New stochastic model for dispersion in heterogeneous porous media: 1. Application to unbounded domains

W. S. Verwoerd [a,*]

[a]*Centre for Advanced Computational Solutions,AGLS Division, Lincoln University, P.O Box 84, Lincoln University, Christchurch 7647, New Zealand*

Abstract

A new model of solute dispersion in porous media that avoids Fickian assumptions and that can be applied to variable drift velocities as in non-homogenous or geometrically constricted aquifers, is presented. A key feature is the recognition that because drift velocity acts as a driving coefficient in the kinematical equation that describes random fluid displacements at the pore scale, the use of Ito calculus and related tools from stochastic differential equation theory (SPDE) is required to properly model interaction between pore scale randomness and macroscopic change of the drift velocity. Solute transport is described by formulating an integral version of the solute mass conservation equations, using a probability density. By appropriate linking of this to the related but distinct probability density arising from the kinematical SPDE, it is shown that the evolution of a Gaussian solute plume can be calculated, and in particular its time dependent variance and hence dispersivity. Exact analytical solutions of the differential and integral equations that this procedure involves, are presented for the case of a constant drift velocity, as well as for a constant velocity gradient. In the former case, diffusive dispersion as familiar from the advection-dispersion equation is recovered. However in the latter case, it is shown that there are not only reversible kinematical dispersion effects, but also irreversible, intrinsically stochastic contributions in excess of that predicted by diffusive dispersion. Moreover, this intrinsic contribution has a non-linear time dependence and hence opens up the way for an explanation of the strong observed scale dependence of dispersivity.

*Key words:*Solute dispersion, porous medium, stochastic model, drift velocity , stochastic differential equation.

Introduction

It has been known for a long time that dispersion of solutes in flow through natural porous media deviates from a simple Fickian behaviour in several ways. In recent years, much research (e.g., [1, 2] and references therein) has focussed on tailing of the solute plume as observed in breakthrough curves, which is of considerable

* Corresponding author. Fax: +64-03-3253839; E-mail: verwoerw@lincoln.ac.nz.

importance for practical modelling of solute transport. However the most prominent non-Fickian feature is that the dispersivity α, which would in a Fickian model be simply a materials constant, is found to depend on the spatial scale of the experimental measurement. That is exemplified by the observations reported by Gelhar [3] and shown in Figure 1, that collects together data from many experiments on many different aquifers. From the point of view of explaining the observed phenomenon of scale dependent dispersivity, there are some striking features.

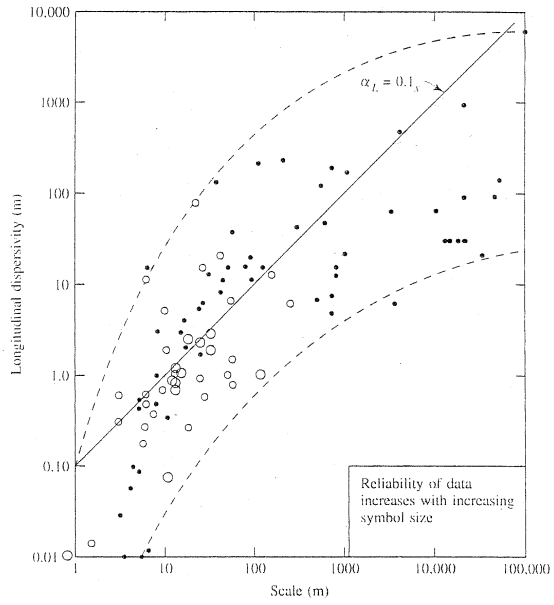

Figure 1: Field measured values of longitudinal dispersivity as function of the scale of measurement. The largest circles represent the most reliable data. Source: Gelhar [3], reproduced from [4].

Firstly, there is the magnitude of the phenomenon. Variations of dispersivity over 6 orders of magnitude are shown. Material properties are known to vary greatly; for example electrical conductivities vary by as much as 23 orders of magnitude, and even hydraulic conductivities of different geological materials differ by as much as 9 orders of magnitude. However, in both those cases the differences are between distinct materials. Variability of a material property for any given material is usually only fractional, as a result of external factors like temperature or pressure.

Secondly, it is a remarkably consistent effect. The differences by up to 3 orders of magnitude of α-values at the same traversal length as shown by Figure 1 are hardly surprising as the data comes from different experiments on different aquifers. What is surprising is that in spite of such variation there is still a clear overall trend, indicated on Figure 1 as a band of values enclosed by dashed lines. For such a consistent trend to emerge there must be a universal underlying mechanism shared by all of the varied observations. Understanding of this mechanism is fundamental to understanding the dispersion phenomenon as a whole.

Thirdly, the dependence on scale is quite complicated. Earlier observations [5] over a more limited range of traversal length Λ suggested a simple proportionality of the dispersivity as given by $\alpha = 0.1\,\Lambda$ represented by the solid line in Figure 1. However the figure shows that this does not fit the data over the whole range.

Moreover, as the plot is logarithmic on both axes, the fact that no straight line fits the data shows that the dependence is more complex than a power law.

One might avoid this conclusion by discounting the less reliable data points and concentrating on the remainder falling mainly in a central range, for which a power law might be adequate. For example, models that link a power law behaviour, over a restricted traversal length range, to fractal dimensions have been presented in the literature [6-8]. Subsequent work [9] extends this to a fitted formula $\alpha = 0.83 (\log \Lambda)^{2.414}$ The effect of replacing a power law with a power of the logarithm, is to create a curve that is non-linear on the log-log plot, with a slope that decreases with increases in scale length.

A similar, but simpler behaviour that would seem roughly consistent with the data, is a combination of two straight lines: one that is steeper than slope 1 for smaller traversal lengths, and another with a slope less than 1 for longer traversal lengths. That implies that the (non-logarithmic) dispersivity curve changes from concave up to concave down with increasing traversal length. More detailed inspection of the data in Figure 1, together with the observation that at laboratory scale Λ -values (below 1 m) dispersivity values are typically of the order 10^{-3} to 10^{-2} m, suggests that even on the logarithmic plot the low range behaviour is concave up, and so an exponential increase rather than a power law is indicated in this range.

While the variability of the data points does not justify detailed conclusions, the overall trend strongly suggests that there is a transitional region near $\Lambda = 100$ m separating distinct low and high range behaviours. An explanation of this behaviour is perhaps the most significant outcome of the work to be presented in this article series.

Many studies of the scale dependence effect have been published before. It has become generally accepted that the heterogeneous nature of natural aquifers is in principle the cause of the growth of dispersivity with the traversal length that is observed. In an approach pioneered by Dagan, Gelhar, Neuman and others [10-13] and extensively reviewed (e.g. [14-18]) macroscopic media properties such as the hydraulic conductivity that appears in solute conservation equations such as the advection-dispersion equation (ADE), are taken as randomly varying. A perturbation approach is used to derive a system of deterministic equations for low-order statistical moments of transport quantities. These equations are then solved analytically or numerically [19-25, 12, 26, 27]. Recent work in this field has moved beyond the low order perturbation approach [28]. The outcome of these stochastic models is generally that the longitudinal dispersivity increases with traversal length and eventually stabilizes at a constant Fickian asymptote, while the transverse dispersivity first reaches a peak and then either decreases to zero or stabilizes at a constant value, depending on the model. The distinction between the behaviour of longitudinal and transverse dispersivity highlights the importance of dimensionality; for example, early work on stratified porous media [29, 30] explored contributions to longitudinal dispersivity due to transverse velocity gradients, an effect only present in higher dimensions, and recent work by Dentz et al has shown first numerically [31] and then analytically [32] that the transverse dispersivity tends asymptotically to small values in two dimensions but stabilizes at much large values in three dimensions.

While the perturbative stochastic approach has been successful for weakly heterogeneous media, deviations have been found where strong heterogeneities are present. For example, strong tailing of tracer breakthrough curves are found in fractured rock [1, 2] and have been interpreted in terms of preferential pathways separating regions of fast and slow solute transport [33, 34]. To deal with such effects, the stochastic model has been extended by including random variation not only in the spatial domain, but also in the time domain giving rise to continuous time random walk (CTRW) models [35, 36]. Another mathematical approach to describe non-Fickian dispersion that has received considerable attention in recent years is the use of fractional derivatives [37-39]. This approach has been shown to be a special case of the CRTW model by Berkowitz et al [40].

Although the stochastic approach to porous dispersion outlined above has been extensively explored and developed to a high level of sophistication both analytically and numerically, there remains a basic issue that it does not address. The common feature in the work referred to is that randomness is introduced, directly or even indirectly such as in the use of fractional derivatives, into the conservation law associated with solute transport. However, when the macroscopic fluid velocity varies as it does in inhomogeneous media, effects of this on

randomness at the more basic level of the kinematics of microscopic fluid motion modelled as random scattering by pore walls, has to be considered.

It may be thought that randomness at this level is adequately taken care of by separating fluid displacements into smooth and random contributions that are treated separately, as in the standard approach to derive Fickian laws, in particular the ADE. However it has to be recognised that in the basic kinematical equation relating displacement and time, $dx = v\,dt$, the velocity acts as a driving coefficient, analogous to the growth rate λ in the population growth equation $dN = \lambda N\,dt$. It is a basic tenet of stochastic differential equation (SPDE) theory [41] that random variation of a driving coefficient cannot in general be represented by merely superimposing random contributions on a deterministic equation. Instead, a more careful analysis leads to special calculus versions such as Ito calculus. In cases where this can be solved analytically, such as population growth, predictions from Ito calculus about probabilities and expectation values differ substantially from those expected from random variation of the deterministic solution.

