Advances in Water Resources 54 (2013) 11–21

Contents lists available at SciVerse ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Using groundwater age distributions to estimate the effective parameters of Fickian and non-Fickian models of solute transport

Nicholas B. Engdahl [a,*], Timothy R. Ginn [b], Graham E. Fogg [a]

[a] Department of Land, Air and Water Resources, University of California, Davis, United States
[b] Department of Civil and Environmental Engineering, University of California, Davis, United States

ARTICLE INFO

Article history:
Received 25 September 2012
Received in revised form 14 December 2012
Accepted 19 December 2012
Available online 29 December 2012

Keywords:
Groundwater age
Solute transport
Parameter estimation
Nonstationary

ABSTRACT

Groundwater age distributions are used to estimate the parameters of Fickian, and non-Fickian, effective models of solute transport. Based on the similarities between the transport and age equations, we develop a deconvolution based approach that describes transport between two monitoring wells. We show that the proposed method gives exact estimates of the travel time distribution between two wells when the domain is stationary and that the method still provides useful information on transport when the domain is non-stationary. The method is demonstrated using idealized uniform and layered 2-D aquifers. Homogeneous transport is determined exactly and non-Fickian transport in a layered aquifer was also approximated very well, even though this example problem is shown to be scale-dependent. This work introduces a method that addresses a significant limitation of tracer tests and non-Fickian transport modeling which is the difficulty in determining the effective parameters of the transport model.

1. Introduction

One of the greatest limitations to accurately predicting solute migration in porous media is the difficulty of estimating the parameters required for a transport model. Non-uniqueness, uncertainty, and a lack of *a priori* data on the transport properties of the medium are just some of the factors that affect parameter estimation but selecting an appropriate model of transport is also a crucial step. It is now well known that natural heterogeneity occurring on multiple scales challenges the application of advection–dispersion equations (ADE) [1]. Mass transfer processes can cause power-law tailing in breakthrough curves [2,3] and preferential pathways allow for faster than expected arrivals [4], neither of which can be modeled with the ADE unless the relevant heterogeneity is included. Transport exhibiting these behaviors is commonly referred to as pre-asymptotic or non-Fickian (though non-Fickian transport can be asymptotic or pre-asymptotic) and this topic accounts for the bulk of the recent literature on contaminant transport. Several alternatives to the ADE have been developed for modeling transport in geologic media including nonlocal pre-asymptotic models [5], discrete or continuous multiple-rate mass transfer [2,6,7], continuous-time random walks (CTRW) [8], fractional differential equations [9], generalized memory function or convolution approaches [10–12], stochastic approaches [13,14], and space-time nonlocal tempered models that combine many of the characteristic elements of the other approaches [15,16]. In the absence of sufficient characterization data, non-Fickian models are more effective in simulating behavior observed in natural systems than the classical Fickian theory; however, the underlying equations can be difficult to solve and determining the parameters for these robust models of transport is also challenging [17,18]. Many of these methods also require numerous parameterizations or assumptions that cannot be validated without some degree of field testing or extensive modeling which limits their predictive capabilities. Neuman and Tartakovsky [1] went so far as to say that, without measured concentrations or mass fluxes, the parameters of non-Fickian models cannot be estimated. Given the complexity of natural porous media, and of the non-Fickian transport equations, it is common to use a 1-D effective model of the transport domain and the most common method for estimating the parameters of such a model is to conduct a tracer test at the site. This approach to parameter estimation is effective but it is fundamentally limited in predictive models because the tracer and solute may not sample the same portion of the aquifer. More importantly, the time or length scales of transport and the tracer test may also differ and nonlocal transport equations can only properly account for non-Fickian behavior when the tracer sufficiently samples the heterogeneity structures relevant for transport [see [19]].

Full scale tracer tests over the entire transport domain are not the only option and many alternatives exist, but these share the limitation that they do not always adequately sample the heterogeneity structure [20] and can mask some of the non-Fickian behavior [17]. For example, in a push–pull test the reversal of

* Corresponding author.
E-mail address: nbengdahl@ucdavis.edu (N.B. Engdahl).

0309-1708/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.advwatres.2012.12.008

the flow direction during the withdrawal phase can mask space nonlocal behavior and some of the information that may be needed for forward modeling is lost [e.g. [21]]. Motivated by similar reasons, Le Borgne and Gouze [18] showed that multiple hydraulic stresses of different types or scales provide a much needed improvement for parameter estimation but doing so is far from common practice. Sometimes the parameters of the transport model can be inferred by relating the physical structure of the aquifer and the transport behavior. For example, Kohlbecker et al. [22] showed that heavy tailed hydraulic conductivity fields may result in heavy tailed velocity distributions and non-Fickian transport and Dentz and Berkowitz [23] investigated connections between the heterogeneity structure and transport for spatially random adsorption; additional references connecting observed responses and the effects of physical/chemical heterogeneities can be found in Seeboonruang and Ginn [24]. Such approaches are promising and merit further investigation but they cannot currently be applied without some degree of calibration. Similarly, Willmann et al. [3] investigated the connection between hydraulic properties and breakthrough curves, finding that the connectivity of the aquifer is more important than the conductivity. This illustrates that, in many cases, the relevant processes or properties needed to derive effective parameters are not understood or resolved highly enough to be useful in predictive models. Less parameterized methods for modeling transport exist like the transfer-function approach [25,26] or the stochastic–convective framework [27]. Though useful for descriptive purposes and some predictive modeling, these methods do not involve characterization of spatially-variable hydraulic properties and thus do not afford robust simulation under transient (changes in flow direction) forcings. Based on the previous work, the most reliable way to determine transport model parameters is a tracer test over multiple temporal and spatial scales where the tracer is affected by the heterogeneity structure in the same way a contaminant would be. Although not a directly measurable physical property, groundwater age metrics may serve as a proxy for tracer data and allow one to infer the parameters of an effective transport model [28].

Groundwater age has already been recognized as a potential calibration tool for groundwater models [29–31]. The governing equations of groundwater age are very similar to the equations of solute transport in porous media [32]. Central to understanding age is the concept that any sample of groundwater contains a distribution of ages [33–38]. The age of a water molecule can be defined as the elapsed time since the molecule entered the aquifer and early work by Raats [25] considered age in transport problems in this fashion. Each water molecule at a given sampling location may have taken a different path to reach the sampling well and has a unique age [31]. The mass density of Ginn [32] can be thought of as the continuous property describing how much of the total mass of the system is concentrated at a specific age at a particular location and time; the mass density is a distributed quantity much like a solute concentration [32]. The importance of age as a distributed quantity was recognized by Campana [39] and other studies have since elaborated on the idea or provided derivations of the governing equations. The first rigorous derivation of the age equations was given by Ginn [32] who defined age as an additional dimension of the problem space in a continuum mechanical framework. This approach was more general than the previous derivations where age was defined as the current time minus some initial time and only the mean ages [e.g. [40]] or the age moments [e.g. [33]] were considered. The additional dimension for age allowed the distribution of mass density to continue changing in the age dimension even if the distribution was stable with respect to time. Furthermore, this formulation with an independent, orthogonal, age dimension allows for transience which is addressed in a recent paper by Cornaton [41] and the feature is clearly critical for transient forcings like seasonal rainfall and, on larger scales, climate change. Recently, Engdahl et al. [28] derived the age equation in a way that explicitly adds the possibility of nonlocal dispersion. The authors derived Fickian and non-Fickian forms of the age equation but also showed that an important relationship exists between the transition probability distributions in the time and age dimensions; in fact, they must be identical. Because of this equivalence, Engdahl et al. [28] speculated that age distributions might be able to function as a tracer to determine the transport properties of an aquifer but a method for doing so was not developed and no examples were given to support this claim.

In this article, we present a theoretical method for estimating the parameters of effective transport models. The spirit of the work is to follow a typical parameter estimation workflow toward the goal of developing accurate one-dimensional (1-D) effective models of transport. The context of this analysis may be most applicable to risk assessment where estimates of the velocity and extents that a plume will reach over a given amount of time are the primary concerns but our analysis is applicable to a wider range of transport problems. We do not attempt to determine the detailed heterogeneity structure of the aquifer but rather the expected response of a portion of the system to an instantaneous pulse. Our goal is to determine if the forward migration of a plume can be reasonably approximated using only basic well data (relative locations and piezomteric head) and groundwater age distributions for Fickian and non-Fickian models of transport. This question is evaluated using two different conceptual representations of a synthetic aquifer and a number of mathematical tools from the non-Fickian and stochastic–convective transport theories which are adapted to the age problem.

