We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



3

A Stochastic Model for Hydrodynamic Dispersion

3.1 Introduction

We have seen in chapter 1 that, in the derivation of advection–dispersion equation, also known as continuum transport model (Rashidi et al. ,1999), the velocity fluctuations around the mean velocity enter into the calculation of solute flux at a given point through averaging theorems. The mean advective flux and the mean dispersive flux are then related to the concentration gradients through Fickian–type assumptions. These assumptions are instrumental in defining dispersivity as a measure of solute dispersion. Dispersivity is proven to be scale dependant.

To address the issue of scale dependence of dispersive fundamentally, it has been argued that a more realistic mathematical framework for modelling is to use stochastic calculus (Holden et al., 1996; Kulasiri and Verwoerd, 1999, 2002). Stochastic calculus deals with the uncertainty in the natural and other phenomena using nondifferentiable functions for which ordinary differentials do not exist (Klebaner, 1998). Stochastic calculus is based on the premise that the differentials of nondifferential functions can have meaning only through certain types of integrals such as Ito integrals which are rigorously developed in the literature. In addition, mathematically well-defined processes such as the Weiner process aid in formulating mathematical models of complex systems. Mathematical theories aside, one needs to question the validity of using stochastic calculus in each instance. In modelling the solute transport in porous media, we consider that the fluid velocity is fundamentally a random variable with respect to space and time and continuous but irregular, i.e., nondifferentiable. In many natural porous formations, geometrical structures are irregular and therefore, as fluid particles encounter porous structures, velocity changes are more likely to be irregular than regular. In many situations, we hardly have accurate information about the porous structure, which contributes to greater uncertainties. Hence, stochastic calculus provides a more sophisticated mathematical framework to model the advection-dispersion in porous media found in practical situations, especially involving natural porous formations. By using stochastic partial differential equations, for example, we could incorporate the uncertainty of the dispersion coefficient and hydraulic conductivity that are present in porous structures such as underground aquifers. The incorporation of the dispersivity as a random, irregular coefficient makes the solution of resulting partial differential equations an interesting area of study. However, the scale dependency of the dispersivity can not be addressed in this manner because the dispersivity itself is not a material property but a constant that depends on the scale of the experiment.

In this chapter we develop one dimensional model without resorting to Fickian assumptions and discuss the methods of estimating the parameters. As of many contracted description of

a natural phenomena the model presented in this chapter has its weaknesses. But we model the fluctuation of the solute velocity due to porous structure and incorporate the fluctuation in the mass conservation of solute. Then we need to characterise the fluctuations so that we can relate them to the porous structure

3.2 The Model Development

The basic assumption on the formulation of this model is that the velocity of solute particles is fundamentally a stochastic variable with irregular but continuous realisations.

Let us consider an infinite cylindrical volume having a cross sectional area, A, within a one-dimensional porous medium with the solute concentration of $C(x, t)$ (Figure 3.1). Within the context of solute transport in heterogeneous groundwater systems, the velocity, $V(x, t)$, the contaminant flux, $J(x, t)$, and $C(x,t)$ are stochastic functions of space and time, having irregular (may be highly irregular) and continuous realisations. Therefore, when formulating the mass conservation model for the solute transport, it is important to use higher order terms in the Taylor expansion.

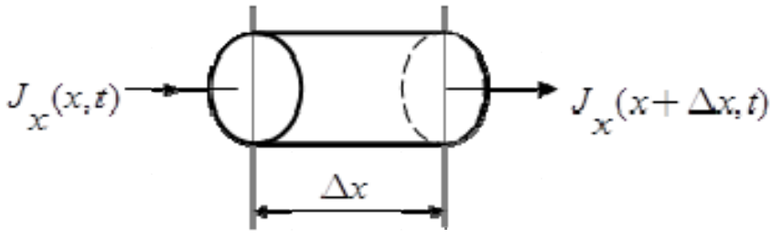


Figure 3.1. An infinitesimal cylindrical volume within a one-dimensional porous medium.