The purpose of this series of articles is to explore whether application of SPDE theory to kinematical displacements in the presence of a variable velocity, can explain non-Fickian dispersion, and in particular the scale dependence effect. In contrast to customary stochastic treatments, coefficients in the conservation law are not explicitly randomised; in fact, solute mass conservation is in this new approach incorporated through an integral formulation rather than a differential equation such as the continuity equation or ADE. Variations in medium properties, such as hydraulic conductivity or non-porous boundaries, are indirectly represented by the drift velocity changes that they produce. These, in turn, are not taken to be random in the present version of the model but instead well defined velocity profiles are chosen as qualitatively representative of the macroscopic changes in drift velocity that can be expected in natural media. The goal is to discover if the interplay between microscopic, random scattering by pore walls and macroscopic drift velocity change, modifies dispersion significantly. Once that has been shown to happen in cases where the velocity changes in a predetermined way, it might become justified to extend the theory to models where medium properties vary randomly.

While an example such as population growth shows that SPDE effects can be large, the exponential growth implied by that equation is rather an extreme case. It is not a priory clear whether the subtle SPDE effects in the context of the flow equation, are significant enough to contribute to an explanation of the enhancement of dispersivity by many orders of magnitude. If so, it may be expected to give a more fundamental insight into the mechanisms responsible for non-Fickian dispersion. It would also be of interest whether other features such as the observed transition between short and long range behaviour can be explained. In order to give answers to these rather fundamental questions, the study presented here deliberately focuses on simplified or idealised systems, in which it is realistic to obtain exact analytic solutions, or at least approximations that allow interpretation in terms of the underlying mechanisms that can cause such a large effect. For example, most of the calculations are restricted to a 1-dimensional model, although it is clear that higher dimensional effects such as the differences between transverse and longitudinal dispersion are important. Instead, the focus here is on building a model in which the stochastic effects under discussion, and their effects, can be clearly expressed and understood.

As a first step in this quest, it has already been demonstrated [42] that the model studied here leads to the familiar ADE equation in the case where the macroscopic drift velocity is constant.

The present paper now presents an analytical solution to the model when applied to the case of a constant velocity gradient. It shows that indeed there are strong effects on dispersion, beyond what would be expected from superimposing random displacements on the kinematical effects that are associated with accelerating or decelerating flow in a deterministic description. While this observation by itself does not explain scale dependent dispersion, it lays the groundwork for further analysis.

Subsequent articles in this series will address velocity profiles that are more realistic in terms of flow through real inhomogeneous porous media, where velocity gradients can only persist over limited (macroscopic) length scales and fluctuations around an average velocity is the norm.

Section 2 below sets out the conceptual and mathematical basis of the model and the various quantities that need to be calculated in order to apply it. Section 3 is devoted mainly to application of the model to the case of a constant gradient drift velocity, but by way of introduction and for comparison the case of a constant drift velocity is also briefly discussed. In the last section , the implications of this comparison are discussed and leads to the conclusion that there are indeed non-Fickian stochastic effects on dispersion of the kind that are required to explain the scale dependence phenomenon.

1. Model formulation

A more detailed exposition of the model used here, has been given by Kulasiri and Verwoerd[43]. The summary below sets out mainly the notation and key concepts and equations.

1.1. Stochastic modelling of fluid flow

The most basic differential equation that describes just the fluid flow without considering transport, is the deterministic kinematical equation

$$d\vec{x} = \vec{v}(\vec{x},t)\,dt \tag{1}$$

Here $d\vec{x}$ is the microscopic displacement of a fluid element. The microscopic velocity $\vec{v}(\vec{x},t)$ is considered to be composed of an underlying smoothly varying macroscopic velocity, which is perturbed by stochastic perturbations. Physically, the perturbations are caused by deflections due to pore walls; inclusion of perturbations can be seen as a modelling device to avoid giving detailed account of the microscopically complex boundary conditions of the flow by merely representing their final effect.

The aim is to model the perturbations as Wiener processes. Most of the work is based on simplifying equation (1) to a scalar equation in one dimension, as further discussed below. This leaves the velocity as a function of time and one spatial coordinate x. Only the time dependence of the perturbations is modelled by Wiener processes, while the spatial dependence is chosen to conform to a predetermined covariance. This preserves a physically plausible spatial correlation of fluid displacements over a finite range such as the extent of a single pore, while each individual trajectory becomes random.

This can be achieved mathematically by using a spectral expansion (also called a Karhunen-Loeve expansion[44]) that builds up a space and time dependent Wiener process $B(x,t,\omega)$ with a known spatial covariance $C(x_1,x_2)$ by superposition of a set of simpler time-dependent Wiener processes $b_n(t,\omega)$ according to the definition

$$B(x,t,\omega) = \sum_{n=0}^{\infty} \sqrt{\lambda_n}\, f_n(x) b_n(t,\omega) \tag{2}$$

The variable ω acts as a label identifying a particular realisation of the stochastic process. In equation(2), $f_n(x)$ is an orthonormal set of functions that satisfy an integral eigenvalue equation in which $C(x_1, x_2)$ forms the integral kernel, and the λ_n are the corresponding eigenvalues. A consequence of this relationship is that

$$C(x_1,x_2) = \sum_{m=0}^{\infty} \lambda_m f_m(x_1) f_m(x_2) \tag{3}$$

A simple plausible choice for $C(x_1, x2)$ would be to let the correlation between velocity values decrease exponentially as their separation increases:

$$C(x_1, x_2) = \exp\left(-\frac{|x_1 - x_2|}{b} \right) \tag{4}$$

The parameter b in this expression is a chosen correlation length, plausibly taken as a typical pore diameter. The detailed functional form is not important for this work, but equation (4) illustrates how the correlation length acts in effect as a length scale for the spatial coordinates. By virtue of equation (3) also the λ_n and/or $f_n(x)$ depend on b as a parameter.

A more extensive discussion of the Karhunen-Loeve expansion is given in [44], but as will be shown below the expansion is largely eliminated from the stochastic model and so the following qualitative summary of its features is enough to place it in context here:

- The random behaviour, and the functional dependence on the independent variable, are separated into factors in each term.
- This functional dependence can be considered known, being carried by a precisely specified set of orthonormal functions.
- One possibility is that a known set of orthogonal functions are chosen, in which case there is an implied covariance given by equation (3) and it would have to be established if this is physically acceptable.
- Alternatively, the covariance function is predetermined, and the orthogonal functions calculated by solving an integral eigenvalue equation using the chosen covariance as the kernel of the equation.

Using these ideas, a (1-dimensional) stochastic version of equation (1) is stated as

$$dx = u(x)\,dt + \gamma^2 \sum_{n=0}^{\infty} \sqrt{\lambda_n} f_n(x) db_n(t, \omega) \tag{5}$$

The coefficient γ was introduced to define the amplitude of the stochastic term, as further discussed below. Here $u(x)$ is now interpreted as a smoothly varying "drift" velocity. As the Wiener increments by construction average to zero over all stochastic realisations, $u(x)$ may be considered to represent a "typical" trend of the velocity similar to the macroscopic velocity obtained by representative elementary volume (REV) averages introduced in a continuum model. This connection will be further explored below.

For the purposes of stochastic modelling we assume that $u(x)$ is a known, stationary velocity field resulting from an appropriate physical law that describes the effects of external fields such as gravity or boundaries such as external containment of the aquifer, etc. The obvious candidate is Darcy's law, but fluid dynamics equations incorporating thermal convection or other effects might be used in a more elaborate model.

With this assumption, equation (5) conforms to the requirements of an Ito Diffusion (ID) as defined in Oksendal [41] chapter 7. For SPDE's of this type, there are powerful methods derived in SPDE theory that allow direct calculation of statistical properties of its realisations without the need to resort to numerical simulation. Exploiting that possibility is one of the themes of this work.

Before discussing that in detail the model will first be extended to describe solute transport as well.

In applying the mathematical theory to a physical situation, some care is needed in handling dimensions and units of measurement. For example, according to the basic rules of Ito calculus (in particular, the Ito formula [41]) if t has the dimension [T], i.e. is measured in time units, b_n would need to have a dimension [T$^{\frac{1}{2}}$]. As a consequence the coefficients in equation (5) would have different dimensions and become hard to compare, e.g. to assign a value to γ that can be interpreted as a simple amplitude of the stochastic perturbation. At a more basic level, it needs to be recognised that an inherent length or time scale has to be assigned to a Wiener process that represents a physical phenomenon – Brownian motion of a microbe in water takes place on different scales than those of continents interacting via plate tectonics. This point is further discussed in [43].

The most consistent way to deal with this problem is to use the convention implicit in the standard mathematical treatment that the position and time variables in equation (5) have been appropriately scaled to

reflect the physical processes that cause the random displacements. In other words, x and t in equation (5) should really be interpreted as x/ℓ_x and t/ℓ_t where ℓ_x and ℓ_t are scale constants to be determined. Doing this, γ can be taken as a dimensionless constant. For simplicity of notation, the scale constants will be suppressed until we are in a position to determine their values later on.

1.2. Modelling solute transport stochastically

A straightforward approach is to take the deterministic equation that describes solute transport in the macroscopic velocity field $u(x)$, and try to reduce it to a form where it becomes an ID by the addition of a Wiener process perturbation.

Solute mass conservation is the basis of the deterministic transport equation as used for example in the ADE approach. Defining the solute flux vector by

$$\vec{j} = \vec{v}\, c(\vec{x},t) - D_m \cdot \nabla c(\vec{x},t) \tag{6}$$

where c is the solute concentration and D_m the diffusion coefficient, the evolution of the concentration is determined by solute conservation as expressed by the equation of continuity:

$$\frac{\partial c}{\partial t} = -\nabla \cdot \vec{j} \tag{7}$$

Eliminating the flux from equations (6) and (7)

$$\frac{\partial c}{\partial t} = -c\nabla \cdot \vec{v} - \vec{v} \cdot \nabla c + \nabla D_m \cdot \nabla c + D_m \nabla^2 c \tag{8}$$

As detailed elsewhere[43] there are both practical and conceptual problems in treating this as a second SPDE. Instead, an integral version of the solute conservation law is formulated below, and is applied after solving (5) as a SPDE.