2. Groundwater age

The literature related to the subjects of non-Fickian transport and groundwater age is too vast to be summarized in a few paragraphs so we assume a basic, conceptual, familiarity with both. Here, only a high level overview that compliments the introduction is provided, covering the important points connecting transport and age. There is a noteworthy separation in the literature between articles dealing with the geochemical methods used to measure age and those presenting mechanistic applications of age in distributed hydrologic models [see [42]]. Few papers have developed robust connections that link the geochemical and mechanical perspectives on age beyond the so-called lumped parameter models. These models assume the shape of an age distribution and environmental tracer data is used to determine the best fit parameters of that model [e.g. [43–46]]. Factors like pumping [47], intra-borehole mixing [29], isotopic fractionation [48], mass transfer [49], and transience [41] can affect the inferred ages of the lumped parameter model but a reasonable representation of the actual age distribution in an aquifer is still possible when the effects of these factors on tracer concentrations are addressed. The tools available for measuring age have greatly improved over the decades and simple mixing models can also be used to combine data from different tracers [e.g. [50]]. Some connections between geochemical and mechanistic models have been made such as Varni and Carrera [33], who showed that the measured radiometric age of a sample is related to the true mean age through the moments of the age distribution, and Massoudieh and Ginn [51], who recently showed that the measured concentration of an ideal radiogenic tracer defines one point on the Laplace transformed age distribution. Though not the focus of this paper, such connections are needed to advance the applications of age distributions because the difficulty in determining age distributions is a significant limitation to the practical application of age data in natural systems. Consequently, most age distributions appearing in the literature have been the result of simulations [e.g. [28,33,35,37,41]]. However, recent research has generated results specifically on how the groundwater age distribution may be constrained [38,46], inferred [20], or formally estimated [52]. Note that [46] and [52] involve age distributions developed from multiple tracer data at field sites.

Groundwater age has been used extensively to determine transport velocities [45], often using a travel time based approach [e.g. [43]]. Age distributions also contain global information about the flow system as a whole and can be used to infer the aquifer structure, in addition to serving as a calibration tool [28,38]. Additionally, it may be possible to characterize flow rates, storage volumes, and mixing using age [53]. The fundamental reason that age can used be an efficient tool for these applications can best be seen from the governing equation of age. The equations that are used to describe age distributions closely resemble the equations of solute transport and the connections between the two provide the theoretical basis for our analysis.

2.1. Governing equations: Fickian dispersion

Engdahl et al. [28] derived the age equation from continuous random walks in five dimensions which is analogous to random walk derivations of the macroscopic transport equations except for the added convective dimension in age. Their general result can recover the Fickian (asymptotic) age equation [32] which is

$$\frac{\partial \rho(\mathbf{x},t,a)}{\partial t} + v_a \frac{\partial \rho(\mathbf{x},t,a)}{\partial a} = \nabla \cdot \mathbf{D_x} \nabla \rho(\mathbf{x},t,a) - \nabla \cdot \mathbf{v_x} \rho(\mathbf{x},t,a) \quad (1)$$

where ρ is the aqueous phase mass density [32], $\mathbf{x}$ is a 3-D position vector, t is time, a is the exposure time or age variable, v_a is the exposure time or "aging" velocity (equal to unity for age), $\mathbf{v_x}$ is the velocity vector in $\mathbf{x}$, and $\mathbf{D_x}$ is a dispersion tensor; note that the gradient operator only acts on the spatial dimensions. For simplicity, we will use the constant coefficient, 1-D (in space) form of (1) which is:

$$\frac{\partial \rho(x,t,a)}{\partial t} + \frac{\partial \rho(x,t,a)}{\partial a} + v_x \frac{\partial}{\partial x} \rho(x,t,a) - D_x \frac{\partial^2}{\partial x^2} \rho(x,t,a) = 0 \quad (2)$$

If the mass density has reached a steady-state with respect to time, which we will assume throughout this article, the time derivative in Eq. (2) is zero and the equation is no longer a function of time; this is the steady state behavior. Under these conditions, Eq. (2) simplifies to:

$$\frac{\partial \rho(x,a)}{\partial a} + v_x \frac{\partial}{\partial x} \rho(x,a) - D_x \frac{\partial^2}{\partial x^2} \rho(x,a) = 0 \quad (3)$$

where we have removed time as a variable. The solution to this equation given a Dirac-delta boundary condition in age is the inverse Gaussian with age replacing time [30,32]. However, Eq. (3) can only describe Fickian dispersion.

2.2. Governing equations: Non-Fickian dispersion

Non-Fickian transport has been widely studied for several decades now but only recently has the theory been applied to groundwater age. Non-Fickian forms of the age equation were derived by Engdahl et al. [28] and similar derivations for continuous-time random walks without age are abundant [see [8]]. The basic concept of the random walk derivation is that probability density functions (PDFs) are specified for displacements, or jumps, in each of the dimensions of the problem space, and the form of the PDFs determines the final behavior of the model. For example, the Fickian model of age (Eq. (2)) is derived from the special case where the spatial jump PDF has finite first and second moments, and the time and age PDFs have finite first moments; the higher moments are assumed to be zero. The example of a non-Fickian governing equation for groundwater age given by Engdahl et al. [28] was derived using a Gaussian spatial jump distribution and a power-law distribution for time-age jumps which produces the time-age fractional model:

$$\frac{\partial^\alpha \rho}{\partial t^\alpha} + \frac{\partial^\alpha \rho}{\partial a^\alpha} + v \frac{\partial \rho}{\partial x} - D \frac{\partial^2 \rho}{\partial x^2} = \frac{t^{-\alpha}}{\Gamma(1-\alpha)} + \frac{a^{-\alpha}}{\Gamma(1-\alpha)} \quad (4)$$

where α is the order of a fractional derivative in the interval $0 < \alpha \leqslant 1$; we have dropped the explicit notation of the variables of ρ for compactness. Without getting lost in the details of non-Fickian constitutive theory, simply recognize that Eq. (4) is a generalization of Eq. (1) to a special kind of time-age nonlocal behavior. The two terms on the right hand side of (4) are remnants of the initial conditions which persist due to the properties of fractional differential operators, but both terms will vanish when $\alpha = 1$, which is the Fickian case. Notice that if the mass density is stable with respect to time (*e.g.* Eq. (3)), then time is no longer a variable and will not appear in (4), but the distributions can still be non-Fickian in age; this is the most significant difference between the family of equations described by Eq. (4) and the equations presented in the previous works on age and CTRW-based transport.

The jump distributions in a random walk for time and age are identical. Engdahl et al. [28] showed this by decomposing the joint PDF of time and age into a conditional probability that recovers identical marginal jump distributions for time and age. This requirement could also be inferred from earlier work by Harvey and Gorelick [54] who showed that the moments of transport and age distributions were identical for steady state flow observed in response to a Dirac-delta pulse boundary condition (Green's function). The significance of this equivalence is that knowledge of the jump PDF for the age dimension also gives the jump PDF for the time dimension. Since age is affected by the same hydrodynamical processes as passive solute transport, the steady state age distribution captures the effects of heterogeneities within the aquifer that will affect transport [30]. This is the motivation for using age distributions in lieu of tracers.

The time-age fractional model of mass density (Eq. (4)) is a mass balance statement with a fractional flux model that corresponds to power law waiting time distributions, or a higher than Fickian probability of long rests, that often manifest as heavy tailed breakthrough curves [e.g. [19]]. Eq. (4) is just one asymptotic form of a family of equations describing sub-diffusion which are characterized by the distribution of jumps in time and age. Many of the other models cannot easily be expressed as a real space differential equation and are solved in Fourier–Laplace space instead. We present (4) because it can easily be compared to Eqs. (1) and (2) but recognize that it is not universal for porous media because different physical mechanisms require different PDFs to correctly model transport or age. Non-Fickian transport in natural systems can also be simulated using the truncated power law model (TPL) [55]. Substituting age for time in Eq. (16) of Dentz et al. [55], the transition probability density is described by:

$$\lambda(a|a_1,a_2,\beta) = \{a_1 \tau^{-\beta} \exp(\tau^{-1}) \, \Gamma(-\beta,\tau^{-1})\}^{-1} \frac{\exp(-a/a_2)}{(1+a/a_1)^{1+\beta}} \quad (5)$$

where Γ is the incomplete Gamma function, $\tau \equiv a_2/a_1$, a_1 and a_2 are the upper and lower cutoff ages, respectively, β is the characteristic exponent that classifies the non-Fickian transport, and $\lambda(a)$ denotes the marginal probability distribution in age, given the required parameters. Recall that time and age transition probability densities can be used interchangeably in the age equations. This model of probability densities allows transport to transition to Fickian behavior after a characteristic or upper threshold length scale (a_2 in Eq.

(5)) is passed and more details about the significance of the parameters are given by Berkowitz et al. [8]. This model can also approximate a power-law model by making the upper cutoff age (a_2) large with respect to the scale of the problem. Although it was not explicitly derived by Engdahl et al. [28] for groundwater age, the authors presented examples where the TPL model was able to provide a better description of simulated age distributions than the power-law model. For steady flow, as shown in Ginn et al. [30], it is straight forward to recognize that the nonlocal model of age using a TPL distribution of age transition densities will be of identical form as the transport case (with the mass density substituting for concentration). This allows one to use the existing CTRW tools to describe steady-state age distributions that are non-local in age without modification. Although we will focus on the continuous random walk based methods, all of the non-Fickian approaches listed in Section 1 should be applicable to the age problem if their various assumptions are satisfied.

