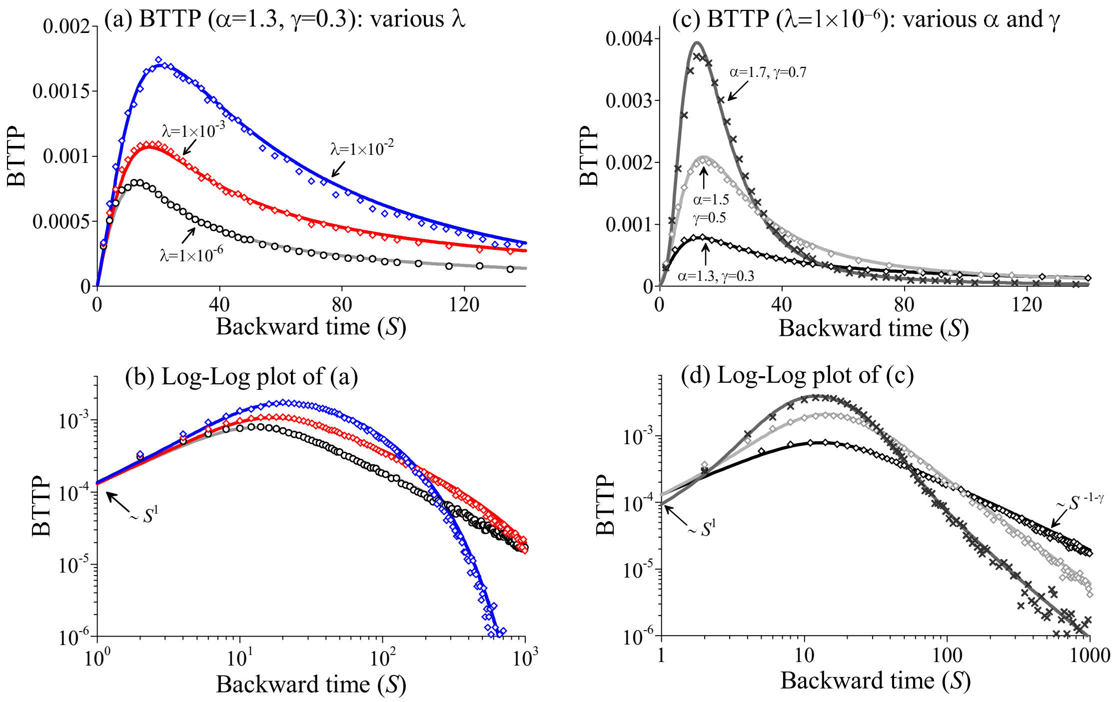


Figure 1. **Solver validation 1**: Lagrangian solutions (symbols) versus the Eulerian solutions (lines) for the 1-*d* backward model (7~~14~~a) with various truncation parameters λ (a), and various subordination index α and time index γ (c). The other model parameters that remain unchanged in these cases are as follows: velocity $V = 1$, scaling factor $\sigma^* = 1$, the spatial truncation parameter $\kappa = 1 \times 10^{-7}$, and the backward travel distance is $L = 10$. (b) and (d) are the log-log plot of (a) and (c), respectively, to show the tailing. Free exit boundary conditions are used in these cases, and parameters are dimensionless here.