Each realisation of the stochastic process $X(t,\omega)$ that is obtained by solving equation (5), represents a possible trajectory of a fluid element through the porous medium. Taking it to be 1-dimensional is meant as a projection of the 3-dimensional flow field onto a single axis.

For example, consider a porous medium contained in a regular cylindrical pipe, and use the cylindrical axis as the X-axis, choosing the macroscopic flow direction as positive. Then, take all trajectories that start from a fixed initial X-value $x = x_0$, by dividing up the circular cross section of the pipe at x_0 into infinitesimal fluid elements. As each of these fluid elements meanders through the porous medium, the X-component of its momentary position varies as a 1-dimensional random walk, i.e. it is represented by a single realisation of the 1-dimensional stochastic process. If the flow is non-laminar at a microscopic scale, another fluid element starting from an identical (y, z) position at $x = x_0$ but at a different time, would describe another realisation.

Suppose that all fluid elements that start from x_0 at one particular time $t = t_0$, and only those, carry identical amounts of the solute, and for simplicity no diffusion is allowed. Then the concentration $c(x, t)$ at a later time t at position x, is the solute amount averaged over all trajectories that traverse the cross-sectional volume element situated at x, in an infinitesimal time interval around t . In other words, this average can be calculated from the probability density that describes the expectation value of the fluid element X-position, over all realisations of the process.

Note that in conceptualising realisations in this way, the ergodic hypothesis is not invoked. We do not use time-dependent stochastic processes to model spatial variations; conceptually, there is only one unchanging porous medium, not an ensemble of them. Spatial variation within the medium, e.g. a position-dependent porosity

or conductivity, is to be included explicitly by taking an assumed drift velocity profile. The stochastic process in this model represents the actual time dependent progress of a fluid element along its trajectory.

In three dimensions, this idea is generalised by considering all infinitesimal fluid elements within a representative elementary volume (REV) as the "particles" whose trajectories are described by individual stochastic realisations that start out from the centre point of the REV.

Consideration of the 3D case highlights a conceptual problem in identifying realisations and trajectories. Only a subset of all possible 3D random walks can occur as trajectories in the physical system, namely those that do not leave the interior of the pores. So taking an average over all realisations is not the same as averaging over all physically meaningful trajectories. Nevertheless, the same average value will be obtained, provided that the microstructure of the porous medium is "sufficiently random" to select the subset without bias.

The problem is less severe in the 1D case, because there cannot be any X-values that cannot be reached in the physical system – if such points existed, they would completely block the flow.

Not all porous media are "sufficiently random". A counterexample is a medium made up of a bundle of straight, uniform, microscopic capillaries. Because here the medium is highly structured, the physically accessible subset of trajectories is a highly biased subset of the collection of realisations. But in this medium there can also be no dispersion – simple plug flow is the only solute transport mode that can occur.

So the stochastic model investigated here only applies to porous media that are sufficiently random, but that can be expected to coincide with the ones of most physical interest. Whether a particular medium is sufficiently random, may have to be decided retrospectively by comparing the predictions of the stochastic model with observed values for that medium.

1.3. Integral formulation of solute conservation

Consider the flow of an incompressible carrier fluid specified by a stationary velocity field $u(x)>0$, and transporting a solute with concentration $c(x, t)$ without diffusion. In the 1D description, c represents the solute mass per unit length. Two distinct cases are considered: (i) the initial value problem where a localised solute distribution, $c(x,0)$, e.g. a gaussian peak, is introduced into the flow at a time $t = 0$; and (ii) the boundary value problem where the value of the concentration is known at the fixed boundary x_0 , for all times earlier than t. In both cases the problem is to find the concentration at a later time t, subject to solute mass conservation.

As shown elsewhere [43] the solution of the initial value problem is

$$c(x,t) = \frac{1}{u(x)} \int_{-\infty}^{\infty} dx' c(x',0)\, u(x')\, P_0(x' \mid x,t)) \tag{9}$$

where we define $P_{t'}(x'|x, t)$ as the probability density over the position variable x', that a fluid element that is known to arrive at the target point (x, t), has originated from a source point x' at a time t'.

Similarly, the boundary value conservation equation is written

$$c(x,t) = \frac{u(x_0)}{u(x)} \int_{-\infty}^{t} dt' c(x_0,t')\, P_{x_0}(t' \mid x,t)) \tag{10}$$

where $P_{x'}(t'|x, t)$ is the probability density over time t', that a fluid element that is known to arrive at the target point (x, t), has originated from x' at a time t'. The two probability densities are not the same, but are related through an appropriate variable transformation.

For a constant drift velocity, the u factors cancel in both (9) and (10) and these equations are self-evident. The presence of the u factors guarantees solute conservation in a variable drift velocity.

For deterministic (non-porous) flow in the absence of diffusion, it is relatively simple to write down either of the required probability densities as an appropriate Dirac delta function expression and show that both

descriptions are equivalent. However, for the stochastic case the probability densities are to be derived from the SPDE and some subtle effects connected to entropy considerations need to be taken into account in the boundary value formulation. Using the boundary value formulation is only essential when the drift velocity is specified piecewise as in the subsequent article, so discussion of these complications will be deferred to that point.

1.4. Calculating a probability density from SPDE theory

In principle numerical simulations of equation (5) could be used to generate a large number of realisations and calculate averages over these, but in simple examples where Ito integration of an SPDE can be done analytically it is found [43] that such results can be unreliable. Fortunately, the Dynkin equation in SPDE theory [41] provides a more direct way to calculate expectation values over all realisations for an ID of the following form

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t \tag{11}$$

This equation in its general form allows the spatial variable and perturbation to be vectors, in which case the coefficients b and σ are a vector and a matrix respectively. The symbols used for them here are to conform to standard notation in the mathematical literature and are not to be confused with other use of b and σ as explicitly defined elsewhere in this article.

The *generator A* of the ID is the (deterministic) differential operator defined by the equation

$$A f(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \tfrac{1}{2}\sum_{i,j}(\sigma\cdot\sigma^T)_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j} \tag{12}$$

Using this concept Dynkin's formula states that

$$E^x[g(X_\tau)] = g(x) + E^x\left[\int_0^\tau A\, g(X_s)\, ds\right] \tag{13}$$

Here, $g(x)$ is an arbitrary function, and the notation E^x means the expectation value taken over all realisations of the ID X_t , that start at $t = 0$ at the particular position x. The integral limit is a stopping time that satisfies $E^x[\tau] < \infty$.

An equivalent differential form of Dynkin's integral formula, sometimes called Kolmogorov's backward equation, is that if we take the left hand side of (13) as the function

$$f(t,x) = E^x[g(X_t)] \tag{14}$$

it is given by the solution of

$$\frac{\partial f}{\partial t} = A f, \qquad t > 0 \quad ; \quad f(0,x) = g(x) \tag{15}$$

To apply these formulas to a given ID, one finds its associated generator to construct the PDE in equation (15). Any solution to this yields a function g of x only, by putting its time argument to zero. Then the expectation value of this function, as its argument evolves over the time-dependent realisations of the ID, is given by equation (14). Notice that the function g for which one can calculate the expectation value cannot be chosen at will, but is dictated by the ID via its generator.

The connection with probability densities is made by using the statistical definition of an expectation value. For example, if the outlined procedure is applied specifically to the fluid displacement ID (5), the result is the expectation value of some function of position of a fluid element, as this position evolves, and the expectation value taken over all fluid trajectories. But this expectation value can by definition also be calculated directly from

the probability density $P_t(x|x', t')$ that a fluid element known to start from a source point $x'=0$ at $t'=0$, will be found at the target point x at a later time t:

$$E^0\left[g(X_t)\right] = \int_{-\infty}^{\infty} dx\, g(x) P_t(x \mid 0,0) \tag{16}$$

Once equation (15) has been solved so f and g are known functions, elimination of the expectation value term between equations (16) and (14) yields an integral equation that can in principle be solved to obtain the probability density.

For the case of a constant drift velocity $u(x) = v_0$, this procedure is relatively straightforward. Equation (15)in this case becomes

$$\frac{\partial f}{\partial t} = v_0 \frac{\partial f}{\partial x} + \frac{1}{2}\gamma^2 \frac{\partial^2 f}{\partial x^2} \tag{17}$$

and has the solution

$$f(x,t) = e^{-a x} e^{a v_0 t - \frac{1}{2} a^2 \gamma^2 t} \tag{18}$$

The integral equation arising from this was shown [42, 43] to be solved by a Gaussian probability density

$$P_t(x \mid x', t') = \mathrm{G}\left((x - x') - v_0 (t - t'), \gamma^2 (t - t')\right) \tag{19}$$

where Γ is the Gaussian defined as

$$\mathrm{G}(x, S^2) = \frac{1}{\sqrt{2\pi S^2}} \exp\left[\frac{-x^2}{2S^2}\right] \tag{20}$$

The relatively simple form of equation (17) is in part due to a remarkable simplification of the generator equation that applies generally when as in equation (5) spatially correlated Wiener processes are used in the stochastic dispersion model, and is not limited to the constant drift velocity case. Comparing equations (5) and (11) we can identify that for this particular ID

$$\sigma(x) = \left(\sqrt{\lambda_0} f_0(x), \sqrt{\lambda_1} f_1(x), \ldots, \sqrt{\lambda_M} f_M(x)\right) \tag{21}$$

where to avoid working with an infinite dimensional vector, it was assumed that the spectral expansion can be truncated at some finite number of terms, M. The generator equation is constructed by evaluating

$$\sigma \cdot \sigma^T = \sum_{n=0}^{M} \lambda_n f_n(x) f_n(x) = C(x,x) \tag{22}$$

For the second equality, equation (3) was used. Not only does M not appear in the final result, but also the need to explicitly solve the eigenvalue problem is completely eliminated in setting up the generator equation.