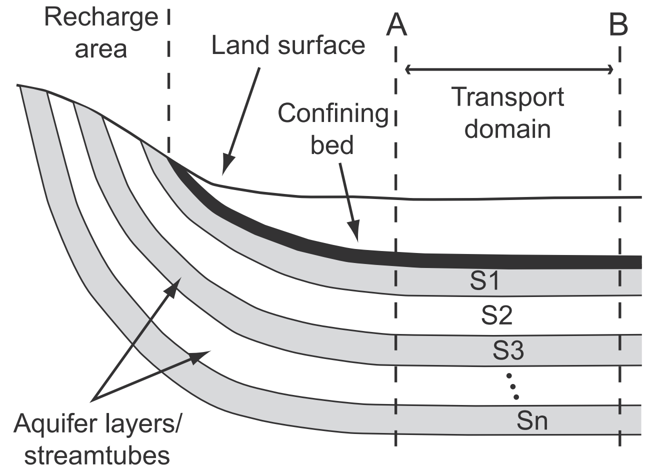

Fig. 2. Conceptual model of a layered aquifer used for the deconvolution approach.

3. Parameter estimation

3.1. Model setup

To illustrate our method, we will consider a confined homogeneous domain and a confined layered aquifer, both of which will use a monitoring network of three wells (Fig. 1). We will assume that age distributions are known in each well and that the wells fully penetrate the aquifer (Fig. 2); note that the homogeneous aquifer is a single layered formation. The sole source of recharge in both cases will represent mountain front recharge (Fig. 2) which is represented mathematically as a Dirac-delta (in age) function boundary condition that is perpendicular to the velocity vector. Flow in this system will be 2-D in the x–y plane and we assume that the velocity within individual layers is constant and that flow is horizontal. Groundwater age, and later transport, will be considered parallel to the direction of the velocity so that the effective model of this system will be strictly1-D. Furthermore, we will assume constant porosity and that the recharge boundary is approximately perpendicular to the flow direction. All of these assumptions can be relaxed, but doing so here would introduce complexity that is excessive for illustrating our proposed method.

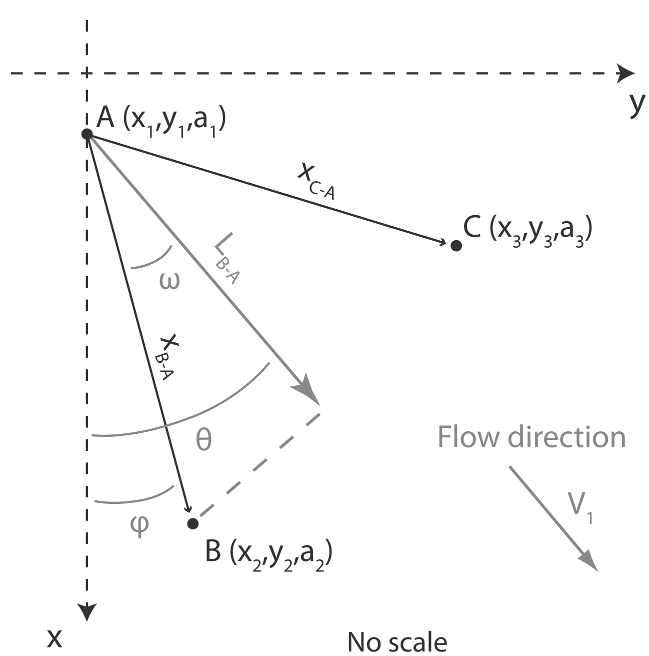

Fig. 1. Layout of the conceptual model of the network of monitoring wells.

3.2. Deconvolution approach

Given two monitoring wells on approximately the same streamline with observed age distributions, it is possible to determine the travel time distribution between them which can then be used to approximate transport [27]. This can be done in any aquifer but the utility of a streamtube based approach is most apparent in heterogeneous systems, particularly if an aquifer is laterally homogeneous but vertically heterogeneous which is a common approximation in many layered aquifers. The homogeneous model is not treated any differently since it is simply a layered aquifer with only one layer.

Consider the basic model of a layered aquifer that is shown in Fig. 2. Within each layer, flow and transport are essentially 1-D and each layer can be conceptually approximated as an independent streamtube that may have a unique velocity, $\nu(s,t)$, length, and dispersivity. If the layers do not interact with each other, this corresponds to the discrete streamtube ensemble of Simmons et al. [56]. For an infinite number of infinitesimal streamtubes, we recover an integral form and the expected breakthrough curve under steady flow for an instantaneous unit mass injection boundary condition is the travel time distribution:

$$C_x(t) = c_0 \int_0^{\infty} \delta(t-\tau) p_x(\tau) \mathrm{d}\tau = p_x(t) \tag{6}$$

where $C_x(t)$ is the expected concentration at location x, c_0 is the initial concentration, and p_x is the travel time distribution from the origin to location x [27]. Travel time distributions along streamlines were also considered earlier by Raats [25], who mathematically defined isochrones (lines or surfaces of constant age) from the flow equations. The transport applications in that work focused on computing solute fluxes in streamtubes, but the basic idea is the same. Conceptually or mathematically, it is straight forward to see that if transport of a unit mass occurs from the recharge zone to the monitoring location under steady flow the travel time and age distributions are identical and we can exchange the two as $p_x(t) \to \rho_x(a)$ as noted by Varni and Carrera [33] and others. We will approximate the behavior of the ensemble shown in Fig. 2 as a single 1-D effective streamtube. The solute concentration, or mass density, in the ensemble is constructed as the flux-weighted average of the individual fluxes from the set of streamtubes (Fig. 2, $s_1, s_2, \ldots, s_n$) contained therein. The weight of each streamtube is the relative fraction of the total flow into the monitoring well contributed by that streamtube. The fully penetrating monitoring wells A and B (Fig. 2) both transect the same ensemble of streamtubes, and we define B as the downstream well. The age distributions at A and B are $\rho_A(a)$ and $\rho_B(a)$, respectively, and the breakthrough curve from the recharge zone to either well can be found by substituting these distributions into Eq. (6); however, for this problem we require $\rho_{B-A}(a)$, the difference in the age distributions between the wells which is also the travel time distribution. For the Dirac-delta boundary condition at the recharge location we observe $C_x(a) = \rho_x(a)$ and we can restate the problem as follows:

$$C_A(t) = \int_0^t \delta(\tau) p_A(t-\tau)\mathrm{d}\tau = \rho_A(t) \tag{7}$$

$$C_B(t) = \rho_B(a) = \int_0^a \rho_A(\tau)\rho_{B-A}(a-\tau)\mathrm{d}\tau \tag{8}$$

where $C_A(t)$ and $C_B(t)$ are the distributions of concentration over time. To reiterate, if we set injection time to zero, these are also the distributions of water over age, $\rho_A(a)$ and $\rho_B(a)$, respectively. The known age distribution at A (given from Eq. (7)) is used as a boundary condition in (8) for the transport problem from A to B and the known distribution at B is the solution. In other words, Eq. (8) is simply the composition of two sequential linear transfer functions to calculate the composite age distribution after transport from the source to A then to B. To find the unknown travel time distribution, ρ_{B-A}, we find the deconvolution of (8) which can be done with a Laplace transform. The Laplace transform (LT) is defined to be

$$L\{f(a)\} = \int_0^\infty \exp(-sa)f(a)\mathrm{d}a$$

and the LT of Eq. (8) with respect to age is

$$L\{\rho_B(a)\} = L\left\{\int_0^a \rho_A(\tau)\rho_{B-A}(a-\tau)\mathrm{d}\tau\right\} \Rightarrow \tilde{\rho}_B(s) = \tilde{\rho}_A(s)\,\tilde{\rho}_{B-A}(s) \tag{9}$$

where L denotes the forward LT operator, s is the Laplace parameter, $\sim$ denotes a LT function, and the convolution theorem allows us to write the final form of (9) as the product of the LT functions. The underlying travel time distribution between the wells can then be determined as

$$\rho_{B-A}(a) = L^{-1}\left\{\frac{\tilde{\rho}_B(s)}{\tilde{\rho}_A(s)}\right\} = L^{-1}\{\tilde{R}(s)\} \tag{10}$$

where L^{-1} denotes the inverse LT, and $\tilde{R}$ is shorthand for the ratio of the LT of the known age distributions. When the effective parameters and physical processes governing transport between the two wells do not vary in time, the solution of (10) should reduce to a model that has the same parameters representing transport between the wells and the length scale should be the distance between the wells. Such a transport model assumes spatial invariance or at least incremental stationarity.