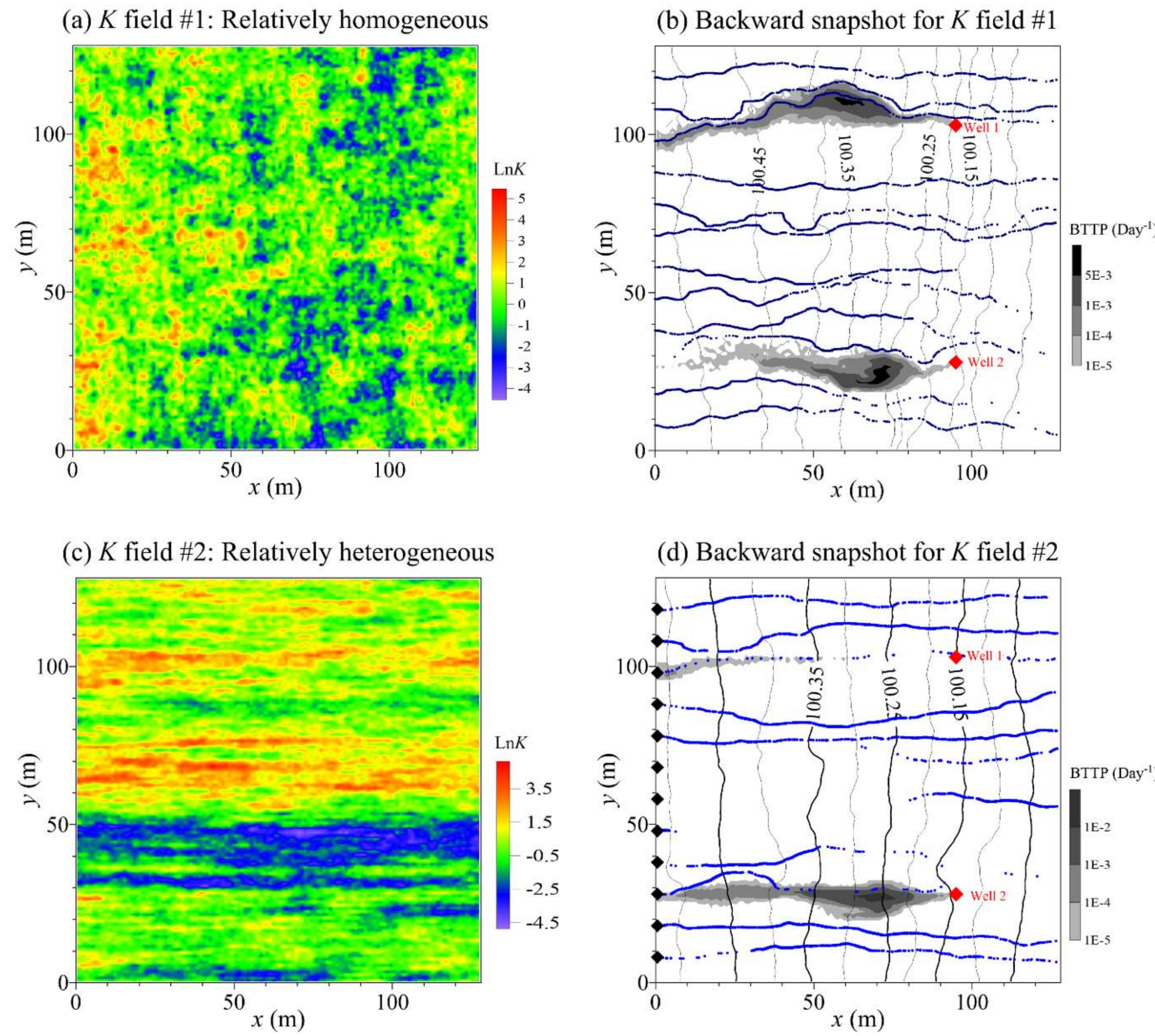


Figure 2. **Solver validation 2**: Two cases of operator-fractional Brownian fields (a) and (c). The corresponding backward particle tracking plume using the Lagrangian solvers for *K* field #1 and #2 is plotted in (b) and (d), respectively. In (b) and (d), black lines represent the hydraulic head calculated by MODFLOW, blue dotted lines denote the streamlines) starting from the left boundary (shown by the black diamonds in (d)), and the red diamonds show the location of two monitoring wells.

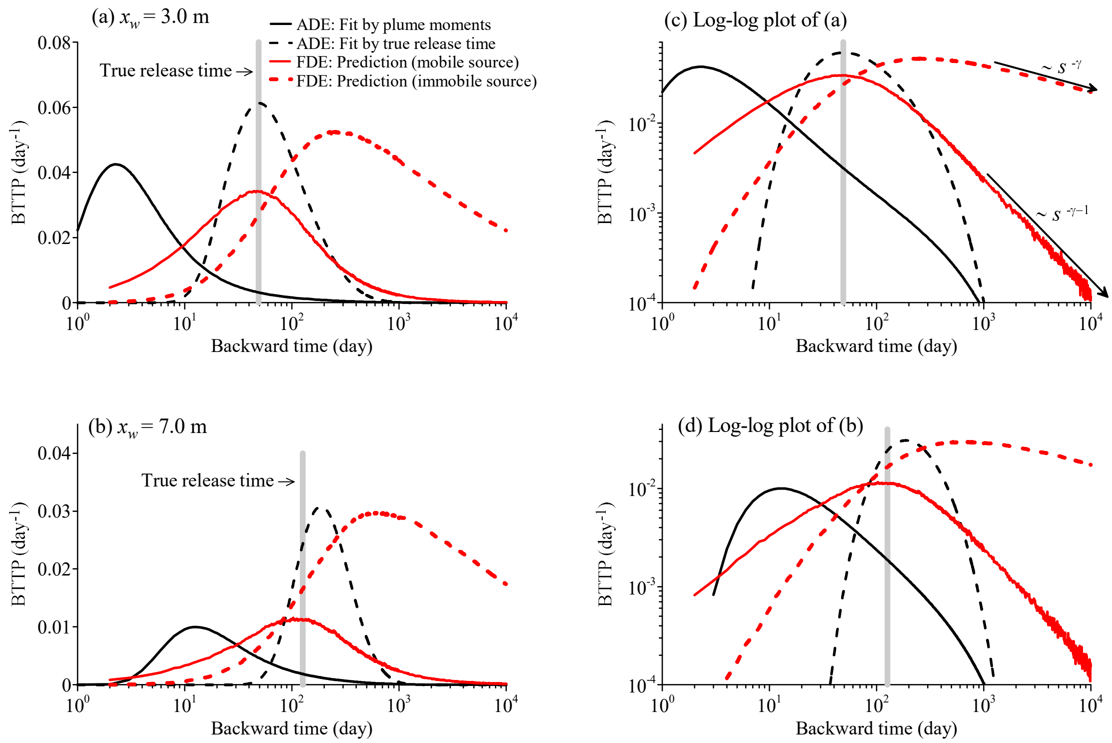


Figure 3. **BTTP Application 1:** MADE-1 aquifer: The calculated BTTP using the adjoint 1-d S-FDE (red lines) and the adjoint 1-d ADE (black line) for the observation well located at $x_w = 3.0$ m (a) and $x_w = 7.0$ m (b). (c) and (d) are the log-log plot of (a) and (b), respectively, to show the tailing behavior. The vertical grey bar denotes the true release time. The solid red line represents the BTTP for a mobile source, and the dashed red line represents the BTTP for an immobile source.

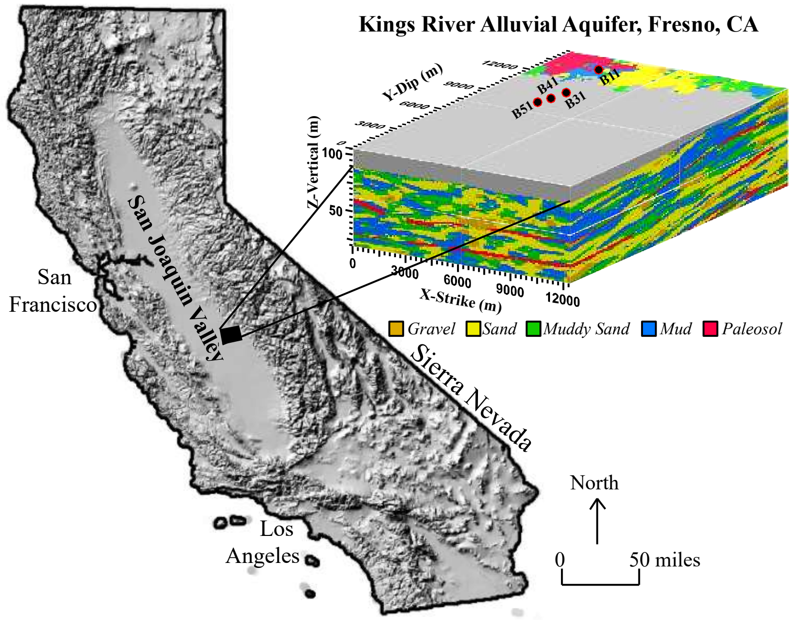


Figure 4. **BTTP Application 2:** KRAA - Location and the multiscale 3-*d* hydrofacies model for the Kings River alluvial aquifer, Fresno County, California.