In fact, it is seen from equation (4) that for the proposed covariance kernel $C(x, x)$ reduces to a constant, the value 1. This reflects the assumption that the covariance of realisations of the $b_n(t, \omega)$ processes of equation (2) at position x, is the same for all positions. This assumption could be avoided by the use of so-called generalised Wiener processes, but as there is no apparent physical reason to do so the present discussion is restricted to the case that $C(x,x) = 1$.

The generator equation (15) in this case reduces to

$$\frac{\partial f}{\partial t} = u(x)\frac{\partial f}{\partial x} + \frac{1}{2}\gamma^2 \frac{\partial^2 f}{\partial x^2} \tag{23}$$

The fact that (23) does not contain any reference to the spatial correlation kernel any more, is very significant. Since solutions of this equation produces expectation values over all representations, it means that while individual realisations of the motion of a fluid element are affected by the functional form of the kernel, all such dependence is averaged out when they are combined. This is not obvious, and might be quite hard to determine from numerical simulations of individual realisations. The ease with which this conclusion follows from the Dynkin equation approach is a striking illustration of its power.

A final remark about the use of SPDE theory to find probability densities for calculating the evolving concentration profile, is that there are some subtle differences between the definitions involved. Inspection of equations (19), (9) and (10) shows that the probability obtained from Dynkin's equation relates to a fixed source point, while the conservation equations require fixed target points. This difference is inconsequential for a constant drift velocity because it turns out that the functional forms of the different probabilities are the same, but for a varying drift velocity that does not hold and more careful analysis is required. More details about that are presented in Appendix A.

1.5. Gaussian initial concentration

In this work, the initial concentration is assumed to have a gaussian spatial profile centred at a location ξ, i.e.

$$c(x',0) = \mathrm{G}(x'-\xi, S^2) \tag{24}$$

A number of arguments can be advanced in favour of this choice:

- The variance of the solute concentration plume is a direct measure of solute dispersion, and by choosing a gaussian form the variance is directly accessible as a parameter of the function that represents it.
- This is particularly useful when the time evolution of the plume stays gaussian. It is well known that this happens when using the ADE equation. In the present model, the combination of a gaussian probability density such as (19) with a gaussian initial concentration, allows equation (9) to take the form of a gaussian convolution integral that can be evaluated analytically to yield a gaussian time evolution as well. In a non-constant drift velocity the probability density may not remain gaussian, but then the evolution of a gaussian profile can demonstrate how this affects dispersion.
- There is no loss of generality in this assumption, as any arbitrary localised concentration can be taken as a superposition of gaussians. In particular, for example, a step function initial concentration as used in breakthrough curve measurements can be seen as an infinite superposition of Dirac delta functions, i.e. zero variance gaussians.

2. Advection and dispersion in unbounded domains

We now turn to the calculation of solute concentration profiles, using the model described in section 1. For comparison it is useful to describe (deterministic) advection and (stochastic) dispersion in the same framework, by taking the probability density for advection (i.e., in the absence of a porous medium) as simply a Dirac delta function that connects the source and target points, as further specified below.

2.1. A constant drift velocity

For advection, it is easily seen that use of the delta function density $P_0(x' \mid x,t) = \delta(x' - x + v_0 t)$ appropriate to $u(x) = v_0$ reduces equation (9) to the plug flow solution for any initial profile (as appropriate, because molecular diffusion is neglected):

$$c(x,t) = c(x - v_0 t, 0) \tag{25}$$

To describe dispersion, putting equations (24) and (19) into (9) yields the gaussian convolution mentioned above, and by completing squares in the exponent it is found [42, 43] that

$$c(x,t) = \mathrm{G}\left(x - X(t), S^2 + \gamma^2 t\right) \quad ; \quad X(t) = \xi + v_0 t \tag{26}$$

This shows that the gaussian peak translates according to the drift kinematics specified by $X(t)$, while its variance increases linearly with time. That agrees with the standard result from the ADE equation, namely diffusion-like dispersion. By appropriate choice of the scaling constants ℓ_x as the correlation length b of equation (4) and $\ell_t = b/v_0$, the microscopic stochastic amplitude γ determines the macroscopic dispersivity α by the relation

$$\alpha = \tfrac{1}{2}\gamma^2 b \tag{27}$$

It follows that the stochastic model presented here is consistent with conventional treatments based on Fickian dispersion and the ADE equation, provided that the drift velocity is constant.

2.2. A constant velocity gradient

The case of solute transport in a changing drift velocity is far more challenging and the simplest case of this where the velocity changes linearly with position, is next considered, i.e.

$$u(x) = v_0 + p(x - x_0) = v_0 \pm \mu^2 (x - x_0) \tag{28}$$

where either p or μ^2 represents the velocity gradient, and the second form just facilitates some intermediate steps where it turns out that the algebra depends on whether the flow accelerates or decelerates.

Although the case of linear acceleration is studied here mainly because it turns out to be mathematically tractable, this case can also be realised physically in a straightforward way, at least over a limited spatial range. For example, consider flow through a pipe filled with a porous medium. If the diameter of the pipe tapers down linearly to a smaller diameter, the flow speeds up linearly along the length of the taper, and would again slow down in the same way if the pipe widens to the original diameter further downstream. Obviously, similar speedup and slowdown flow regions must also occur in flow through a natural porous medium where there is constriction of the flow, e.g. due to impenetrable rocks embedded in a porous matrix. The case represented by equation (28) is an idealisation of this where the acceleration region is unbounded.

The kinematics of transport is compactly described by introducing the time lapse function $w(x, x')$, defined as the time taken to move from the source point to the target point at the drift velocity u:

$$w(x,x') = \int_{x'}^{x} \frac{dx}{u(x)} \tag{29}$$

Inverting this yields χ, the source function that calculates the source point at $t'=0$, given a target point and time:

$$t - w(x,x') = 0 \quad \Rightarrow \quad x' = \chi(x,t) \tag{30}$$

The probability densities that describe advection are specified by

$$P_0(x' \mid x,t) = \delta(x' - \chi(x,t))$$
$$P_{x_0}(t' \mid x,t) = \delta(t' - t + w(x,x_0)) \qquad (31)$$

Putting the specific case of equation (28) into these general definitions yields

$$\chi(x,t) = x_0 + \frac{v_0}{p}\left[e^{-pt} - 1\right] + \left(x - x_0\right)e^{-pt} \qquad (32)$$

For the initial concentration (24) it is appropriate to choose $x_0 = \xi$, and then this equation becomes

$$\chi - \xi = e^{-pt}(x - X(t)) \quad ; \quad X(t) = \xi + \frac{v_0}{p}\left[e^{pt} - 1\right] \qquad (33)$$

The expression $X(t)$ has a straightforward physical interpretation, as the target point that results from kinematical displacement from the particular source point $x'=\xi$. This is easily confirmed by putting $\chi = \xi$ in equation (32). In other words, $X(t)$ represents the trajectory followed by the fluid element that contained the peak value of the gaussian concentration at the initial time. Equation (26) reflects an analogous situation for the case of a constant flow velocity.

The evolution of the concentration in a linearly accelerating flow, without diffusion or dispersion, is found by putting equations (31) to (33) into (9) to obtain

$$c(x,t) = \mathrm{G}(x - X(t), \sigma(t)^2) \quad ; \quad \sigma(t) = S\,e^{pt} \qquad (34)$$

Equation (34) demonstrates some striking features of advection in a velocity gradient :

- A concentration peak that starts off as a gaussian, retains its gaussian shape while propagating.
- The gaussian remains normalised, i.e. the total solute mass is conserved.
- The peak of the gaussian moves at the speed predicted by the kinematics of the fluid flow.
- The width of the gaussian grows exponentially in an accelerating flow ($p > 0$) and shrinks exponentially in a decelerating flow ($p < 0$).
- For $p = 0$ i.e. constant drift velocity, the plug flow solution is regained.

The physical reason that the plume extension grows or shrinks is clearly just the differential in speed ahead and behind the peak position in the accelerating flow. This effect appears superficially like dispersion, but is merely a result of the kinematics of the fluid motion. Unlike "real" dispersion, it is fully reversible in the sense that if, after propagating for a time Δt the acceleration is instantaneously reversed, equation (34) implies that the gaussian plume will return to its original extension after a further time interval Δt. This effect is referred to as *kinematical dispersion* in what follows.

The fact that the shape of the solute profile remains gaussian during advection, is a peculiarity of the linear acceleration. For any other dependence of the velocity on position, distortions of the gaussian shape will result.

We now turn to stochastic dispersion. The examples discussed so far confirm that the model gives sensible results, but hardly justify the elaborate framework. It is only when it is applied to dispersion in an accelerating flow that its full power will become apparent.