The mass balance for change of solute of the infinitesimal cylindrical volume for a small time interval, Δt , could be written as,

$$\Delta C(x,t) n_e A \Delta x = (J_x(x,t) - J_x(x+\Delta x,t) n_e A \Delta t \ ,$$

$$\left(\frac{\Delta C}{\Delta t}\right)_{x,t} = \frac{(J_x(x,t) - J_x(x+\Delta x,t))}{\Delta x}, \tag{3.2.1}$$

where n_e is the effective porosity of the material, $\Delta C(x,t)$ and Δt are infinitesimal changes of solute concentration, $C(x,t)$, and time, respectively.

For the convenience, let us indicate $J_x(x,t)$ as J_x and $J_x(x+\Delta x,t)$ as $J_{x+\Delta x}$.

The Taylor series expansion can be shown as

$$J_{x+\Delta x} - J_x = \frac{1}{1!}\frac{\partial J_x}{\partial x}\Delta x + \frac{1}{2!}\frac{\partial^2 J_x}{\partial x^2}(\Delta x)^2 + \frac{1}{3!}\frac{\partial^3 J_x}{\partial x^3}(\Delta x)^3 + R(\varepsilon) \ ,$$

where $R(\varepsilon)$ is the remainder of the series.

Let us assume that the terms up to second order would be sufficient to adequately represent the irregular behaviour of the flow, thus, the higher order derivatives of the flux greater than three of the flux are negligible. Therefore, equation (3.2.1) can be written as,

$$\frac{\partial C}{\partial t} = -\frac{\partial J_x}{\partial x} - \frac{1}{2}\frac{\partial^2 J_x}{\partial x^2} dx + R_c(x,t), \tag{3.2.2}$$

where $R_c(x,t) = -\frac{1}{6}\frac{\partial^3 J_x}{\partial x^3}(dx)^2 - \frac{1}{24}\frac{\partial^4 J_x}{\partial x^4}(dx)^3 - R(\varepsilon)$.

Substituting $dx = h_x$, we can write equation (3.2.2),

$$\frac{\partial C}{\partial t} = -\frac{\partial J_x}{\partial x} - \frac{h_x}{2}\frac{\partial^2 J_x}{\partial x^2} + R_c(x,t). \tag{3.2.3}$$

Multiplying the both sides of equation (3.2.3) by *dt*, we obtain,

$$dC = -\left(\frac{\partial J_x}{\partial x} + \frac{h_x}{2}\frac{\partial^2 J_x}{\partial x^2}\right)dt + R_c(x,t)\,dt. \tag{3.2.4}$$

Equation (3.2.4) represents the mass conversation of the solute within the infinitesimal cylindrical volume shown in Figure 3.1. $C(x,t)$ describes the average solute concentration over the cylindrical volume and the smallest possible Δx would increase the accuracy of the stochastic model. However, if Δx is in the same order of magnitude of a typical grain size of the porous media, $C(x,t)$ would lose its meaning. Therefore, it is important to use a realistic Δx value for the scale under consideration for the computational solutions.

Compared to the first term on the right hand side of equation (3.2.4), it is assumed that $R_c(x,t)\,dt \approx 0.$ (However, this assumption needs to be tested for any given situation, especially when porous media is extremely heterogeneous.) Under this assumption,

$$dC = -\left(\frac{\partial J_x}{\partial x} + \frac{h_x}{2}\frac{\partial^2 J_x}{\partial x^2}\right)dt. \tag{3.2.5}$$