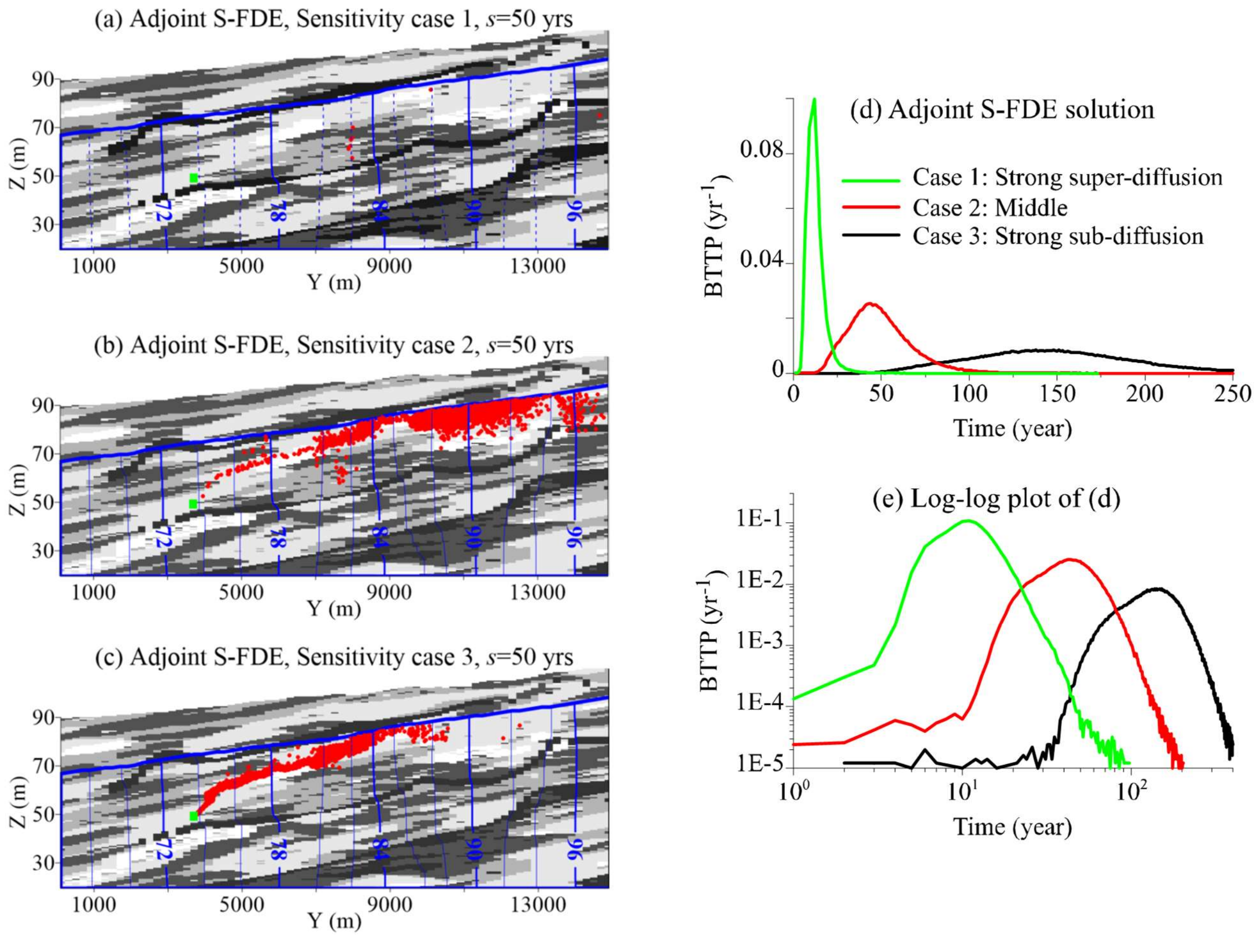


Figure 5. **BTTP Application 2:** KRFKings River alluvial aquifer (KRAA): A Snapshot snapshot (project of particle plumes) on within the vertical cross--section located at along strike the *X*-strike direction, with a coordinate of *X*=3,700 m shown in the hydrofacies model in **Figure 4**-. This snapshot was obtained throughof backward particle tracking at theover a backward time of *s*=50 yrs using the adjoint S-FDE (411a) for Case 1 (a), Case 2 (b), and Case 3 (c). The green rectangle in each plot represents the well screen (with a length of 0.5 m) where the groundwater sample is collected. In all cases, 5,000 particles were released initially at *s*=0. Plot (d) showsThe right plots show the corresponding BTTPs for these three cases, and plot (e) snapshot of backward particle tracking at time *s*=50 yrs using the adjoint ADE with the dispersivity $\alpha_L = \alpha_T$ =0.4 m (d), 0.04 m (e), and 0.004 m (f)is the log-lot version of (d). In all cases, 5,000 particles were released initially at *s*=0. The green rectangle in each plot represents the well screen (with a length of 0.5 m) where the groundwater sample is collected.