The first step in this process is to put equation (28) into the generator equation (23). Its solution is facilitated by transforming to scaled position and time coordinates z and T given by

$$z = \frac{1}{\mu\gamma}\left(\pm\mu^2\,(x - x_0) + v_0\right) \quad ; \quad T = \mu^2 t \qquad (35)$$

The definitions of scaled coordinates are slightly different for accelerating flow (the + sign) and decelerating flow (the – sign case) to give

$$\frac{1}{2}\frac{\partial^2 f}{\partial z^2} \pm z\frac{\partial f}{\partial z} = \frac{\partial f}{\partial T} \tag{36}$$

Separating the variables by substituting an exponential T-dependence gives a hypergeometric equation that can be solved in terms of Hermite polynomials $H_n(z)$ [45]. As detailed elsewhere [43] this allows the integral equation (16) to be solved by use of their orthonormality relations, leading to the probability densities

$$\begin{aligned} P_T^{+}(z \mid z', 0) &= \frac{e^{-z'^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{e^{-(n+1)T}}{n!2^n} H_n(z')H_n(z) \\ P_T^{-}(z \mid z', 0) &= \frac{e^{-z^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{e^{-nT}}{n!2^n} H_n(z')H_n(z) \end{aligned} \tag{37}$$

While these are formal solutions, they do not show a relationship to the gaussian expression (19) that was found for a constant velocity, and to which each of these have to reduce in the limit $p \to 0$. Moreover, there are convergence problems even for a simple calculation of the moments of the probability distribution if term by term integration of the series is attempted.

To solve this problem, an analytical summation formula for the bilinear hermite polynomial sum in equation (37) can be obtained by suitable substitution in the Mehler formula [46]. This yields

$$\sum_{n=0}^{\infty} \frac{e^{-nz}}{n!2^n} H_n(x)\, H_n(y) = \frac{1}{\sqrt{1-e^{-2z}}} \exp\left(-\frac{x^2\, e^{-2z} - 2\,x\,y\,e^{-z} + y^2\, e^{-2z}}{1-e^{-2z}} \right) \tag{38}$$

Application of this formula to each part of equation (37) allows the two distinct probability expressions to be collected together by including the sign of the velocity gradient in the definition of the scaled time parameter T. In other words, also allowing for an arbitrary time origin, redefine T as

$$T = p(t - t') \tag{39}$$

in which case the following closed expression is found for the respective probability densities:

$$P_T^{\pm}(z \mid z', 0) = \mathrm{G}\left(z - z'\, e^{T}, \pm\tfrac{1}{2}(e^{2T} - 1) \right) \tag{40}$$

The re-emergence of a gaussian form, albeit in scaled variables, suggests a connection to the previous results. Various limiting forms express these connections explicitly. For example, in the limit $T \to 0$, equation (40) clearly reduces to a zero variance normalised gaussian, i.e. a Dirac delta function that guarantees that when used in (9) the input concentration is regained as it should be.

A more significant connection is the location of the gaussian peak. When the relation between source and target points, $t - t' = w(x, x')$, is translated to the scaled coordinates by use of (35), it takes the form

$$z = z'\, e^{T} \tag{41}$$

This shows that the peak position in equation (40)is analogous to that in both (19) and (31).

To complete the picture, we establish that probability density obtained here reduces to the previous results in both the deterministic and the constant flow velocity limits. To obtain the deterministic limit of equation (40), we let $\gamma \to 0$, but in this limit the variable transformation (35) becomes singular so that we have to transform back to

the original (x, t) variables before taking the limit. This results in a gaussian with variance proportional to γ^2 , and when $\gamma \to 0$ the resulting δ-function is identical with that in (31).

The constant flow velocity limit can be obtained similarly by first reverting to (x, t) because when p=0, the variable transformation is also singular. Substituting equation (28) and (35) and using $\left(e^{nT}-1\right) \xrightarrow[0]{p} np(t-t')$ it is found that

$$\lim_{p\to 0}\left(z - z' e^{T}\right)^2 = \mp \frac{p}{\gamma^2}\left[(x-x') - v_0(t-t')\right]^2 \tag{42}$$

Writing out the exponential term in the gaussian of (40), and applying the same limit to the numerator in the argument of the exponential as well, it is seen that the factor p cancels and the expression reduces to the one in the constant velocity gaussian given by (19). To complete the derivation the limit needs to be taken in the normalisation prefactor of the gaussian as well, and for this it is noted that a factor p is also contributed by the integration variable, $dz = (p/\gamma\mu)\ dx$. The final result is

$$\lim_{p\to 0} P_T^{\pm}(z \mid z', 0)\, dz = \mathrm{G}\left((x-x') - v_0(t-t'), \gamma^2 (t-t')\right) dx = P_t(x \mid x', t')\, dx \tag{43}$$

This demonstrates complete correspondence between both stochastic probability densities. There is nevertheless a significant new feature that arises from the velocity gradient probability. Consider the following normalisation integrals calculated from equation (40):

$$\int_{-\infty}^{\infty} dz\, P_T(z \mid z', 0) = 1 \quad ; \quad \int_{-\infty}^{\infty} dz'\, P_T(z \mid z', 0) = e^{-T} \tag{44}$$

Both results apply to both the accelerating and decelerating cases, transforming the integrals by using the following integration variable in the form applicable to each case:

$$y = \frac{z - z' e^{T}}{\sqrt{\pm(e^{2T}-1)}} \tag{45}$$

For a constant drift velocity, it does not matter whether the integration is taken over the source or target variable, and the value 1 is obtained in both cases (as also happens in (44) in the constant velocity limit p=0 implying T=0). However for a varying drift velocity, using a particular density expresses different situations when integrated over different variables. Formulated verbally, the first equation in (44) expresses the fact that given the presence of a fluid element at the source point (z',*0*) fluid mass conservation ensures that it must be found at some target point z at the later time T; so summing the probabilities over all possible target points gives 1. On the other hand, if we select any target point z at random, there is no guarantee that there exists a source point from which a fluid element will proceed to z – it is logically conceivable that some points are unreachable, and in such a case summing probabilities over all source points will give 0. The second equation in (44) shows that indeed for the changing flow velocity, this probability sum is generally less than 1.

This observation demonstrates the necessity for selecting the appropriate probability density for the type of integration performed. As shown in Appendix A the spatial density needed for initial value problems is obtained from the second equation in (A3). Applied here to the scaled variables it yields:

$$P_0^{\pm}(z' \mid z, T) = e^{T}\ \mathrm{G}\left(z - z' e^{T}, \pm\tfrac{1}{2}(e^{2T}-1)\right) \tag{46}$$

The time evolution of the initial solute concentration (24) is calculated by using (46) in the conservation equation (9). In practice it is found easiest to transform back to the physical coordinates (x, t). The resulting integrand is complicated, but as the terms in the exponent contributed by the concentration and the probability are

both quadratic expressions in the integration variable x', it is once more a gaussian convolution integral that can be performed in the usual way. The result obtained from the *Mathematica* symbolic algebra program [47] is as follows, expressed in terms of the auxiliary variable $\phi = \gamma^2 / \mu^2$ that might be referred to as a stochastic speedup ratio :

$$c(x,t) = M(x,T)\mathrm{G}\left(x - X(T),\ S^2 e^{2T} \pm \tfrac{1}{2}\phi(e^{2T} - 1)\right) \tag{47}$$

Here $X(T)$ is the kinematical peak position from equation (33) and M is a modulation factor:

$$M(x,T) = 1 - \frac{\pm\phi(e^{2T} - 1)\, p(x - X(T))}{2\sigma^2(T)\, u(x)} \tag{48}$$

The dominating term in (47) is a gaussian centred on the deterministic trajectory $X(t)$, as was the case for both equation (34) (accelerating flow, no dispersion) and equation (26) (constant velocity, with dispersion). The main difference is that here the variance of the gaussian is given by

$$\sigma^2(T) = S^2 e^{2T} \pm \tfrac{1}{2}\phi(e^{2T} - 1) \tag{49}$$

In the deterministic limit, $\gamma \to 0$, so $\phi \to 0$ and only the first term remains, in agreement with equation (34). This means that the first term represents kinematical dispersion while the second one represents the intrinsic stochastic dispersion. Indeed, in the constant velocity limit $p \to 0$ it is found that the intrinsic dispersion term reduces to the diffusive dispersion expression given by the second expression in (34), as it should.

Physical scaling should in principle also be restored to the results of this section as it was done at the end of section 2.1, but that is left as an implicit understanding in order to keep the algebra as simple as possible.

The fact that there is a position dependent modulation factor implies that with dispersion in a velocity gradient, the concentration profile does not remain strictly gaussian. At the (moving) position of the gaussian peak, M=1, but is perturbed from this value on either side of the peak. Excluding the vicinity of the stagnation point $u(x)$ =0, M changes only slowly compared to the exponential change of the gaussian factor as both the numerator and denominator of its perturbation change linearly with x. Hence only a mild distortion of the gaussian shape is expected. In principle, this distortion is a skewing of the peak, since the perturbation of M has opposite signs on both sides of the peak position $X(t)$. The direction of skewing will be backwards for acceleration and forwards for deceleration. Numerical trials confirm that for plausible parameter choices skewing or other non-gaussian distortion is so small as to be hardly visually recognisable.

3. Discussion and conclusions

While non-gaussian distortion of a gaussian input solute concentration is generally small, the phenomenon is significant in principle. It is an example of the fact that dispersion in the presence of a velocity gradient is different from merely superimposing stochastic variation on the deterministic evolution. It was shown in equations (34) and (26) that taken separately, the effects of flow acceleration and dispersion produce perfectly symmetric gaussian concentration peaks; and yet (47) shows that together they give rise to a skewed, quasi-gaussian peak.

This demonstrates the essential difference between SPDE solutions and the mere addition of random effects to a deterministic differential equation, which is the basis of the conventional treatment that leads to the ADE equation.

The same point is made in a more dramatic fashion by considering in more detail the time behaviour of the plume variance given in equation (49). Separately, stochastic dispersion produces a variance that increases linearly with time (equation (26)) and flow acceleration gives an exponential increase (equation (34)). Together,

this same exponential increase is present as the first term in equation (49), but in addition the stochastic term also acquires its own exponential time dependence.