The contaminant flux can be expressed in terms of the velocity in the *x* direction and the concentration of the contaminant is expressed as,

$$J(x,t) = V(x,t)\,C(x,t). \tag{3.2.6}$$

The velocity is modelled as,

$$V(x,t) = \bar{V}(x,t) + \xi(x,t), \tag{3.2.7}$$

where $\bar{V}(x,t)$ is the mean velocity of the flow that may be described by Darcy's law,

$$\bar{V}(x,t) = -\frac{K(x)}{n_e(x)}\frac{\partial p}{\partial x},$$

$K(x)$ = an average value of the hydraulic conductivity of the region,

$n_e(x)$ = porosity of the material,

p = pressure head, and

$\xi(x,t)$ = stochastic perturbation of the fluid velocity described by noise correlated in space and δ - correlated in time such that,

$$E\left[\xi(x,t)\right]=0, \tag{3.2.8}$$

$$E\left[\xi(x_1,t_1)\xi(x_2,t_2)\right]=q(x_1,x_2)\delta(t_1-t_2), \tag{3.2.9}$$

where $q(x_1,x_2)$ is the velocity covariance kernel in space, and $\delta(t_1-t_2)$ is the Dirac's delta function.

Since, $\Delta t=t_2-t_1$, we can rewrite equation (3.2.9) as,

$$E\left[\xi(x_1,t_1)\ \xi(x_2,t_1+\Delta t)\right]=q(x_1,x_2)\delta(\Delta t)\ . \tag{3.2.10}$$

In equation (3.2.6) the velocity, $V(x,t)$, is modelled as a stochastic process with the fluctuating component ($\xi(x,t)$) superimposed on the expectation of velocity (mean velocity). The expected value of the velocity $\bar{V}(x,t)$ can be obtained from other considerations without involving Darcy's law. The use of Darcy's law is only one way of prescribing $\bar{V}(x,t)$ within a given region of the porous medium. The fluctuating velocity ($\xi(x,t)$) is assumed to be a zero-mean stochastic process as given by equation (3.2.7), and the two-time correlation function is expressed as a product of a function of space ($q(x_1,x_2)$) and a function of time ($\delta(t_1-t_2)$) in equation (3.2.9) and (3.2.10). (As $\xi(x,t)$ is a zero-mean process, the two-time correlation function and the covariance are the same). The spatial function, $q(x_1,x_2)$, reflects the contribution of the porous structure to the fluctuations and helps characterise the structural influence on the flow. The Dirac's delta function makes the stochastic process behave as a white noise process along the time line. For low velocity situations, such as in aquifers, this is not an unreasonable assumption, but the alternative approach of assembling a time correlation function along the time line would complicate the mathematics a fair deal, and unless the flow is turbulent, there is no justification to take that approach.

Substituting equations (3.2.5) and (3.2.6) into equation (3.2.4), we can obtain,

$$\begin{aligned} dC &= -\frac{\partial}{\partial x}\left[\bar{V}(x,t)C(x,t)+C(x,t)\xi(x,t)\right]dt-\frac{h_x}{2}\frac{\partial^2}{\partial x^2}\left[\bar{V}(x,t)C(x,t)+C(x,t)\xi(x,t)\right]dt, \\ &= -\left(\frac{\partial}{\partial x}+\frac{h_x}{2}\frac{\partial^2}{\partial x^2}\right)\left[\bar{V}(x,t)C(x,t)+C(x,t)\xi(x,t)\right]dt. \end{aligned} \tag{3.2.11}$$

An operator in space can be defined such that,

$$S=-\left(\frac{\partial}{\partial x}+\frac{h_x}{2}\frac{\partial^2}{\partial x^2}\right) \tag{3.2.12}$$

for a given h_x.

Therefore, we can write equation (3.2.11) as,

$$dC = S\left(\overline{V}(x,t)C(x,t)\right)dt + S\left(C(x,t)\xi(x,t)\right)dt. \tag{3.2.13}$$

Equation (3.2.13) is a stochastic (partial) differential equation and both terms on the right hand side need to be integrated as Ito integrals to obtain the concentration. $\xi(x,t)dt$ can be transformed by $d\beta(t)$ which is a Wiener process increments in Hilbert space for a given x. The explanation on the transformation of $\xi(x,t)dt$ to $d\beta(t)$ can be found in Jazwinski (1970).