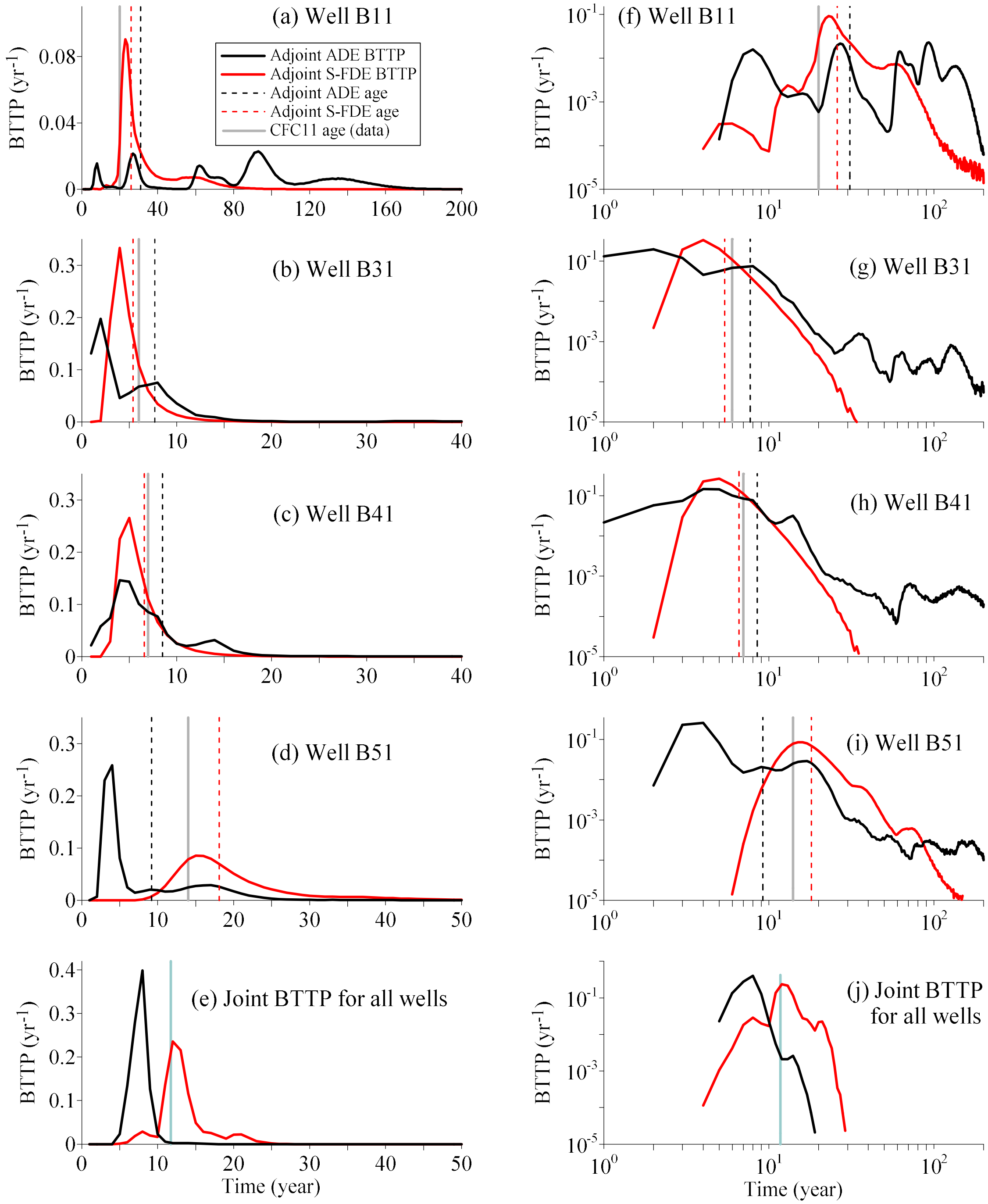
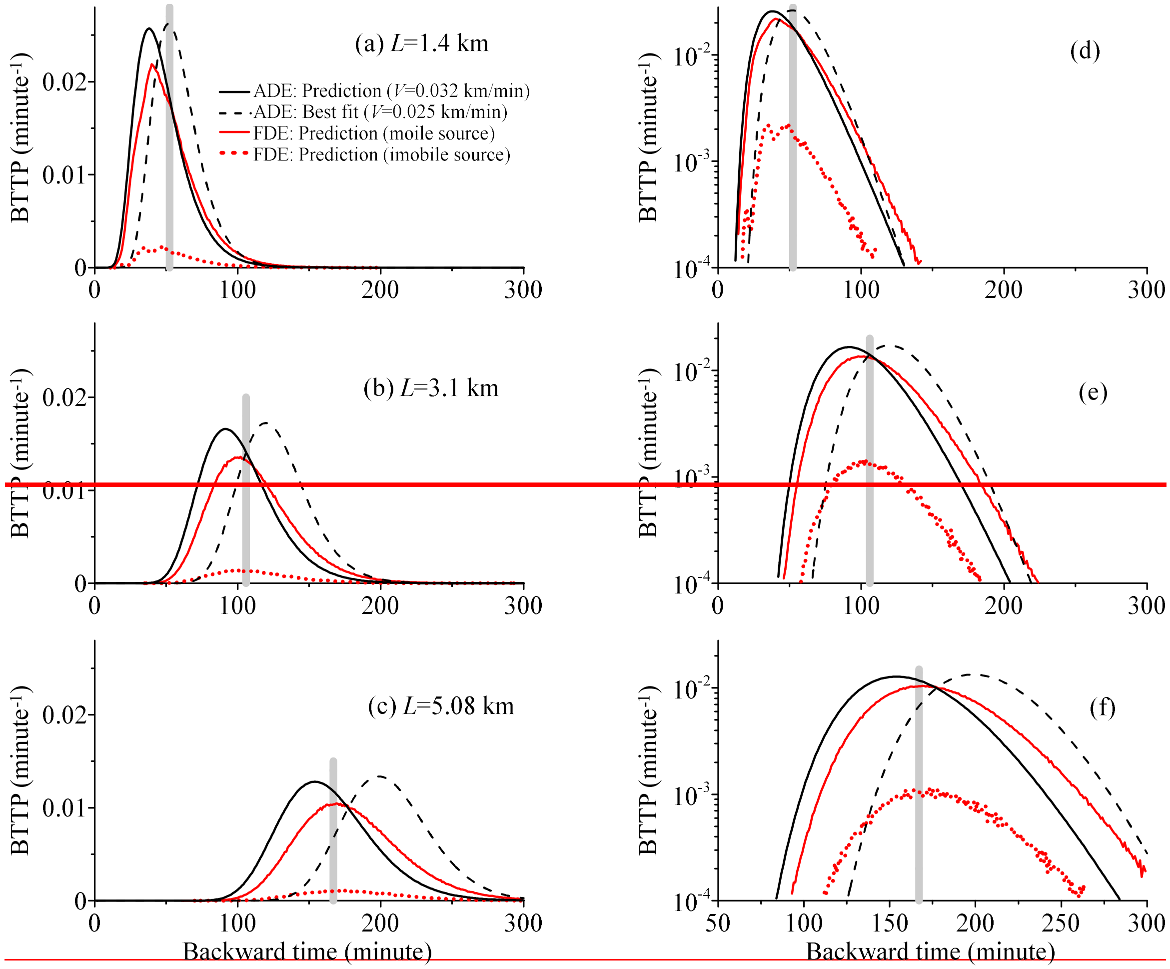


Figure 6. **BTTP Application 2:** KRF: the simulated BTTP using the adjoint S-FDE (red line) and the adjoint ADE (black line) for Well B11 (a), B31 (b), B41 (c), and 51 (d). The right plot is the log-log version of the left plot, to show the tailing. The vertical lines show the CFC-11 age measured in the lab (vertical grey line), estimated by the adjoint S-FDE (dashed red line), and estimated by the adjoint ADE (dashed black line).