To demonstrate the invariance of (10), consider the LT of the general solution of the time-steady age equations for a Dirac-delta boundary condition with an arbitrary memory function, $M(t)$, that describes the distribution of transition times in the system [55]. This result is a generalization of the LT of the classical, Fickian solution (the inverse Gaussian) to allow for a broad distribution of resting or trapping times. Engdahl et al. [28] showed that age distributions can substitute for transition time distributions, so we will express the memory function in age as $M(a)$. The LT of the general solution is:

$$\tilde{\rho}_i(x_i, s) = \exp\left[-\frac{x_i v}{2D_L}\left(\sqrt{1 + 4\frac{sD_L}{\tilde{M}(s)v^2}} - 1\right)\right] \tag{11}$$

where x_i is the 1-D position from the source to the ith well, D_L is the longitudinal dispersion coefficient, $\tilde{M}$ is the LT of the memory function, s is the Laplace variable dual to age, and we have exchanged concentration for the mass density [modified after 55]. Eq. (11) can represent Fickian behavior when $\tilde{M}(s) = 1$ and represents non-Fickian behavior when the memory function takes on a different form. If the age distributions at A and B and the space between the wells are assumed to be described by a single memory function, the effective representation of the domain is stationary and should only be a function of separation distance. This can be demonstrated by inserting (11) into (10), using $\tilde{\rho}_A(x_A, s)$ and $\tilde{\rho}_B(x_B, s)$, and simplifying using the properties of exponentials which gives:

$$\tilde{R}(s) = \exp\left[\frac{v}{2D_L}(x_B - x_A)\right]\exp\left[\left(\sqrt{1 + 4\frac{sD_L}{\tilde{M}(s)v^2}}\right)\frac{v}{2D_L}(x_A - x_B)\right] \tag{12}$$

Define $\chi = (x_B - x_A)$ and (12) can be rewritten as

$$\tilde{R}(s) = \exp\left[-\chi\frac{v}{2D_L}\left(\sqrt{1 + 4\frac{sD_L}{\tilde{M}(s)v^2}} - 1\right)\right] \tag{13}$$

which is the LT general solution, Eq. (11), evaluated at the separation distance between the wells, χ, and the inverse LT of (13) is the solution of (10). Given two wells and their age distributions we can determine the travel time distribution in the aquifer between them and express that distribution as a function of arbitrary separation distance as long as the memory function can be presumed stationary. The travel time distribution can then be combined with a suitable boundary condition in Eq. (6) to determine the expected breakthrough curve.

It is particularly important to recognize that the heterogeneity structure of a system may not be suited to description with a stationary memory function in many practical situations. This implies that the effective model parameters change between the wells and the assumption of stationarity is violated; however, a representation of the travel time can still be found from the deconvolution of the age distributions from Eq. (10). This kind of approximation is strictly a transfer function approach [e.g. [25,26]] and should not be expected to perform well at scales other than the separation distance between the wells because of non-stationarities. Some other limitations of this approach are the assumption that the wells sample the same aquifer over the same ensemble of streamtubes, and that the wells are approximately on the same streamline. If these conditions are satisfied then (10) is an exact description of the travel time between the wells, but it is an approximation for the separation distance between the wells under all other circumstances.

4. Examples

In this section we will describe our two example models and investigate the ability of the deconvolution method to recover the correct age distributions; this is presented separately for the homogeneous (Section 4.1) and layered (Section 4.2) aquifers. The practical utility of the deconvolution method for predicting solute migration is then examined in Section 4.3. These two examples were selected because they represent the end-members of the range of behaviors for our synthetic domain and this is discussed in Section 5.

4.1. Uniform aquifer

The conceptual model for this example is a homogeneous domain, or that the number of layers (streamtubes in the ensemble) is one. We will assume that flow is horizontal between the wells and that the recharge boundary (located beyond the monitoring domain) is perpendicular to the velocity. The model domain in this example has a prescribed hydraulic gradient of 0.05, a hydraulic conductivity of 10 m/d, and kinematic porosity of 0.27; these values result in a seepage velocity of about 1.88m/d. We also assume that the velocity vector is oriented 17.7° from the *x*-axis (θ = 17.7 in Fig. 1) so flow is 2-D relative to the coordinates of the domain, but only 1-D along the flowpath. The locations of the monitoring wells are given in Table 1 where Well A is used as the origin of the coordinate system; Well A is located 5000 m downstream from the recharge boundary. We assume 1-D transport aligned with the velocity vector and a constant dispersion coefficient of 500m^2/d. At each of the monitoring locations, the complete age distribution was generated using the analytical solution of Eq. (3) for a Dirac-delta boundary condition in age (the inverse Gaussian) and the mean ages of these are given in Table 1. The synthetic "data" that will be used in both of our examples will only consist of the position of the wells relative to well A, and the age distribution in each well. Generally, we will only consider transport between wells A and B, but the age distribution in C was also simulated for further comparison of the models (Fig. 3).

Looking at Fig. 1, it should be evident that the important distance in this problem is not the absolute distance between the wells but rather the difference in downstream distance parallel to the velocity (L_{B-A} in Fig. 1). The downstream distance between the wells is the projection of the straight-line distance between the wells onto a line parallel to the velocity. This distance is L_{B-A} in Fig. 1 where θ is the angle of the velocity relative to the coordinate system of the model, ϕ is the angle from the *x*-axis to the line connecting wells A and B, and ω is defined as $\omega = \theta - \phi$ which is the angle needed to correctly project the distance between the wells. To find the downstream distance between the wells, the angle of the velocity vector must be known. The first moments of the age distributions (Table 1) and the locations of the monitoring wells can be used to approximate the direction of flow from finite differences and the angles are then easily found. Using finite differences of the mean ages and position for all three wells, the correct angle was recovered, and scaling the distance from well A to B by $\cos(\omega)$ recovers the correct downstream distance between the wells which is 2500 m.

The ratio of the Laplace transforms of the age distributions (Eq. (10)) was numerically approximated using a modification of the routines included with the CTRW toolbox [57]. The travel time distribution from well A to B (ρ_{B-A}) was found by numerically inverting the ratio of the LT functions (Eq. (10)). An effective velocity was determined based on the known length between the wells and the first moment of the deconvolution distribution. The dispersion parameter was determined by minimization of the errors of a trial solution against the deconvolution to find the best fit parameters of the deconvolution (Fig. 4); the dashed line is ρ_{B-A} from Eq. (10). There was a small amount of departure from the analytical solution near the peak of the distribution but this is most likely a result of the numerical approximation of the forward and inverse Laplace transforms. Despite the minor discrepancy in the peak, the parameters of the fitted model agreed very well with the specified model parameters and the age distributions in all three wells were recovered almost exactly using the effective parameters.

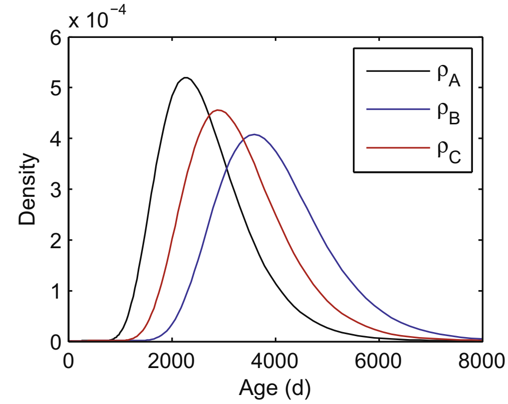

Fig. 3. Age distributions for the homogeneous example problem. The length scales from the recharge location to the monitoring wells increases from left to right.

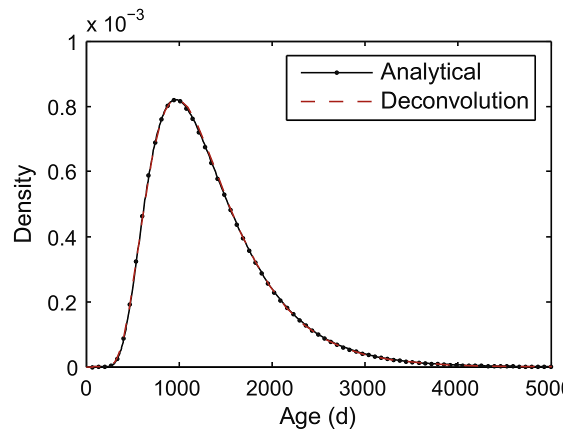

Fig. 4. Comparison of the fit of the deconvolution distribution (ρ_{B-A}) to the simulated age distribution for the homogeneous example.

Since this aquifer is truly homogeneous, the same parameters could have been inferred from a single age distribution if the distance of the well from the recharge location is known. By definition, a homogeneous system is invariant so the effective parameters from a single age distribution provide an exact description of the entire system. This example is very simple in order to provide a straight forward demonstration of the method but seldom, if ever, would we expect to see noise-free, inverse Gaussian age distributions in natural systems. This section is included as a proof of concept and we are not implying that the inverse Gaussian is a robust model of age; however, this section shows that the proposed method is functional.

Table 1
Well locations, simulated mean ages and the associated model parameters from the homogeneous example problem. *V* and *D* are the Fickian velocity and dispersion coefficients, respectively. The values for Wells A, B, and C are the best fit descriptions of the age distributions. For the wells, the downstream distance is from the recharge source to the well and the distance for the deconvolution represents transport between wells A and B.