The transformation of $\xi(x,t)dt$ to $d\beta(t)$ can be understood if we interpret $\xi(x,t)dt$ to be fluctuating component of the travel length of a solute particle during of time interval dt. As we see later this fluctuating travel length component can be expressed as a random vector field ($d\beta(t)$) in an orthonormal set of co-ordinate space (Hilber space). The co-ordinates depend on the covariance kernel $q(x_1, x_2)$, and these are called basis functions in Hilbert space (Hernandez, 1995).

Equation (3.2.13) can be written as,

$$dC = S\left(\overline{V}(x,t)C(x,t)\right)dt + S\left(C(x,t)\,d\beta(t)\right), \tag{3.2.14}$$

And

$$C(x,t) = \int_0^t S\left(\overline{V}(x,t)C(x,t)\right)dt + \int_0^t S\left(C(x,t)\,d\beta(t)\right). \tag{3.2.15}$$

We call equation (3.2.15) a stochastic solute transport model (SSTM) in which $C(x,t)$ is stochastic solute concentration at a given point in space and time; $\overline{V}(x,t)$ is the expected value of stochastic velocity; h_x approximates the spatial discretization interval, dx; and $d\beta(t)$ can be considered as a "noise" term which models the velocity fluctuations.

In this model, instead of the Fickian assumption (equation (1.2.3)), a covariance kernel for the velocity fluctuations $q(x_1, x_2)$ is assumed. The emphasis on modelling of the velocity fluctuation as a random field of second order has its own strengths and weaknesses. The expectation is that by directly incorporating the velocity fluctuation in the solute mass conservation, we expect to reduce the scale dependency because the assumed covariance kernel is a function of the pore structure. The major weakness is that we usually do not have velocity data to construct the kernel for velocity fluctuations. On the other hand, different velocity kernels can be assumed based on plausible reasoning. In this chapter, for the illustration purposes, we assume that the covariance kernel to be exponentially decaying function, $\sigma^2 e^{\frac{-|x_1 - x_2|}{b}}$, where σ^2 is the variance and b is the correlation length.

In SSTM, the first term on the right hand side of equation (3.2.15) can be considered as the drift term of the stochastic partial differential equations (SPDE). The velocity, $\overline{V}(x,t)$, can be considered as the mean velocity of a local region but for the simplicity, we can assume that $\overline{V}(x,t)$ is a constant in this chapter. This assumption is reasonable in practice as we

often calculate the Darcian velocity over a region. As $C(x,t)$ is a stochastic variable, at a given point in space, the integral can be evaluated as an Ito integral to simplify the numerical solution. This argument can also be applied to the dispersion term, the second integral of the right hand side of equation (3.2.15), but $d\beta(t)$ term has to be evaluated. (We use the term "dispersion" instead of "diffusion" to denote the second integral of the right hand side of equation (3.2.15) to allude to the physical nature that is being represented by the term. But in the mathematical literature, the term "diffusion" is used to denote the stochastic (Ito or Stratonovich) integral of analogous stochastic differential equations because it is a diffusion process in mathematical sense.) It should also be noted that $\overline{V}(x,t)$ could also vary from location to location at a higher scale, and we can assume that we can identify regions of sufficiently large magnitudes within which $\overline{V}(x,t)$ can be treated as a smoothly varying function if not a constant. However, this assumption is not necessary if we know the mean velocity profiles based on more aggregate properties. $\xi(x,t)$, on the other hand, relaxes the assumption expressed by equation (1.2.3) and allows us to include a much more realistic assumption as to how the noise behaves. There is some experimental evidence to suggest that the normalised longitudinal velocity covariances for different flow rates had similar exponential behaviour with respect to time in a homogenous porous media (Moroni and Cushman, 2001). However, there is no experimental evidence to date as to how the covariance of velocity fluctuation around the mean velocity, $\xi(x,t)$, behaves with respect to space, but the exponential decay in this situation is quite plausible.