Once more, simple superposition fails and we might describe the result as an interaction between dispersion and the flow acceleration, which is described by the full SPDE treatment. The extent of the interaction may be judged by inspecting the plot in Figure 2.

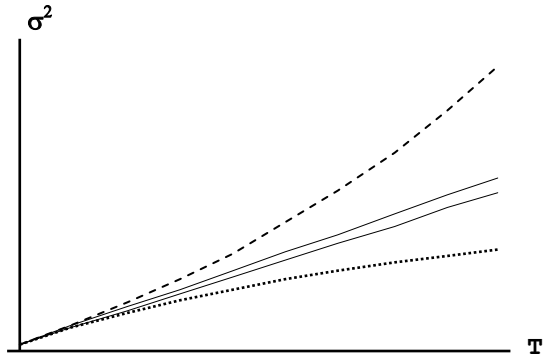

Figure 2. Gaussian plume variance in changing flow velocity. Solid lines show diffusive and kinematical effects only; dashed line is total dispersion for accelerating, and dotted line for decelerating flow

It shows the calculated plume variance according to equation (49) in accelerating flow as a dashed line and in a decelerating flow as a dotted line. For comparison, the two solid lines show the result that would be obtained by superimposing a linear (diffusive) dispersion on two kinematical dispersion rates, for the same acceleration and deceleration rates respectively. The relatively small separation of the solid lines shows that for the parameters chosen for the plot, kinematical dispersion is quite small compared to the straight line diffusive effect that would hold in a constant flow regime over the same time interval. Nevertheless, the interaction of this small kinematical effect and the dispersion produces a much larger final effect, enhancing dispersion for acceleration and suppressing it for deceleration.

The non-linear time dependence is also very significant from the point of view of traditional models of solute transport. One might formally extract a time dependent dispersivity expression by taking the time derivative of equation (49), but it does not reduce any more to a pure materials property that describes the porous medium, as it did for constant flow in the discussion leading up to equation (27). The fact that it becomes time dependent opens the way to an explanation of the "scale dependence" observed in practical dispersivity measurements, as was discussed in section 0. When the spatial scale of an experiment is increased, the time interval over which dispersion is observed also increases, and so scale- and time dependence are just different ways to describe the same phenomenon.

Nevertheless, the connection to the observed scale effect is still quite tenuous. The mathematical treatment in this article applies to a semi-infinite range of linearly increasing flow velocity, limited only on one side by the presence of a stagnation point. This is not a physically realistic situation. That is clear, for example, from the presence of exponential growth for a positive acceleration, which was found in both the deterministic and stochastic solutions. Obviously such exponential growth in any physical quantity is physically sustainable only over a limited spatial or time range. In the physically realistic situation where there is a stable average flow velocity over larger scales, one would expect that intervals of velocity growth would alternate with intervals where the velocity decreases. So velocity fluctuation rather than a sustained velocity growth is physically relevant.

The observation that one can make from Figure 2 that dispersion tends to be enhanced by acceleration and suppressed by deceleration, suggests that a key issue is whether these effects cancel over the extent of a fluctuation. If so, the ADE-type approach where it is assumed that dispersion may be described in terms of an average flow velocity should be in order; but if not, additional effects from fluctuations will change the dispersion behaviour and might give rise to the observed scale effects.

In section 2.2 it was demonstrated that the kinematical component of the dispersion is reversible, and indeed cancels over the extent of a fluctuation (see discussion of equation (34)). However, equation (49) implies that this reversibility is lost when stochastic dispersion is present. That is seen by considering the following simplified situation. A gaussian peak that starts from an initial variance S^2 and propagates for a time τ in a flow with acceleration coefficient $p = P > 0$, attains a variance value of $S^2 e^{2P\tau} + \frac{1}{2}\phi(e^{2P\tau} - 1)$ according to (49). If at that moment the acceleration is reversed, i.e. $p = -P$, the variance after a further time interval τ will by the same argument be given by

$$\sigma^2(2\tau) = \left[S^2 e^{2P\tau} + \tfrac{1}{2}\phi(e^{2P\tau} - 1) \right] e^{-2P\tau} - \tfrac{1}{2}\phi(e^{-2P\tau} - 1) = S^2 + \phi(1 - e^{-2P\tau}) \qquad (50)$$

Kinematical dispersion is represented by the deterministic case, $\phi = 0$, for which cancellation takes place and the original variance is restored. However, equation (50) shows that intrinsic dispersion does not cancel and in fact there is a net increase in dispersion.

Of course, a plume that consecutively penetrates an accelerating and a decelerating region is not quite the same as the hypothetical reversal of the acceleration at a given time, but if the intrinsic dispersion is not even reversible in the simpler situation it is plausible that it will also not cancel over the extent of a velocity fluctuation.

Subsequent articles in this series investigate this further, but need to use approximations to achieve that. The fact that it was possible to perform all derivations for both a constant velocity and a constant velocity gradient analytically without simplifying assumptions beyond those in the underlying stochastic model, creates a baseline against which the further results can be compared.

Acknowledgement

The encouragement and support of my colleague Don Kulasiri is gratefully acknowledged by the author.

Appendix A.

The probability densities delivered by the Dynkin equation approach and those required by the integral conservation laws are not identical. There are four related but different densities:

$P_t(x|x', t')$ – Probability density with respect to x, that a fluid element that is known to have been at the source point x' at a time t', will be found at the target point x at a later time $t > t'$.

$P_{t'}(x'|x, t)$ – Probability density with respect to x', that a fluid element that is found at the target point x at a time t, originated from x' at the earlier time $t' < t$.

$P_x(t|x', t')$ – Probability density with respect to t, that a fluid element that is known to have been at the source point x' at a time t', will be found at x at a later time $t > t'$.

$P_{x'}(t'|x, t)$ – Probability density with respect to t', that a fluid element that is found at the target point x at a time t, originated from x' at the earlier time $t' < t$.

The first of these is the density that is obtained from Dynkin's equation. For this density, since a fluid element known to be at the source point *x'* at *t'* must end up somewhere at a time *t*, the probabilities over all possible target points must add up to one, i.e.

$$\int_{-\infty}^{\infty} dx\, P_t(x \mid x', t') = 1 \tag{A1}$$

Mathematically, this equation states that when the function *P* of the four variables (*x, t, x', t'*) is integrated out over only one of them, namely *x*, all three of the others also drop out of the result. That can only happen if the function is such that it is possible to find a transformation to a new variable $y = y(x, t, x', t')$ in terms of which

$$\int_{y_{-\infty}}^{y_{\infty}} dy\, P(y) = 1 \tag{A2}$$

where $y_{\pm\infty} = y(\pm\infty, t, x', t')$. In other words, the single variable y connects the point (*x',t'*) with another point (*x,t*) with $t > t'$, and $P(y)$ gives the probability that these points are related as source and target points respectively, as a density in the combined variable. In a sense, $P(y)$ is a universal probability density from which the others are derived. To retrieve $P_t(x|x', t')$ one would transform the integration variable in equation (A2) back to *x*, and by the same token $P_{t'}(x'|x, t)$ is obtained if we transform the integration variable to *x'*. So

$$\begin{aligned} P_t(x \mid x', t') &= P\big[y(x,t,x',t')\big]\frac{\partial y}{\partial x} \\ P_{t'}(x' \mid x, t) &= P\big[y(x,t,x',t')\big]\frac{\partial y}{\partial x'} \end{aligned} \tag{A3}$$

The time domain probability densities are obviously given by the same argument as

$$\begin{aligned} P_x(t \mid x', t') &= P\big[y(x,t,x',t')\big]\frac{\partial y}{\partial t} \\ P_{x'}(t' \mid x, t) &= P\big[y(x,t,x',t')\big]\frac{\partial y}{\partial t'} \end{aligned} \tag{A4}$$

The integrations involving these densities are usually performed by transforming the integration variable to *y*. Transforming the integration range is straightforward for the spatial densities but more subtle for the time densities.

a. The constant drift velocity case

From equation (19) it becomes clear that the appropriate choice for the integration variable in this case is

$$y(x,t,x',t') = \frac{(x-x') - v_0 T}{\sqrt{2\gamma^2 T}} \quad ; \quad T \equiv t - t' \tag{A5}$$

The choice of square root was made such that $\partial y/\partial x = 1$ and $y_{\pm\infty} = y(\pm\infty, t, x', t') = \pm\infty$ so that equations (A3) - (A5) lead back to (19).

The spatial density required for initial value problems is obtained from the second equation in (A3) . In this case $\partial y/\partial x' = -1$, but also $y_{\pm\infty} = y(x, t, \pm\infty, t') = \mu\infty$ so that the integration limits need to be exchanged. This cancels the negative sign and hence in this special case an identical expression to (19) is recovered for the spatial density.

The *t*-domain probability density derived from this via (A4) is

$$P_x(t \mid x',t') = \frac{\gamma^2}{\sqrt{\pi}} \frac{x - x' + v_0 T}{(2\gamma^2 T)^{3/2}} \exp\left[\frac{-(x - x' - v_0 T)^2}{2\gamma^2 T} \right] \qquad \text{(A6)}$$

and apart from the sign the same expression is obtained for $P_{x'}(t' \mid x, t)$.

The appropriate integration range for y needs to be found next. For the case of $P_x(t \mid x', t')$, t is the variable while t' plays the role of a constant parameter. Causality requires that the target time must be later than the source time, i.e. $t \in [\, t', \infty\,] \Rightarrow T \in [\, 0, \infty\,]$. For $P_{x'}(t' \mid x, t)$ on the other hand, the roles of t and t' are reversed; so causality is now expressed by $t' \in [-\infty, t\,] \Rightarrow T \in [\infty, 0\,]$, implying that exchange of integration limits will be required.