	X-position (m)	*Y*-position (m)	Downstream distance (m)	Estimated length (m)	Mean age (d)	*V* (m/d)	*D* (m^2/d)
Well A	0	0	5000	5001	2660	1.88	500.0
Well B	2385	750	7500	7501	3990	1.88	500.0
Well C	414	2647	6200	6200	3298	1.88	500.0
Deconvolution	N/A	N/A	2500	2499	1329	1.88	500.1
Transport simulation	N/A	N/A	2500	2500	1329	1.88	500.0

4.2. Layered aquifer

The layered aquifer introduced in Section 3.2 (Fig. 2) will now be superimposed on the conceptual model of Fig. 1 to introduce a simple heterogeneity structure, but the location of the monitoring wells are unchanged. The age distributions in each well will now be represented with an ensemble of 100 independent streamtubes; recall that each layer is laterally infinite and homogeneous. The age distribution in each streamtube is described by an inverse Gaussian distribution (the solution to equation 3) but each streamtube is assigned a different length and velocity. A Gaussian distribution of streamtube lengths and a Gamma distribution of velocities were used to ensure that the simulated distributions exhibited non-Fickian behavior (Fig. 5). The velocity and length for each of the 100 streamtubes were selected by independently sampling each distribution and the values show no significant correlation (Fig. 5c). The mean values of these distributions were similar to the values used in the homogeneous example but the calculated means of the streamtube ensemble are different because 100 random samples of the length and velocity distributions are not enough to consistently reach the central limit behavior. The arithmetic mean velocity of the ensemble was $1.81 m/d$ and the mean streamtube length to well A was 4994 m. Note that the lengths shown in Fig. 5 correspond to the distance from the recharge area to well A; wells B and C use the same distributions of lengths and velocities but the length of each streamtube was increased by the downstream distance from A to the other wells which are 2500 and 1200 m, respectively. The dispersion coefficient in each streamtube was kept constant at $500 m^2/d$ for simplicity. The simulated age distributions of the streamtube ensemble are shown in Fig. 6 for the three monitoring locations. The distributions exhibit faster than Fickian arrivals and power-law tailing; both of these features are characteristic of non-Fickian behavior. However, it is crucial to recognize that non-Fickian transport characteristics can be caused by a variety of subsurface features (see Section 1) and the method used here is simply one way of creating the desired behaviors.

The angle to the velocity field was found using the locations and mean ages of the three wells, as in Section 4.1, and the estimated angle was identical to the estimate in the previous section. The downstream distance between wells A and B was found using this angle which recovered the correct spacing of 2500 m (L_{B-A}, Fig. 1). In the homogeneous example, the age distributions were Fickian and could be described exactly by analytical solutions but non-Fickian distributions in natural systems rarely allow for an exact analytical description. Non-Fickian analytical models were fitted to each simulated age distribution (i.e. a parametric approximation) and used in this analysis to simplify the forward and inverse Laplace transforms. Doing so serves a dual purpose because, under the assumption that the domain is stationary, a set of parameters that simultaneously describes the age distributions in both wells should exist (Eqs. (11), (12), 13) and failure to find such parameters would suggest a lack of stationarity. Non-stationarity will be apparent if the fundamental physical parameters of each well's fit differ. The CTRW toolbox was used to find a joint estimate (i.e. joint optimization) of the best fit parameters for the age distributions in wells A and B (Fig. 7). Using the TPL model for the transition probability densities, the best fit parameters were $2.22 m/d$, $640 m^2/d$ for the velocity and dispersion coefficients, respectively (Table 2). These parameters seem to provide a reasonable approximation of the system (Fig. 7), but we were unable to find a set of parameters that gave a more precise description of the distributions at all ages. The same joint estimation procedure was used for wells A and C but, similar to wells A and B, the fitted model generally split the difference between the two wells and did not provide a very accurate description; this is an early indication of non-stationarity.

We begin by assuming that the system is stationary and that both wells are described by the same model parameters. If these assumptions are met, the modeling is greatly simplified because Eq. (13) is satisfied. The travel time distribution between wells A and B (ρ_{B-A}) was found from the deconvolution of the jointly fitted model (parametric approximation) of the age distributions in wells A and B. The parameters describing the inferred travel time distribution (from the deconvolution) were found using the parameter estimation routine of the CTRW toolbox where the solution was evaluated at the downstream separation distance between the wells (Table 2). The best fit parameters for this stationary deconvolution closely matched those used for the parametric approximations of the age distributions, and the velocity and dispersion coefficients were similar to those used to describe the age distribution in well A (Table 2). However, the resulting models did not provide good descriptions of the age distributions in the remaining wells except for the memory function. The poor fit relative to the known age distributions suggests that the parameters are incorrect or that there is scale dependency.

Eq. (10) does not require that the parameters of each distribution be constant and the assumption of stationarity was only introduced to derive Eq. (13). Different parameters can be used to describe each age distribution to find the travel time distribution

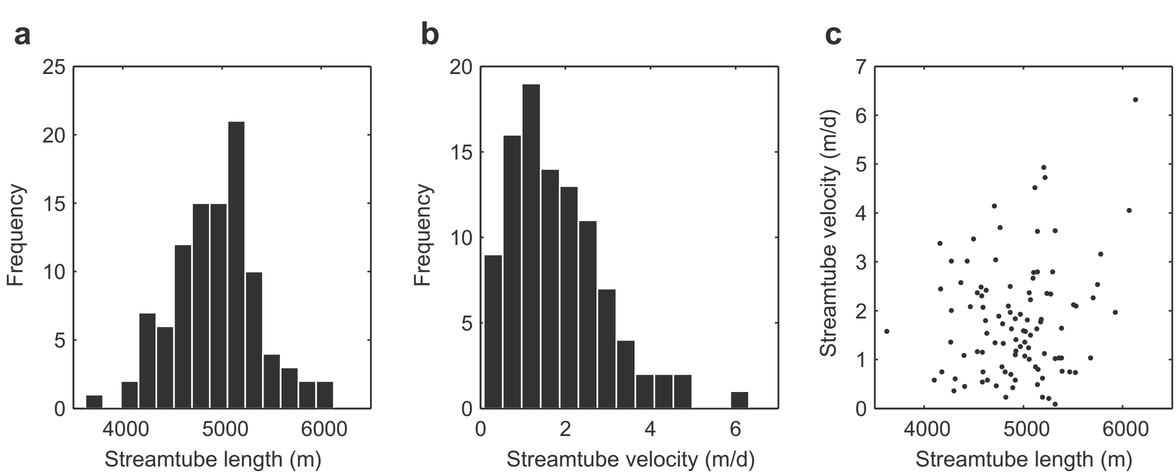

Fig. 5. The distribution of values used for the streamtube based, layered aquifer example for (a) streamtube length to well A, (b) streamtube velocity, and (c) the relationship between the parameters. The length distributions for wells B and C are linear shifts of the distribution for well A of 2500 m and 1200 m, respectively.

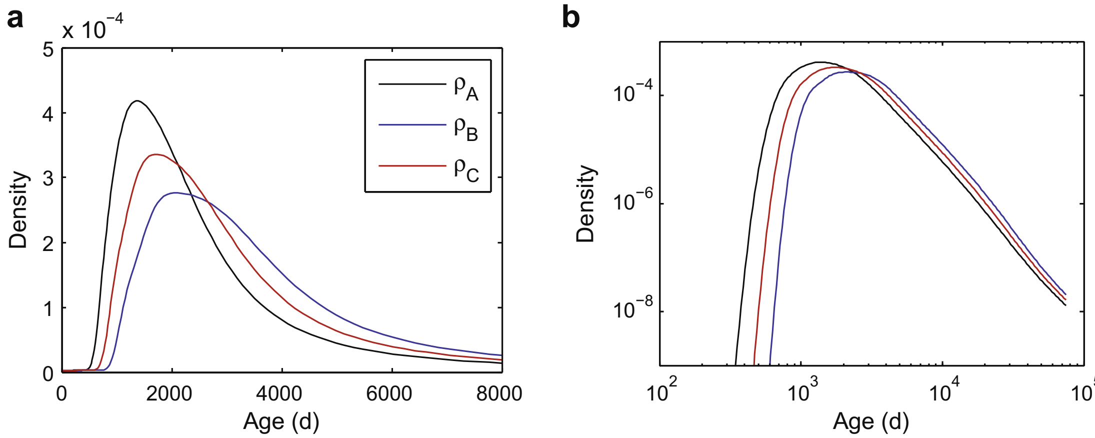

Fig. 6. Simulated age distributions from the streamtube ensemble in (a) linear and (b) logarithmic space. The straight line in log–log space is power-law tailing.