(a) L=1.4 km
(b) L=3.1 km
(c) L=5.08 km
(d)
(e)
(f)
ADE: Prediction (V=0.032 km/min)
ADE: Best fit (V=0.025 km/min)
FDE: Prediction (moile source)
FDE: Prediction (imobile source)
BTTP (minute-1)
Backward time (minute)

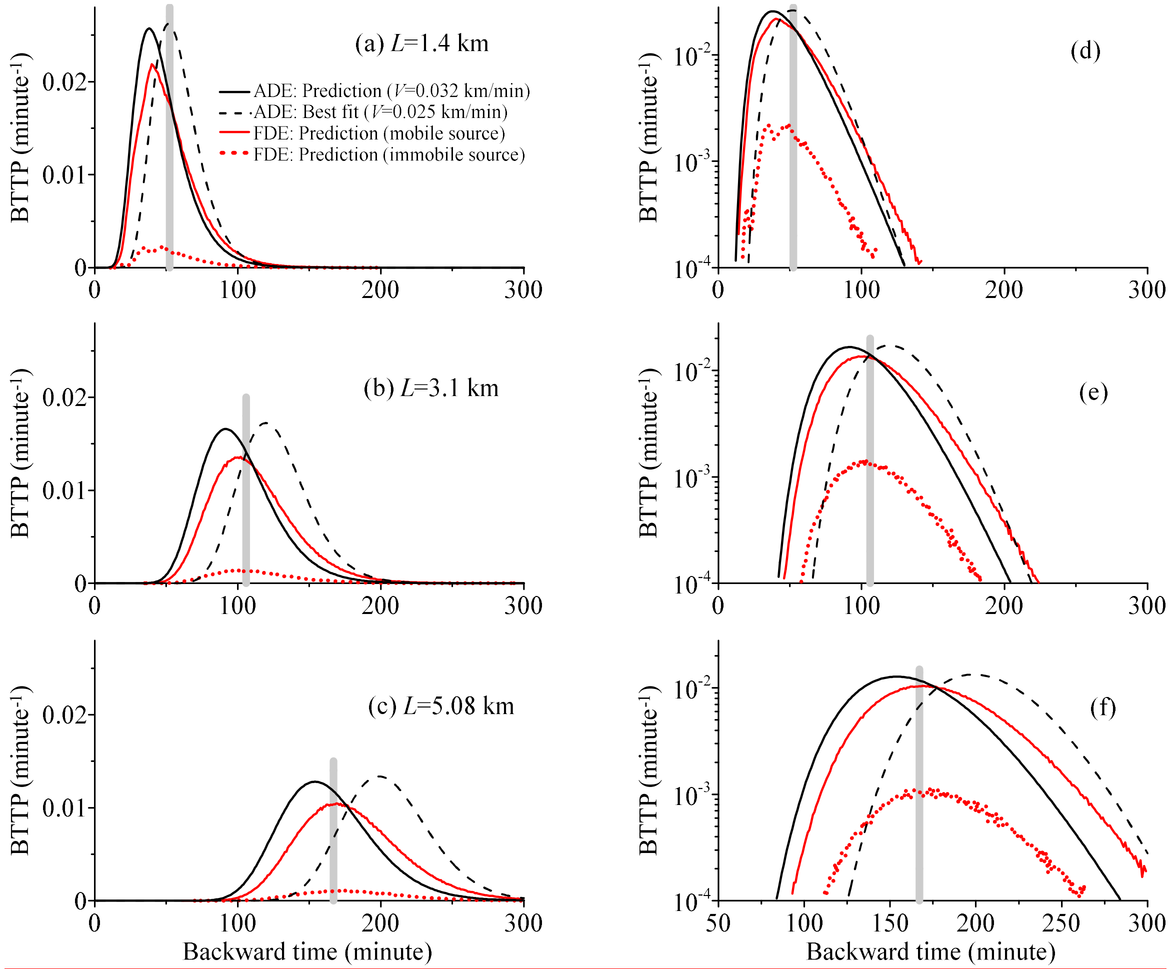


Figure 7. **BTTP Application 3** - Red Cedar River: the simulated BTTP using the adjoint S-FDE (red lines) and the adjoint ADE (black lines) for the backward travel distance of *L*=1.4 km (a), 3.1 km (b), and 5.08 km (c). The right plot is the semi-log version of the left plot, to show the tailing. The vertical bar in each plot shows the true release time. In the legend, "FDE: Prediction (mobile source)" represents the predicted BTTP using the adjoint S-FDE for a mobile source, and "FDE: Prediction (immobile source)" represents the predicted BTTP for an immobile source.

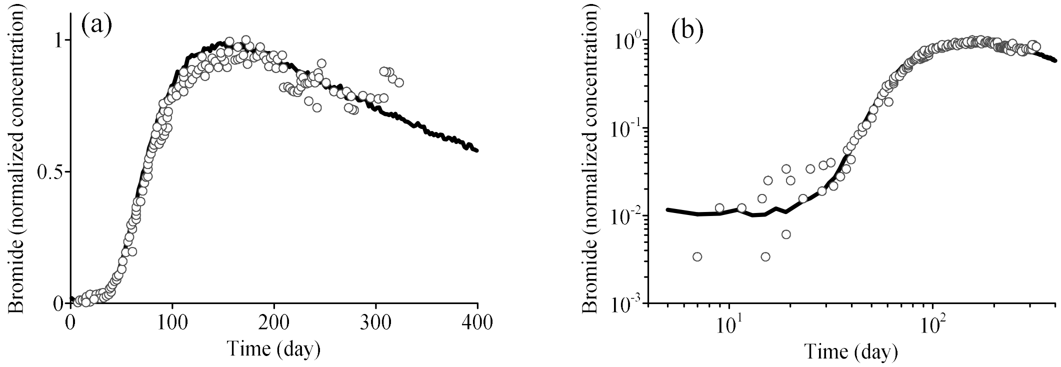


Figure 8. **BLP Application 1**: SHOAL test site: the measured (symbols) vs. the best-fit (line) bromide breakthrough curve using the vector model S-FDE (1a). (b) is the log-log plot of (a), to show the BTC tail.