3.3 Construction of $d\beta(t)$ Random Fields Using Spectral Expansion

We summarise the pertinent theoretical background here to explain the assumptions in our model of velocity fluctuations. Let K be a compact set in R^d, over which we define a second order random field, having a covariance function, $q(x,y)$ which is assumed to be square integrable over $K \times K$ (see Hernandez,1995):

$$\iint_{K\times K} |q(x,y)|^2 dx\,dy < +\infty .$$

Let $A: L_2(K) \rightarrow L_2(K)$ be the integral operator

$$A\,\phi(x) = \int_K q(x,y)\,\phi(y)\,dy \tag{3.3.1}$$

with the set of eigen values $\{\lambda_i : i = 1, 2, 3.....\}$ and the set of orthonormal eigenfunctions $\{\psi_i : i = 1, 2, 3 ...\}$. This means

$$A\psi_i = \lambda_i \psi_i, \quad i = 1,2,3..., \tag{3.3.2}$$

$$\text{and } \langle \psi_i, \psi_j \rangle = \delta_{ij} \quad i,j = 1,2,3....., \tag{3.3.3}$$

$$\text{where } \left\langle \psi_i, \psi_j \right\rangle = \int_K \psi_i(x)\,\overline{\psi_j(x)}\,dx . \tag{3.3.4}$$

If q(x,y) is continuous then $\left\langle A\psi_i, \psi_i \right\rangle \geq 0$ for $\psi_i \in L_2(K)$.

Now we can make use of Mercer's theorem (Hernandez, 1995) to express $q(x,y)$ in terms of eigenfunctions (ψ_i) and the corresponding eigenvalues (λ_i) ,

$$q(x,y)=\sum_{k=1}^{\infty}\lambda_k\psi_k(x)\overline{\psi_k(y)}\,, \tag{3.3.5}$$

and convergence is uniform and absolute over $K \times K$. Now let us define a set of (complex) random variables $\{Z_i : i=1, 2, 3...\}$ with the following moments:

$$EZ_n=0, \quad n=1,2,\ldots \text{ , and}$$

$$EZ_n\overline{Z_m}=\lambda_n\delta_{nm}, \quad n=1,2,\ldots..\,.$$

Then the Karhunen-Loeve theorem states that random field $\gamma(x)$ can be expressed as a series expansion of Zi s,

$$\gamma(x):=\sum_{n=1}^{\infty}\psi_n(x)Z_n\,, \tag{3.3.6}$$

which converges in quadratic mean for any x in K. In addition, the following conditions hold,

$$E\gamma(x)\overline{Z_n}=\lambda_n\psi_n(x), \quad n=1,2,\ldots.. \text{ and,}$$

$$\lim_{x\to x^0}E\left|\gamma(x)-\gamma(x^0)\right|^2=0 \qquad \text{for any} \quad x^0 \text{ in } K.$$

What Karhunen-Loeve (KL) expansion does is to model a random field by using two separate mathematical objects: continuous functions, which stem as the solutions of an integral equation (equation (3.3.1)) and random variables. If we assume the random variables (Z_i) to be standard Gaussian ($N(0,1)$), then Karhunen-Loeve expansion takes the following form,

$$\gamma(x):=\sum_{n=1}^{\infty}\sqrt{\lambda_n}\,\psi_n(x)Z_n\,. \tag{3.3.7}$$

KL expansion provides a way of modelling a random field of a single variable, which is considered to be the space variable, in terms of a set of orthonormal basis functions of a Hilbert space and a discrete set of standard Gaussian variables. We extend this development to model the velocity fluctuation of equation (3.2.6), $\xi(x,t)$, which is spatially and temporally correlated in such a way that $\xi(x,t)\in H^0=L_2(R)\times[0,T]$ where H^0 is a separable Hilbert space in which we assume $\xi(x,t)$ to be spatially correlated through a symmetric and positive definite covariance function $q(x_1,x_2)$, and δ-correlated in time (or white in time). As mentioned before this means that velocity changes in a porous medium are influenced by the porous matrix but behaves like white noise with respect to time, i.e., correlation in time is a Dirac delta function. Suppose that the mean velocity $\overline{V}(x,t)$ of a region is 1 m per day, then within a second a solute particle travels 0.01 mm on the average, and within 0.001 days it travels 1 mm, which increases the probability of solute particle changing its course. We can assume that in slow moving fluids, the velocity fluctuation at