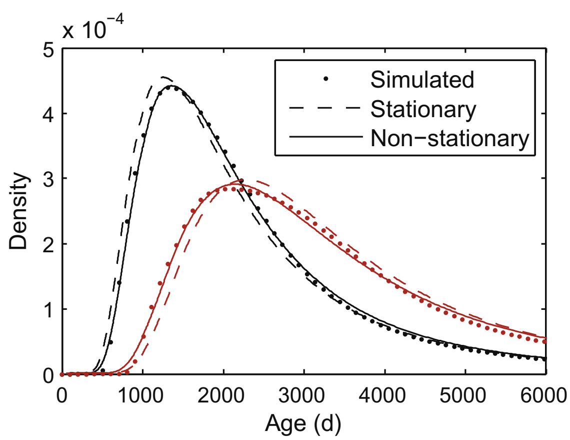

Fig. 7. Comparison of the descriptive models of the age distributions in well A (black) and well B (red). The stationary model represents a joint fit to both wells and the non-stationary model treated the parameters describing the age distributions in each well independently. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from (10) but this will only be valid at the separation distance between the wells. Since the effective models will differ for each of the wells, we will refer to this as the non-stationary deconvolution method. The effective parameters used to find the non-stationary deconvolution were based on those labeled Well A and Well B in Table 2 and provide better approximations of the shapes of the age distributions (Fig. 7). The parameters that best described the non-stationary deconvolution gave the same estimates of the memory function but differed significantly from the velocity and dispersion coefficients that describe the three wells. Curiously, the non-stationary deconvolution provided little improvement over the stationary model in terms of its multi-scale performance. Based solely on a comparison of the fitted parameters, the deconvolution approach does not do a good job of recovering the simulated age distributions regardless of whether a stationary or non-stationary model is used. However, recovering the age distributions is not the primary goal of this study. We cannot conclude whether or not the deconvolution method can describe transport in the example domain without comparing the effective transport models to a forward model of transport.

4.3. Comparison to transport

The accuracy of the model parameters found in Sections 4.1 and 4.2 was evaluated by comparing the effective models to a forward transport model that uses the actual parameter values for the homogeneous and layered aquifer examples. The homogeneous model is translation invariant, so the analytical solution of the transport equation (Eq. (2) without an age dimension) for a Dirac-delta pulse of unit concentration, evaluated at the distance between wells A and B, was used for the transport simulation. Unsurprisingly, the model based on the parameters of the deconvolution matched the solution for transport almost exactly for the homogeneous model (Table 1). The transport model based on the parameters of the deconvolution precisely mimics the age distribution (ρ_{B-A}) in Fig. 4 with the same small discrepancies described in Section 4.1 from numerical approximation of the Laplace transforms. Note that Fig. 4 is expressed in terms of age and density, but the transport simulation is identical for time and normalized concentration, respectively. Overall, the deconvolution provides excellent estimates of the transport behavior in our homogeneous example but we note that this is an idealized example to show the basic functionality of the deconvolution method; natural systems will certainly exhibit more complex behaviors and contain noise which will affect the accuracy of this approach.

Table 2
Estimated parameters for the streamtube based example. V and D are the CTRW based velocity and dispersion coefficients, respectively, which differ from their Fickian equivalents. β is the characteristic exponent and a_1 and a_2 are the upper and lower cutoff scales of age, respectively. The values for Wells A, B, and C are the best fit descriptions of the age distributions.

	Downstream distance (m)	Estimated length (m)	Mean age (d)	V (m/d)	D (m^2/d)	β	a_1	a_2
Well A	5000	4987	2764	2.22	558	1.738	2.46	5.6
Well B	7500	7484	4148	2.24	780	1.737	2.56	5.7
Well C	6200	6186	3429	2.19	673	1.736	2.55	5.6
Stationary deconvolution	2500	2500	1500	1.93	575	1.738	2.52	6.0
Non-stationary deconvolution	2500	2500	1536	2.00	313	1.739	2.48	5.8
Transport simulation	2500	2500	1494	2.03	300	1.740	2.48	5.6

The forward model of transport in the layered aquifer was generated using the streamtube ensemble for a Dirac-delta pulse of unit mass between wells A and B. According to our conceptual model (Fig. 2), all the streamtubes have a uniform length of 2500 m over this interval so the layers only differ in their velocities. The parameters that provide the best fit description of the ensemble averaged breakthrough curve are given in Table 2 under the heading "transport simulation". Relative to the transport model, the stationary deconvolution (a single set of parameters representing both wells) overestimates the dispersion coefficient and slightly underestimates the velocity. The mean age is approximated very well, but the overall distribution is shifted ahead of the transport simulation and exhibits greater positive skewness (Fig. 8, blue solid line). The non-stationary deconvolution gave a slight overestimate of the mean age (Table 2) but provided a better approximation of the transport parameters. The peak of the non-stationary deconvolution was below that of the transport model and this mass deficit near the peak is balanced by overestimation in the tailing. However, this approach provided a good approximated of the shape of the breakthrough curve and exhibited less skewness than the stationary approach (Fig. 8). The non-stationary deconvolution clearly provides the better approximation of the transport behavior in the layered aquifer.

The over- and underestimation of the non-stationary deconvolution is most likely due to the minor errors in the analytical representations of the simulated age distributions. Inspecting the fits in Fig. 7, the analytical representations (solid lines, Fig. 7) also slightly overestimate the tails and have small deviations near the peaks; this is consistent with the result of the non-stationary deconvolution. Comparing the results from the homogeneous and layered examples, it is expected that a more accurate parametric description of the simulated age distributions (Fig. 7) would give a more accurate result. We were unable to find a better fit using the TPL model but it may be possible to improve the fit using a different non-Fickian modeling technique (see Section 1). Even though our fits are imperfect, the non-stationary method works best because it most accurately represents the age distributions in both wells and, to some extent, accounts for scale dependent changes within the domain which are discussed in Section 5. Any effects of the heterogeneity structure that both wells have encountered are "filtered out" by the deconvolution method and the remaining travel time distribution is representative of transport between the two points, even in this non-stationary system.

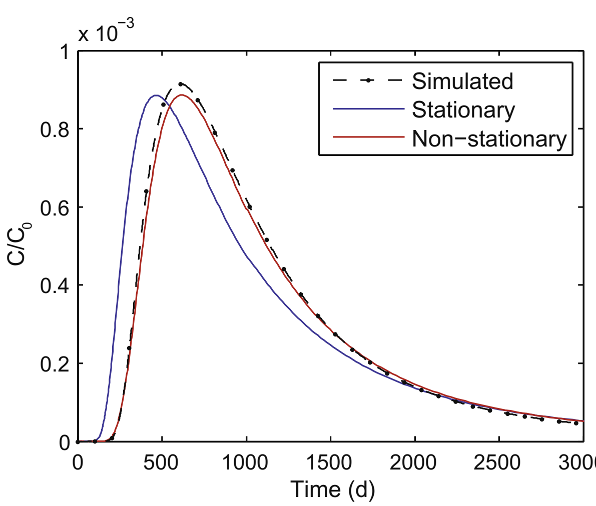

Fig. 8. Performance of the models in a forward transport simulation of the streamtube ensemble at 2500 m. The stationary deconvolution (blue solid) provides an order of magnitude estimates but is outperformed by the non-stationary deconvolution (red solid). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Discussion and summary

Instead of presenting a suite of well-engineered problems that highlight the performance of the proposed method, we have shown two end-member examples that demonstrate the influence of mixing and highlight some of the practical challenges that make inferring transport parameters difficult. In homogeneous systems, the proposed method provides an exact description of the age distributions and transport behavior within the system at any scale, but few natural aquifers are truly homogeneous. In heterogeneous systems, the effectiveness of the method depends on the heterogeneity structure of the aquifer and whether or not that structure can be described with a stationary model. The design of our deceptively simple layered aquifer fundamentally violates the assumption of stationarity. This can be seen in the variance of the streamtube ensemble relative to the homogeneous, Fickian, example (Fig. 9). At small scales the variance of the streamtube ensemble grows linearly with scale but then experiences a non-linear transition into a different regime at about the scale of our monitoring network. The length scales we considered span the transitional regime which makes them difficult to describe because it is the source of the non-stationarity; though it is difficult to see in Fig. 9, the growth rate of the variance at 7500 m is roughly twice the growth rate at 2500 m. In this example, as long as a problem resides entirely within a scale range where the growth of the variance is linear, the system will appear stationary, but the non-stationary deconvolution can approximate transport across the transitional regime in the variance.

The scale dependency of this system is exaggerated most by the fact that we did not allow the streamtubes to interact; the streamtube ensemble has no lateral mixing and none of the mass in the system can experience acceleration or deceleration that will decrease lateral concentration gradients. If the layers were able to interact, the efficiency of the mass exchange between the layers would attenuate the growth rate of the variance. At early times, transport appears nearly Fickian because the longitudinal extent of the mass in the fast streamtubes encompasses the extent of all of the slow moving mass. At later times, the mass in the fastest streamtubes no longer overlap with any of the mass in the slower streamtubes. This results in the enhanced growth rate of the variance and a separated plume. If lateral mixing was permitted, the effects of the non-stationarity would be diminished, and the

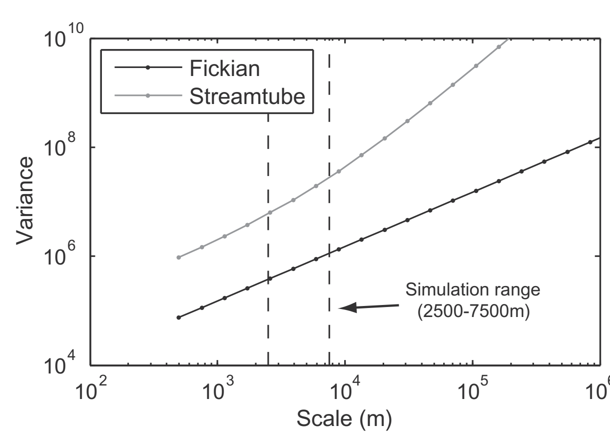

Fig. 9. Scale dependency of the streamtube ensemble. The length scales in this study are in the transitional regime which is difficult to model because it is where the non-stationary of this system manifests. However, the non-stationary deconvolution was still able to capture the effects of the scale dependency.

variance would not grow as fast at long scales, but the system would still exhibit non-Fickian behavior. Avoiding a lengthy discussion on mass transfer and mixing process, it is sufficient to think of mixing conceptually as an adjustment dial that controls the long distance growth rate of the variance as well as the difference between the magnitude of the variance of the homogeneous and layered models at small scales. In this way, the homogeneous domain and the streamtube ensemble are the end-member scenarios of perfect mixing and no mixing, respectively, but we stress that more complicated representations of heterogeneity structures will have different effects on the growth rate of the variance.