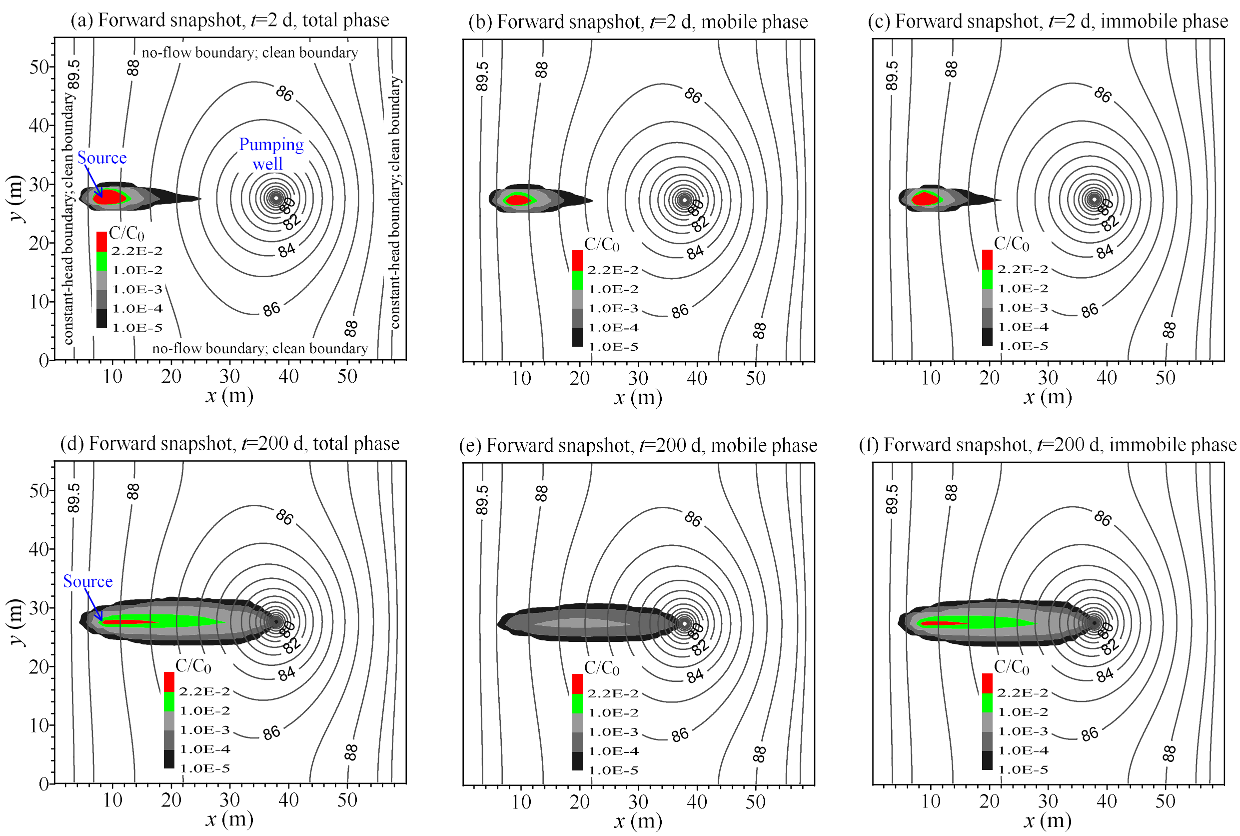


Figure 9. **BLP Application 1**: SHOAL test site: the modeled forward snapshot for the total phase (a), mobile phase (b), and immobile phase (c) at time t=2 days. (d), (e), and (f) show the snapshot at time t=200 days.