one small interval in time is independent of that of within another interval in time. By incorporating appropriate discretisation time and space intervals in the solution process, the white-in-time assumption leads to a simpler but physically plausible model of velocity fluctuation. Therefore, we can express $\xi(x,t)$ as a second-order random field with the moments as given in equation (3.2.9).

Based on stochastic calculus considerations (Unny, 1985), $\xi(x,t)\,dt$ can be considered as a Wiener process increment $d\beta(t)$ in Hilbert space H^0 for a given *x;* and to obtain strong solutions to the stochastic partial differential equation (3.2.14) numerically, this representation is particularly useful. The orthonormal functions, $\{\psi_i : i=1, 2, 3...\}$, provide a basis for H^0, and then we can use the standard Wiener increments ($db_i(t)$) to replace random variables in *KL* expansion (equation (3.3.7)) to construct the random field $d\beta(t)$:

$$d\beta(t) := \sum_{n=1}^{\infty} \sqrt{\lambda_n}\, \psi_n(x)\, db_n(t) = \xi(x,t)dt\,. \tag{3.3.8}$$

Intrinsic in this model is the fact that for any given element $e \in H^0$, the inner product $<\beta(t), e>$ is a real Wiener process having the correlation,

$$E\left[<\beta(t),e_1><\beta(t),e_2>\right] = \iint_{K\times K} q(x_1,x_2)\,e_1(x_2)\,e_2(x_1)\,dx_1\,dx_2\,. \tag{3.3.9}$$

An approximation for equation (3.3.8) can be constructed by considering the first *m* terms in the expansion which converges to a desired accuracy. As the increments in standard Wiener process have zero means and *dt* variances, care must be given to the choice of the time increments within the context of the problem. Smaller the time increments, higher the accuracy of KL expansion as in the case of solving stochastic differntial equations numerically (Kloeden and Platen, 1992).

Equation (3.3.8) gives a series expansion for the stochastic process $d\beta(t)$ but eigen functions and eigenvalues must be found for a given kernel. This task is not straight forward and often involves solving the Fredholm integral equations. Let us assume, based on previous discussion, an exponentially decaying kernel in the *x*-direction in this chapter,

$$q(x_1,x_2) = \sigma^2 e^{\frac{-|x_1-x_2|}{b}}\,. \tag{3.3.10}$$

where x_1 and x_2 are two neighbouring points on *x*. The correlation length *b* signifies the extent to which the correlation of the stochastic process is decayed along the *x*-axis. σ^2 is the variance which acts as an amplitude factor. To obtain the orthornormal functions for the kernel in equation (3.3.10), the following integral equation must be solved within the domain [0, *a*], where *a* is the length (scale) of the experiment,

$$\int_0^a q(x_1,x_2)\;\; f_i(x_2)\;\; dx_2 = \lambda_i f_i(x_1)\,. \tag{3.3.11}$$

Equation (3.3.11) can be written as,

$$f_i(x_1) = \frac{1}{\lambda_i}\int_0^a \sigma^2 e^{\frac{-|x_1-x_2|}{b}}\;\; f_i(x_2)\;\; dx_2\,. \tag{3.3.12}$$

As the upper limit of integration is a constant, equation (3.3.10) is a homogenous Fredholm integral equation of the second kind.

We now discuss the general solution of the integral equation generated by equation (3.3.10) considering the solutions of a general Fredholm integral equation of the form (Polyanin et al., 1998),

$$y(x) + A1 \int_a^b e^{\theta|x-t|}\, y(t)\, dt = f(x)\,, \tag{3.3.13}$$

where $A1$ and θ are constants.