To determine the corresponding y-range, consider the schematic plots of equation (A5) shown in Figure 3 for the three cases that occur.

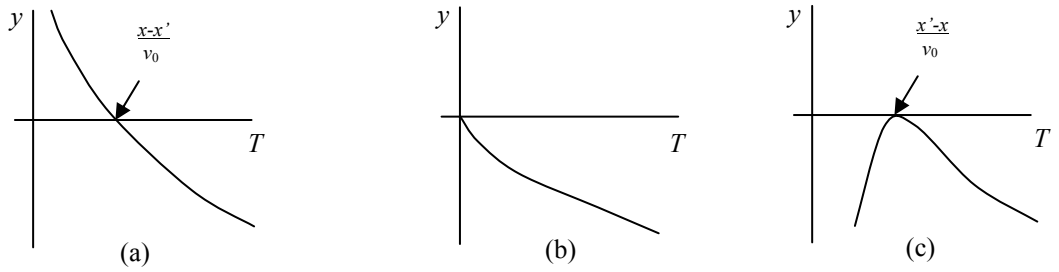

Figure 3. Schematic plots of the constant velocity integration variable y from equation (A5) as function of T, for (a) $x > x'$ (b) $x = x'$, and (c) $x < x'$.

The simplest situation is when $x > x'$. Then part (a) of Figure 3 shows that $T \in [\, 0, \infty\,] \Rightarrow y \in [-\infty, \infty]$.

As an example, the probability of **ever** finding the fluid element that departs from (x', t') at the target position x, is given by

$$p(x \mid x',t') = \int_{t=t'}^{\infty} dt\, P_x(t \mid x',t') = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dy\, e^{-y^2} = 1 \qquad \text{(A7)}$$

Equation (A7) merely says that in a constant flow velocity, a fluid element is certain to pass through every point downstream of the source at some time or another.

A more surprising result is obtained when $x = x'$. This case is illustrated by Figure 3 (b) showing that now $T \in (\, 0, \infty\,] \Rightarrow y \in (\, 0, -\infty\,]$ and so the y-integral extends over only half the range to yield $p(x|x', t') = ½$. To interpret this result, it is best to consider also the 3rd possibility, when $x < x'$.

That is shown in Figure 3 (c). A new complication arises in this figure – for any allowed y value, there are now two T values to consider, i.e. $T(y)$ is a multi-valued function with two branches. However, for the "left-hand" branch in the figure, $T < (x' - x)/v_0$ and this would make the probability density of equation (A6) negative.

Hence only the "right hand" branch, $T \geq (x' - x)/v_0$ is physically meaningful. Translated to the individual time variables, this means that for $x<x'$ their ranges are more restricted than before, to $t \in [t' + (x-x')/v_0, \infty]$ and $t' \in [-\infty, t - (x-x')/v_0]$ respectively. And the corresponding range for y is again $y \in (0, -\infty]$ leading to $p(x|x', t') = ½$ for all $x<x'$.

The results for the probabilities in the various ranges are collected together in Figure 4. The figure shows the target point probability $p(x|x', t')$, plotted as either a function of x for a fixed source point or as function of x' for a fixed target point.

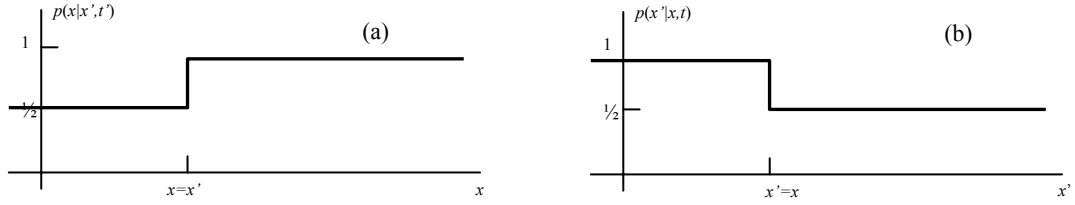

Figure 4. Constant drift velocity; (a) Target point probability dependence on the target point *x,* for a fixed source point *x'* (b) The target point probability as function of source point

Since the source point probability $p(x'|x, t)$ turns out to be identical, the same plots apply to it as well.

For a fixed source point *x'*, the probability that a fluid element will ever be found at a target point downstream from *x'* is 1, just as it would be in deterministic flow. For the source point itself or any point upstream from it, this probability is ½ as a result of stochastic displacements, instead of 0 as it would be in deterministic flow. The source point probability (for a chosen target point) shows just the opposite discontinuity.

The result is somewhat counter-intuitive, as one might have anticipated a decreasing probability of finding the fluid element as one moves further upstream from the source point. This is, of course, true of the probability *density* which has a gaussian factor and represents the probability of finding the element at a certain position during an infinitesimal time interval. However, integrating over all time, the result should rather be compared to Polya's theorem [48] according to which a random walk in 1 dimension eventually reaches every point on the axis, even though the individual displacements have a gaussian distribution. The present result, in effect a random walk within a uniform flow, could be seen as an interpolation between that result and the zero-to-one discontinuity resulting from deterministic displacement in a uniform flow.

In addition to the intrinsic interest of the probability in gaining insight into the cumulative effect of stochastic displacements, its calculation also serves as a prototype of how the integrals required for calculating solute concentrations are to be performed. Indeed, writing down the equation analogous to (A7) for the source point probability $p(x'|x, t)$ and comparing that with (10) shows that this probability might also be interpreted as the stationary concentration profile resulting from a sustained, constant solute injection of magnitude 1 at the position $x=x'$.

b. The case of a drift velocitygradient

The analysis of time domain probability densities for linear velocity change is similar to the constant velocity treatment in principle. However it is more complicated not only because the universal probability variable y is more complex, but in addition the appearance of an additional spatial reference point (the stagnation point) means that instead of the 3 cases ($x > x'$, $x = x'$, $x < x'$) there are now 10 different cases covering all the permutations of x, x' and the stagnation point. For each of these, both the t- and the t'-dependent densities and integration limits are required, and all of this needs to be repeated for both accelerating flow ($p > 0$) and decelerating flow, giving a total of 40 different scenarios in total.

As above, the first task is to calculate the derivatives $\partial y/\partial t = (\partial y/\partial T)\,(\partial T/\partial t)$ and similarly for $\partial y/\partial t'$. Using equation (40) and bearing in mind that for this T is defined by (39) rather than (35), we obtain the probability density as

$$\left.\begin{matrix} P_x(t \mid x',t') \\ P_{x'}(t' \mid x,t) \end{matrix}\right\} = L\frac{p}{\sqrt{\pi}}\frac{e^{2T}}{\left[\pm(e^{2T}-1)\right]^{3/2}}(z-z'e^{-T})\exp\left[-\frac{(z-z'e^{T})^2}{\pm(e^{2T}-1)}\right] \qquad \text{(A8)}$$

Here L is a sign multiplier; the different signs of $\partial T/\partial t$ and $\partial T/\partial t'$ contribute opposite signs, but as for the constant drift velocity case, reversal of the integration limits for t' cancels this. Depending on the relative values of z and z', a further sign multiplier is contributed by the selection of multivalued branches, so the factor L is left to be specified later.

The causality ranges of the time variables are, respectively,

$t \in [\, t', \infty\,]$ $\quad$ $t' \in [-\infty, t\,]$

$\Rightarrow T \in [\, 0, \infty\,]$ $\quad$ $\Rightarrow T \in [\infty, 0]$ $\quad$ for $p>0$

$\quad T \in [\, 0, -\infty\,]$ $\quad$ $\quad T \in [-\infty, 0]$ $\quad$ for $p<0$

The relation between these ranges and those of y are obtained from schematic plots of equation (45), shown in Figure 5 for cases where $z>0$ and also $z'>0$. Note that the dimensionless position variable z defined by equation (35) guarantees, according to (28), that the stagnation point will fall at $z = 0$.

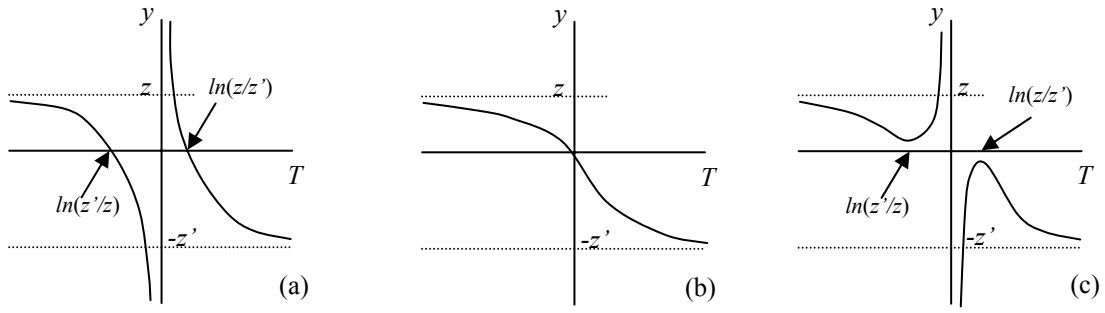

Figure 5 Schematic plots of the velocity gradient integration variable y from equation (45) as function of T, for (a) $z > z'$ (b) $z = z'$, and (c) $z < z'$.

First consider $z > z' > 0$, shown in Figure 5(a). It is seen that the full range of y-values is covered by neither $T \in [\, 0, \infty\,]$ nor $T \in [\, -\infty, 0]$, implying that the probability that a chosen target point is reached is not 1 any more. In fact, for accelerating flow i.e. $p > 0$:

$$p(x \mid x', t') = \int_{t=t'}^{\infty} dt\, P_x(t \mid x', t') = \frac{1}{\sqrt{\pi}} \int_{-z'}^{\infty} dy\, e^{-y^2} = \tfrac{1}{2}\, erfc(-z') \qquad \text{(A9)}$$

As for the constant flow velocity, the probability does not depend on the target point, but here it does depend on the source point.