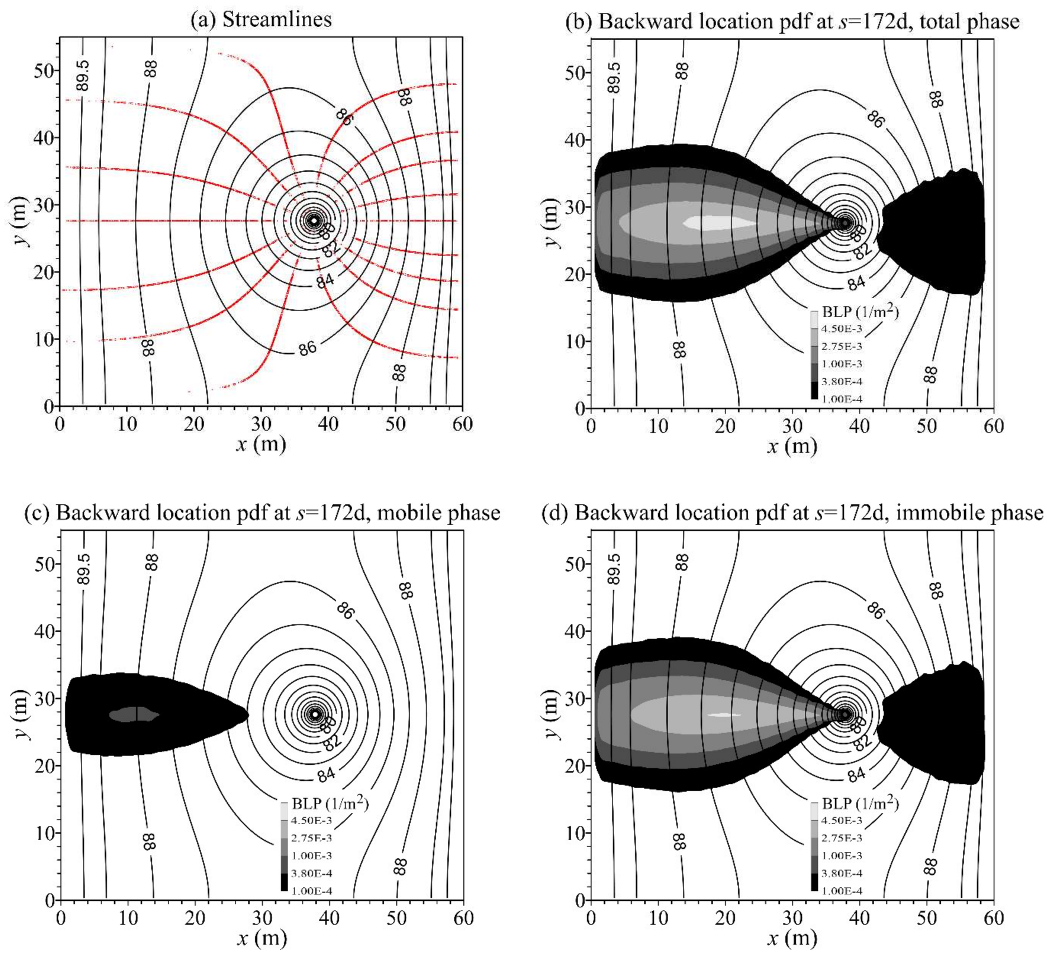


Figure 10. **BLP Application 1**: SHOAL test site: the modeled backward streamlines starting from the pumping well (a), and the calculated backward location probability density function (BLP) for pollutants located initially in the total phase (b), mobile phase (c), and immobile phase (d). It is noteworthy that there is a low concentration blob on the east side of the pumping well, due to the divergent flow in the backward model.

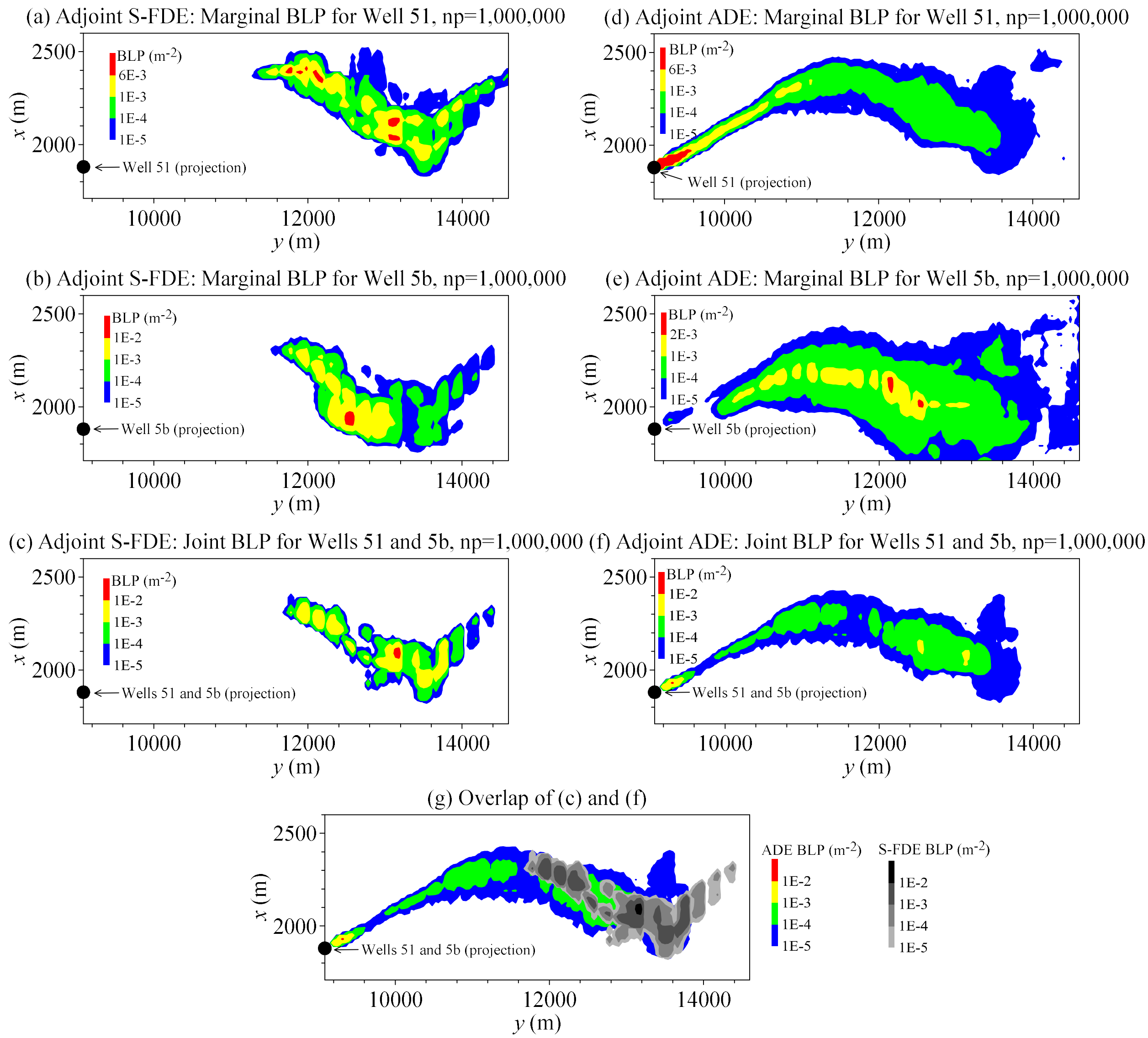


Figure 11. **BLP Application 2**: KRF: the simulated BLP using the adjoint S-FDE for Well B51 (a), B5b (b), and the adjoint BLP for Wells B51 and B5b (c). The adjoint ADE results are shown on the right plots. (g) is the overlap of plot (c) and (f). In the legend, “np” denotes the number of particles released in the Lagrangian solver.