Although the layered example is scale dependent, the deconvolution method provided clues to the existence of the non-stationarity early on in our analysis because a set of parameters that simultaneously described both wells could not be found. Even though this study was not concerned with describing the heterogeneity structure of the underlying aquifer, the proposed method revealed important information about whether or not we can assume stationarity. This information is crucial because it dictates which methods will work well and the extent of characterization that is needed for a more detailed study beyond the simplified effective models used here. If it is found that a non-stationary approach must be used, transport can still be described at the known (or estimated) length scales in the domain but other length scales may not be described well with the same model. However, the non-stationary approach also can be used to supplement the monitoring well data. Given three monitoring wells, we can generate travel time or age distributions for up to three more length scales using different pairings of the monitoring wells; in our example, age distributions at 1200, 1300, and 2500 m can be approximated in addition to the known age distributions at 5000, 6200, and 7500 m. These distributions can then be used to help determine how the effective parameters vary with scale and assist in developing a nonlocal model that better accounts for the non-stationarity of the system. It may be possible to improve the methods shown here using more robust models of non-Fickian transport, but the goal of this study was to consider the practical nature of the proposed method as a tool, not to determine which non-Fickian models work best.

The age based methods we presented are promising because they are improvements to artificial tracer tests in a number of ways. First, the age based estimates of the transport parameter estimates are not based on a point source. This removes the possibility of the tracer being released into an anomalously high or low velocity region and eliminates the requirement of maintaining a constant concentration or constant solute flux in the injection well which can affect the analyses of the tracer test. The problem of incomplete mass recovery is also avoided because we are not using a solute and are considering the distribution of the mass of water itself. Second, in a uniform flow field, even if it is a layered aquifer, the wells do not need to be on the same streamline for the age based methods. This is an important difference because if the wells are not aligned with the downstream flow direction in a natural gradient test (non-pumping) an incomplete picture of the breakthrough curve will be obtained if the solute mass is even recovered at all. In the age case, the important distance is that from the well to the recharge source regardless of whether or not the wells are on the same streamline; however, if the aquifer is layered (i.e. Fig. 2) but not laterally homogeneous one must recognize that the wells must be, approximately, on the same streamline. Finally, the age distributions sample the full range of temporal and spatial scales of transport within the example domain; in other words, transport cannot occur in locations where there is no water and all water in the aquifer has age. Thus, the unlikely but preferable scenario where the tracer precedes the contaminant happens automatically with groundwater age. Age distributions are able to provide a more complete picture of the transport behavior and heterogeneity structure of the aquifer than can be found from point to point estimates based on artificial tracers.

Throughout this article we have assumed that suitable estimates of age distributions at the wells are obtainable. Although determining age distributions is not the subject of this article, it is the greatest limitation to our approach and any other approach that directly uses age distributions. We mentioned some of the studies that relate multiple tracers to age distributions in Section 2, including both theoretical [51] and applied [38,46]. Additional research is needed to improve our abilities to determine age distributions from geochemical data collected from monitoring wells [53], including refinements of measurement capabilities for challenging isotopes such as 81-Kr. Measurements of multiple geochemical tracers, sampled repeatedly over time at the same locations, should be invaluable to such efforts and may allow methods like those of Massoudieh and Ginn [51] and Massoudieh et al. [52] to give progressively better approximations of the age distributions. Combinations of age indicators (presence or absence of a compound) and measured isotopic concentrations may also be valuable tools for constraining age distributions, but it will also be necessary to consider the uncertainty and effects of measurement error on the age distributions since these factors will propagate through the method. It is the hope of the authors that the robustness, accuracy, and simplicity of the deconvolution method presented here will motivate further research on determining age distributions from geochemical data, which will allow our method to be applied to field studies in the future.

In summary, we used a deconvolution based method to determine the effective transport properties of two simplified aquifers using only basic well information and groundwater age distributions. We applied Fickian and recently developed non-Fickian models of groundwater age to the simulated distributions and showed that the method provides nearly exact estimates of the transport properties in homogeneous aquifers at all scales. In heterogeneous aquifers, the method provides reasonable order of magnitude estimates of the transport properties under the assumption of stationarity but provides very good approximations of the transport parameters when the non-stationary behavior is accounted for. This method is limited by its assumptions and the difficulty in determining age distributions, but age based methods offer a number of improvements and advantages over traditional tracer tests. Clearly, the method presented here requires further investigation in a wider range of systems to address the affects that different kinds of heterogeneity structures will have on the method, and how to account for transport in multiple spatial dimensions. However, our analyses suggest that it is reasonable to conclude that age distributions can be effective tools for determining the transport behavior in aquifers.

Acknowledgements

The authors thank the anonymous reviewers for timely and constructive feedback that improved the quality of this manuscript. The project described was supported by Award Number P42ES004699 from the National Institute of Environmental Health Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Environmental Health Sciences or the National Institutes of Health.

References

[1] Neuman SP, Tartakovsky DM. Perspective on theories of non-Fickian transport in heterogeneous media. Adv Water Resour 2009;32(5):670–80. http://dx.doi.org/10.1016/j.advwatres.2008.08.005.