The reason for this is obviously that if the source point is moved close to the stagnation point, there is an increasing probability that the fluid element will be pushed across it by stochastic fluctuations and into the negative velocity regime, where the ambient flow tends to carry it away from the target point.

This example illustrates that applying the same analysis procedure leads to significantly different, but physically plausible, results for a velocity gradient compared to a constant velocity. Consideration of all 40 scenarios mentioned above has been performed and the behaviour of target point probabilities analysed in a similar way as done above for a constant drift velocity. However, the main result relevant for calculating concentration profiles is that the sign factor L in equation (A8) turns out to be +1 except for two sequences of the source point x', target point x and stagnation point Ω: when either $x < \Omega < x'$ in an accelerating flow, or $x' < \Omega < x$ in a deceleraing flow, the sign factor is $L = -1$. Integration ranges for y are obtained from Figure 5 as needed in each case.

References

[1] M. W. Becker and A. M. Shapiro, Tracer transport in fractured crystalline rock: Evidence of nondiffusive breakthrough tailing, *Water Resources Research* **36** (2000) pp. 1677-1686.

[2] M. W. Becker and A. M. Shapiro, Interpreting tracer breakthrough tailing from different forced-gradient tracer experiment confgurations in fractured bedrock, *Water Resources Research* **39** (2003) pp. 1024.

[3] L. W. Gelhar, *Water Resources Research* **22** (1986) pp. 135S-145S.

[4] C. W. Fetter, Contaminant Hydrogeology, Prentice-Hall, London, 1999.

[5] P. Lallemand-Barres and P. Peaudecerf, *Bulletin, Bureau de Recherches Géologiques et Miniéres* **3/4** (1978) pp. 277-284.

[6] S. P. Neuman, Universal scaling of hydraulic conductivities and dispersivities in geological media., *Water Resources Research* **26** (1990) pp. 1749-1758.

[7] S. P. Neuman and Y.-K. Zhang, A quasi-linear theory of non-Fickian and Fickian subsurface dispersion, 1. Theoretical analysis with application to isotropic media, *Water Resources Research* **26** (1990) pp. 887-902.

[8] Y.-K. Zhang and S. P. Neuman, A quasi-linear theory of non-Fickian and Fickian subsurface dispersion, 2. Application to anisotropic media and the Borden site, *Water Resources Research* **26** (1990) pp. 903-913.

[9] M. Xu and Y. Eckstein, Use of weighted least-squares method in evaluation of the relationship between dispersivity and field scale, *Ground Water* **33** (1995) pp. 905-908.

[10] G. Dagan, Stochastic modeling of groundwater flow by unconditional and conditional probabilities, *Water Resources Research* **18** (1982) pp. 813.

[11] L. W. Gelhar and C. L. Axness, Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resources Research* **19** (1983) pp. 161.

[12] S. P. Neuman, C. L. Winter and C. M. Newman, Stochastic theory of field-scale Fickian dispersion in anisotropic porous media, *Water Resources Research* **23** (1987) pp. 453.

[13] C. L. Winter, C. M. Newman and S. P. Neuman, A perturbation expansion for diffusion in a random velocity field, *SIAM J. Appl. Math.* **44** (1984) pp. 411.

[14] J. H. Cushman, The Physics of Fluids in Hierarchical Porous Media: Angstroms to Miles, Kluwer Academic Publishers, Norwell, MA, 1997.

[15] G. Dagan, Flow and transport in porous formations, Springer Verlag, Berlin, 1989.

[16] L. W. Gelhar, Stochastic subsurface hydrology, Prentice Hall, Englewood Cliffs, NJ, 1993.

[17] Y. Rubin, Transport of inert solutes by groundwater: recent developements and current issues, in G. Dagan and S. P. Neuman, eds., Subsurface Flow and Transport: A Stochastic Approach, Cambridge University Press, Cambridge, 1997, pp. 115-132.

[18] D. Zhang, Stochastic Methods for Flow in Porous Media: Coping with uncertainties, Academic Press, San Diego, CA, 2002.

[19] O. Amir and S. P. Neuman, Gaussian closure of one-dimensional unsaturated flow in randomly heterogenous soils, *Transport in Porous Media* **44** (2001) pp. 355.

[20] W. D. Graham and D. McLaughlin, Stochastic analysis of nonstationary subsurface solute transport, 1. Unconditional moments, *Water Resources Research* **25** (1989) pp. 215.

[21] A. Guadagnini and S. P. Neuman, Nonlocal and localized analysis of conditional mean steady state flow in bounded, randomly nonuniform domains: 1. Theory and computational approach, *Water Resources Research* **35** (1999) pp. 2999.

[22] K. C. Hsu, D. Zhang and S. P. Neuman, Higher-order effects on flow and transport in randomly heterogenous porous media, *Water Resources Research* **32** (1996) pp. 571.

[23] P. Indelman, Averaging of unsteady flows in heterogeneous media of stationary conductivity, *J. Fluid Mech.* **310** (1996) pp. 39.

[24] S.-G. Li, D. McLaughlin and H.-S. Liao, A computationally practical method for stochastic groundwater modeling, *Advances in Water Resources* **26** (2003) pp. 1137-1148.
[25] S. P. Neuman, Eulerian-Lagrangian theory of transport in space-time nonstationary velocity fields: exact nonlocal formalism by conditional moments and weak approximations, *Water Resources Research* **29** (1993) pp. 633.
[26] Y. Rubin, Stochastic modeling of macrodispersion in heterogenous media, *Water Resources Research* **26** (1990) pp. 133.
[27] D. Zhang, Numerical solutions to statistical moment equations of groundwater flow in nonstationary, bounded heterogeneous media, *Water Resources Research* **34** (1998) pp. 529.
[28] D. Zhang and Z. Lu, An efficient, high order perturbation approach for flow in random porous media via Karhunen-Loeve and polynomial expansions, *J Comp Phys* **194** (2004) pp. 773-794.
[29] L. W. Gelhar, A. L. Gutjahr and R. L. Naff, Stochastic analysis of macrodispersion in a stratified aquifer, *Water Resources Research* **15** (1979) pp. 1387-1397.
[30] A. L. Gutjahr, L. W. Gelhar, A. A. Bakr and J. R. Macmillan, Stochastic analysis of spatial variability in subsurface flows, 2, Evaluation and application, *Water Resources Research* **14** (1978) pp. 953-959.
[31] M. Dentz, H. Kinzelbach, S. Attinger and W. Kinzelbach, Temporal behaviour of a solute cloud in a heterogeneous porous medium, 3. Numerical simulations, *Water Resources Research* **38** (2002) pp. 1118.
[32] S. Attinger, M. Dentz and W. Kinzelbach, Exact transverse macro dispersion coefficient for transport in heterogeneous porous media, *Stochastic Environmental Research and Risk Assessment* **18** (2003) pp. 9-15.
[33] V. Cvetkovic, H. Cheng and X. H. Wen, Analysis of non-linear effects on tracer migration in heterogeneous aquifers using travel time statistics, *Water Resources Research* **32** (1996) pp. 1671-1680.
[34] E. K. Webb and M. P. Anderson, Simulation of preferential flow in three-dimensional, heterogeneous conductivity fields with realistic internal architecture, *Water Resources Research* **32** (1996) pp. 533-545.
[35] M. Dentz and B. Berkowitz, Transport behavior of a passive solute in continuous time random walsk and multirate mass transfer, *Water Resources Research* **39** (2003) pp. 1111-1131.
[36] M. Dentz and J. Carrera, Effective solute transport in temporally fluctuating flow through heterogeneous media, *Water Resources Research* **41** (2005) pp. 8414.
[37] Y. Pachepsky, D. A. Benson and W. Rawls, Simulating scale-dependent solute transport in soil with the fractional advective-dispersive equation, *Soil Sci. Soc. Am. J.* **64** (2000) pp. 1234-1243.
[38] S. Schumer, D. A. Benson, M. M. Meerschaert and S. W. Wheatcraft, Eulerian derivation of the fractional advection-dispersion equation, *J. Contam. Hyd.* **48** (2001) pp. 69-88.
[39] S. Schumer, D. A. Benson, M. M. Meerschaert and B. Baeumer, Multiscaling fractional advection-dispersion equations and their solutions, *Water Resources Research* **39** (2003) pp. 1022.
[40] B. Berkowitz, J. Klafter, R. Metzler and H. Scher, Physical pictures of transport in heterogeneous media: advection-dispersion, random walk and fractional derivative formulations, *Water Resources Research* **38** (2002) pp. 1191.
[41] B. Øksendal, Stochastic differential equations - An introduction with applications, Springer Verlag, Berlin, 1998.
[42] W. S. Verwoerd and D. Kulasiri, Theory of diffusions applied to stochastic flow in porous media, *Mathematical and Computer Modelling* **38** (2003) pp. 1453-1459.
[43] D. Kulasiri and W. Verwoerd, Stochastic dynamics - Modeling solute transport in porous media, Elsevier, New York, 2002.
[44] R. G. Ghanem and P. D. Spanos, Stochastic finite elements: a spectral approach, Springer Verlag, Berlin, 1991.
[45] M. Abramowitz and I. A. Stegun, Handbook of mathematical functions, Dover Publications, New York, 1965.

[46] H. M. Srivastava and H. L. Manocha, Treatise on Generating Functions, Ellis Horwood Ltd; Halsted Press., Chichester, 1984.
[47] S. Wolfram, *Mathematica*, Wolfram Research, 2002.
[48] G. Polya, *Mathematische Annalen* **84** (1921) pp. 149-160.