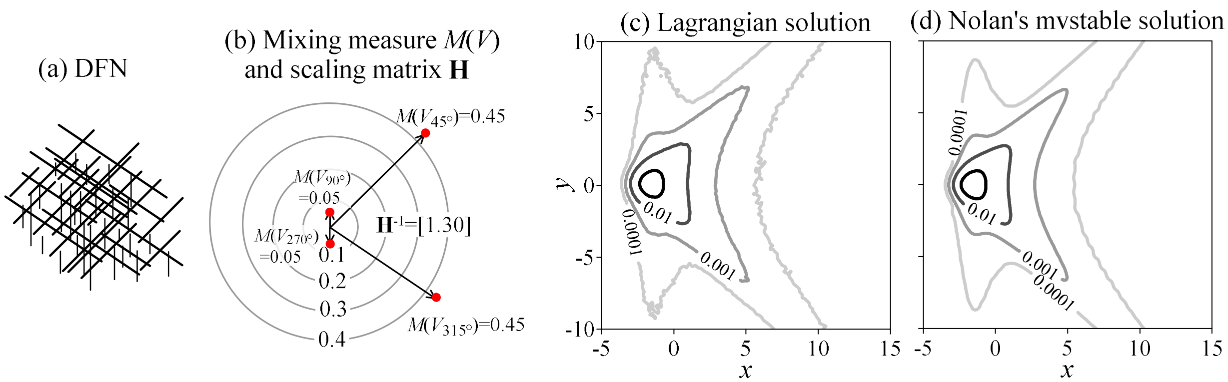


Figure 12. **Solver validation**: (a) shows the schematic diagram of a 2-*d* discrete fracture network. (b) is the polar plot of the discrete mixing measure and the scaling matrix. (c) is the Lagrangian solution of the multi-scaling S-FDE. (d) is Nolan's (1998) multivariate stable distribution.

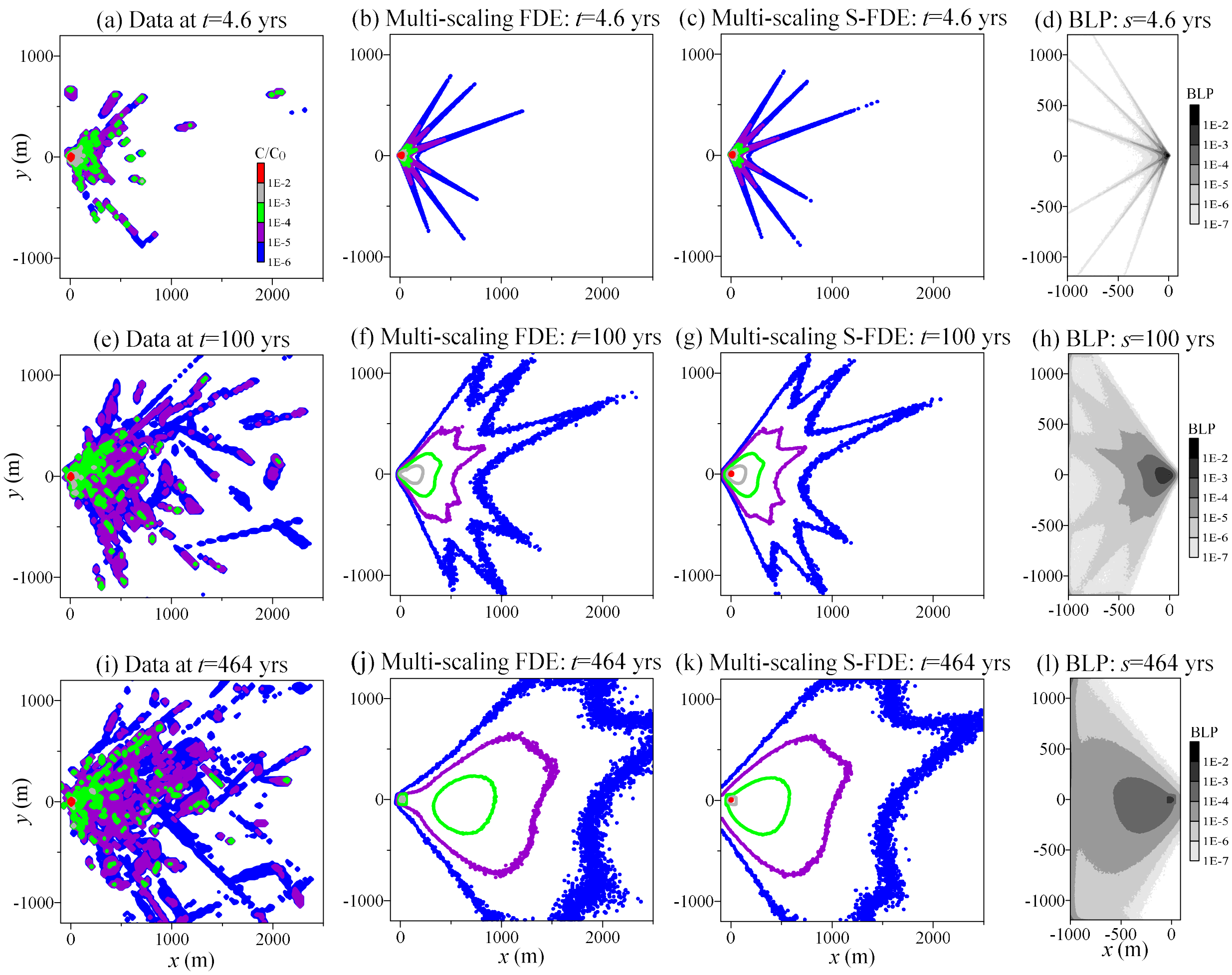


Figure 13. **Application of the multi-scaling S-FDE in DFNs**: (a) shows the average plume snapshot at time *t*=4.6 yrs from Monte Carlos simulations of pollutant transport in DFNs (Reeves et al., 2008). (b) and (c) are the best-fit solution using the multi-scaling FDE and multi-scaling S-FDE, respectively. (d) shows the resultant BLP using the multi-scaling S-FDE. The middle row (e)~(h) shows the result at a later time *t*=100 yrs, and the bottom row (i)~(l) shows the result at a later time *t*=464 yrs. Note that the model solutions in the middle and bottom rows are prediction results using parameters fitted in the top row.