[2] Haggerty R, Mckenna SA, Meigs LC. On the late-time behavior of tracer test breakthrough curves. Water Resour Res 2000;36(12):3467–79.
[3] Willmann M, Carrera J, Sánchez-Vila X. Transport upscaling in heterogeneous aquifers: what physical parameters control memory functions? Water Resour Res 2008;44(12):1–13. http://dx.doi.org/10.1029/2007WR006531.
[4] Zheng C, Gorelick SM. Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale. Ground Water 2003; 41(2): 142–155.
[5] Cushman JH, Hu X, Ginn TR. Nonequilibrium statistical mechanics of preasymptotic dispersion. J Statist Phys 1994;75(5/6):859–78.
[6] Haggerty R, Gorelick SM. Multiple-rate mass transfer for modeling diffusion and surface reactions in media with pore-scale heterogeneity. Water Resour Res 1995;31(10):2383–400.
[7] Carrera J, Sanchez-Vila X, Benet I, Medina A, Galarza G, Guimerà J. On matrix diffusion: formulations, solution methods and qualitative effects. Hydrogeol J 1998;6:178–90.
[8] Berkowitz B, Cortis A, Dentz M, Scher H. Modeling non-Fickian transport in geological formations as a continuous time random walk. Rev Geophys 2006;44:1–49. http://dx.doi.org/10.1029/2005RG000178.1.
[9] Benson DA, Wheatcraft SW, Meerschaert MM. The fractional-order governing equation of Levy motion. Water Resour Res 2000;36(6):1413–23.
[10] Dentz M, Berkowitz B. Transport behavior of a passive solute in continuous time random walks and multirate mass transfer. Water Resour Res 2003;39(5):1–20. http://dx.doi.org/10.1029/2001WR001163.
[11] Schumer R, Benson D, Meerschaert M, Baeumer B. Multiscaling fractional advection-dispersion equations and their solutions. Water Resour Res 2003;39(1):1–11. http://dx.doi.org/10.1029/2001WR001229.
[12] Ginn TR. Generalization of the multirate basis for time convolution to unequal forward and reverse rates and connection to reactions with memory. Water Resour Res 2009;45(12):1–9. http://dx.doi.org/10.1029/2009WR008320.
[13] Morales-Casique E, Neuman SP, Guadagnini A. Non-local and localized analyses of non-reactive solute transport in bounded randomly heterogeneous porous media: theoretical framework. Adv Water Resour 2006;29:1238–55. http://dx.doi.org/10.1016/j.advwatres.2005.10.002.
[14] Riva M, Guadagnini A, Neuman SP, Janetti EB, Malama B. Inverse analysis of stochastic moment equations for transient flow in randomly heterogeneous media. Adv Water Resour 2009;32:1495–507.
[15] Meerschaert MM, Zhang Y, Baeumer B. Tempered anomalous diffusion in heterogeneous systems. Geophys Res Lett 2008;35:L17403. http://dx.doi.org/10.1029/2008GL034899.
[16] Zhang Y, Benson DA, Reeves DM. A tempered multiscaling stable model to simulate transport in regional-scale fractured media. Geophys Res Lett 2010;372:L11405. http://dx.doi.org/10.1029/2010GL043609.
[17] Gouze P, Le Borgne T, Leprovost R, Lods G, Poidras T, Pezard P. Non-Fickian dispersion in porous media: 1. Multiscale measurements using single-well injection withdrawal tracer tests. Water Resour Res 2008;44(6):1–15. http://dx.doi.org/10.1029/2007WR006278.
[18] Le Borgne T, Gouze P. Non-Fickian dispersion in porous media: 2. Model validation from measurements at different scales. Water Resour Res 2008;44(6):1–10. http://dx.doi.org/10.1029/2007WR006279.
[19] Zhang Y, Benson DA, Reeves DM. Time and space nonlocalities underlying fractional-derivative models: distinction and literature review of field applications. Adv Water Resour 2009;32(4):561–81. http://dx.doi.org/10.1016/j.advwatres.2009.01.008.
[20] Larocque M, Cook PG, Haaken K, Simmons CT. Estimating flow using tracers and hydraulics in synthetic heterogeneous aquifers. Ground Water 2009;47(6):786–96. http://dx.doi.org/10.1111/j.1745-6584.2009.00595.x.
[21] Becker MW, Shapiro AM. Interpretating tracer breakthrough tailing from different forced-gradient tracer experiment configurations in fractured bedrock. Water Resour Res 2003;39(1):1024.
[22] Kohlbecker MV, Wheatcraft SW, Meerschaert MM. Heavy-tailed log hydraulic conductivity distributions imply heavy-tailed log velocity distributions. Water Resour Res 2006;42(4):1–12. http://dx.doi.org/10.1029/2004WR003815.
[23] Dentz M, Berkowitz B. Exact effective transport dynamics in a one-dimensional random environment. Phys Rev E 2005;72(3):1–12. http://dx.doi.org/10.1103/PhysRevE.72.031110.
[24] Seeboonruang U, Ginn TR. Upscaling heterogeneity in aquifer reactivity via exposure-time concept: forward model. J Contam Hydrol 2006;84(3–4):127–54. http://dx.doi.org/10.1016/j.jconhyd.2005.12.011.
[25] Raats PAC. Convective transport of solutes by steady flows 1. General theory. Agric Water Manage 1978;1:201–18.
[26] Jury WA. Simulation of solute transport using a transfer function model. Water Resour Res 1982;18(2):363–8.
[27] Simmons CS. A stochastic–convective transport representation of dispersion in one-dimensional porous media systems. Water Resour Res 1982;18(4):1193–214.
[28] Engdahl NB, Ginn TR, Fogg GE. Non-Fickian dispersion of groundwater age. Water Resour Res 2012;48(7):1–13. http://dx.doi.org/10.1029/2012WR012251.
[29] Zinn BA, Konikow LF. Effects of intraborehole flow on groundwater age distribution. Hydrogeol J 2007;15(4):633–43. http://dx.doi.org/10.1007/s10040-006-0139-8.
[30] Ginn TR, Haeri H, Massoudieh A, Foglia L. Notes on groundwater age in forward and inverse modeling. Transport Porous Med 2009;79(1):117–34. http://dx.doi.org/10.1007/s11242-009-9406-1.
[31] Sanford W. Calibration of models using groundwater age. Hydrogeol J 2011;19(1):13–6. http://dx.doi.org/10.1007/s10040-010-0637-6.
[32] Ginn TR. On the distribution of multicomponent mixtures over generalized exposure time in subsurface flow and reactive transport: foundations, and formulations for groundwater age, chemical heterogeneity, and biodegradation. Water Resour Res 1999;35(5):1395–407.
[33] Varni M, Carrera J. Simulation of groundwater age distributions. Water Resour Res 1998;34(12):3271–81.
[34] Fogg GE, LaBolle EM, Weissmann GS [Groundwater vulnerability assessment: hydrogeologic perspective and example from Salinas Valley, California]. In: Assessment of non-point source pollution in the Vadose Zone. In: Corwin DL, Loague K, Ellsworth TR, editors. Geophysical monograph series, vol. 108. Washington, DC: AGU; 1999. p. 45–61.
[35] Weissmann GS, Zhang Y, LaBolle EM, Fogg GE. Dispersion of groundwater age in an alluvial aquifer system. Water Resour Res 2002;38(10). http://dx.doi.org/10.1029/2001WR000907.
[36] Cook PG, Love AJ, Robinson NI, Simmons CT. Groundwater ages in fractured rock aquifers. J Hydrol 2005;308(1–4):284–301. http://dx.doi.org/10.1016/j.jhydrol.2004.11.005.
[37] Cornaton F, Perrochet P. Groundwater age, life expectancy and transit time distributions in advective–dispersive systems: 1. Generalized reservoir theory. Adv Water Resour 2006;29(9):1267–91. http://dx.doi.org/10.1016/j.advwatres.2005.10.009.
[38] Leray S, de Dreuzy J-R, Bour O, Labasque T, Aquilina L. Contribution of age data to the characterization of complex aquifers. J Hydrol 2012. http://dx.doi.org/10.1016/j.jhydrol.2010.06.052.
[39] Campana ME. Generation of ground-water age distributions. Ground Water 1987;25(1):51–8.
[40] Goode D. Direct simulation of groundwater age. Water Resour Res 1996;32(2):289–96.
[41] Cornaton FJ. Transient water age distributions in environmental flow systems: the time-marching Laplace transform solution technique. Water Resour Res 2012;48(3):1–17. http://dx.doi.org/10.1029/2011WR010606.
[42] Bethke CM, Johnson TM. Groundwater age and groundwater age dating. Annu Rev Earth Planet Sci 2008;36(1):121–52. http://dx.doi.org/10.1146/annurev.earth.36.031207.124210.
[43] Maloszewski P, Zuber A. Determining the turnover time of groundwater systems with the aid of environmental tracers. J Hydrol 1982;57:207–31.
[44] Bohlke JK, Denver JM. Combined use of groundwater dating, chemical, and isotopic analyses to resolve the history and fate of nitrate contamination in two agricultural watersheds, Atlantic Coastal Plain, Maryland. Water Resour Res 1995;31(9):2319–39.
[45] Cook PG, Böhlke JK. Determining timescales for groundwater flow and solute transport. In: Cook PG, Herczeg A, editors. Environmental tracers in subsurface hydrology. Boston, MA: Kluwer Academic Publishers; 2000. p. 1–30.
[46] Corcho Alvarado JA, Purtschert R, Barbecot F, Chabault C, Rueedi J, Schneider V, et al. Constraining the age distribution of highly mixed groundwater using 39 Ar: A multiple environmental tracer (3H/3He, 85Kr, 39Ar, and 14C) study in the semiconfined Fontainebleau Sands Aquifer (France). Water Resour Res 2007;43(3):1–16. http://dx.doi.org/10.1029/2006WR005096.
[47] Zinn BA, Konikow LF. Potential effects of regional pumpage on groundwater age distribution. Water Resour Res 2007;43(6):1–17. http://dx.doi.org/10.1029/2006WR004865.
[48] LaBolle EM, Fogg GE, Eweis JB. Diffusive fractionation of 3H and 3He in groundwater and its impact on groundwater age estimates. Water Resour Res 2006;42(7):1–11. http://dx.doi.org/10.1029/2005WR004756.
[49] Neumann RB, Labolle EM, Harvey CF. The effects of dual-domain mass transfer on the tritium-helium-3 dating method. Environ Sci Technol 2008; 42(13):4837–43.
[50] Solomon DK, Genereux DP, Plummer LN, Busenberg E. Testing mixing models of old and young groundwater in a tropical lowland rain forest with environmental tracers. Water Resour Res 2010;46:W04518. http://dx.doi.org/10.1029/2009WR008341.
[51] Massoudieh A, Ginn TR. The theoretical relation between unstable solutes and groundwater age. Water Resour Res 2011;47(10):1–6. http://dx.doi.org/10.1029/2010WR010039.
[52] Massoudieh A, Sharifi S, Solomon DK. Bayesian evaluation of groundwater age distribution using a radio-active tracers and anthropogenic chemicals. Water Resour Res 2012;48:W09529. http://dx.doi.org/10.1029/2012WR011815.
[53] DOE. Basic research needs for geosciences: facilitating 21st century energy systems. Report from the workshop held February 21–23, 2007. Office of Basic Energy Sciences, U.S. Department of Energy. <http://www.sc.doe.gov/bes/reports/list.html>.
[54] Harvey CF, Gorelick SM. Temporal moment-generating equations: modeling transport and mass transfer in heterogeneous aquifers. Water Resour Res 1995;31(8):1895–911.
[55] Dentz M, Cortis A, Scher H, Berkowitz B. Time behavior of solute transport in heterogeneous media: transition from anomalous to normal transport. Adv Water Resour 2004;27(2):155–73. http://dx.doi.org/10.1016/j.advwatres.2003.11.002.
[56] Simmons CS, Ginn TR, Wood BD. Stochastic–convective transport with nonlinear reaction: mathematical framework. Water Resour Res 1995; 31(11):2675–88.
[57] Cortis A, Berkowitz B. Computing “anomalous” contaminant transport in porous media: the CTRW MATLAB toolbox. Ground Water 2005;43(6):947–50. http://dx.doi.org/10.1111/j.1745-6584.2005.00045.x.