Then the function $y = y(x)$ obeys the following second order linear non-homogeneous differential equation with constant coefficients,

$$y''_{xx} + \theta(2A1-\theta)y = f''_{xx}(x) - \theta^2 f(x)\,. \tag{3.3.14}$$

The boundary conditions for this ordinary differential equation (ODE) have the form,

$$y'_x(a) + \theta y(a) = f'_x(a) + \theta f(a)\,, \text{and} \tag{3.3.15}$$

$$y'_x(b) - \theta y(b) = f'_x(b) - \theta f(b)\,. \tag{3.3.16}$$

Polyanin et al. (1998) explain why the given ODE (equation (3.3.16)) under these boundary conditions determines the solutions of the original integral equation (3.3.12).

Case 1. For $\theta(2A1-\theta) < 0$, the general solution of equation (3.3.13) is given by,

$$y(x) = C_1 Cosh(k\,x) + C_2 Sinh(k\,x) + f(x) - \frac{2A1\,\theta}{k}\int_0^x Sinh[k(x-t)]\, f(t)\, dt\,, \tag{3.3.17}$$

where $k = \sqrt{\theta(\theta - 2A1)}$, and C_1 and C_2 are arbitrary constants.

Case 2. For $\theta(2A1-\theta) > 0$, the general solution of equation (3.3.13) is given by

$$y(x) = C_1 Cos(k\,x) + C_2 Sin(k\,x) + f(x) - \frac{2A1\,\theta}{k}\int_0^x Sin[k(x-t)]\, f(t)\, dt\,, \tag{3.3.18}$$

where $k = \sqrt{\theta(2A1-\theta)}$.

Case 3. For $\theta = 2A1$, the general solution of equation (3.3.13) is given by,

$$y(x) = C_1 + C_2\, x + f(x) - 4\,A1^2 \int_a^x (x-t)\, f(t) dt\,. \tag{3.3.19}$$

The constants C_1 and C_2 are determined by the boundary conditions.

These results (Cases 1, 2, and 3) can be applied to the integral equation (3.3.11) by changing notation, and observing that

$$(2A1-\theta) = \left(\frac{1}{b}\right)\left(\frac{2\sigma^2}{\lambda} + \frac{1}{b}\right) > 0\,. \tag{3.3.20}$$

The eigenvalues of equation (3.3.13) for constant σ^2 over [0, a] are given by

$$\lambda_i = \frac{2\theta\sigma^2}{\omega_i^2 + \theta^2}, \tag{3.3.21}$$

where $\theta = \frac{1}{b}$, and

ω_i s are the roots of the following equation,

$$Tan(\omega_i a) = \frac{2\omega_i\theta}{\omega_i^2 - \theta^2}. \tag{3.3.22}$$

The orthonormal basis functions of the Hilbert space associated with the given exponential kernel (equation (3.3.10)) are the normalised eigenfunctions,

$$f_i(x) = \frac{1}{\sqrt{N}}\left(Sin(\omega_i x) + \frac{\omega_i}{\theta} Cos(\omega_i x) \right), \tag{3.3.23}$$

where,

$$N = \frac{1}{2}a\left(1 + \frac{\omega_i^2}{\theta^2} \right) - \frac{1}{4\omega_i}\left(1 - \frac{\omega_i^2}{\theta^2} \right) Sin(2\omega_i a) - \frac{1}{2\theta}\left(Cos(2\omega_i a) + 1 \right). \tag{3.3.24}$$

As seen from equation (3.3.21), the eigenvalues are dependent on the correlation length b and the variance σ^2 directly and to the length of the experiment, a, indirectly through equation (3.3.22). Equation (3.3.22) is a transendental equation with roots ω_i. Firstly, equation (3.3.22) has to be solved numerically for a given a and b, then the eigenvalues λ_i s and finally the basis functions of H^0. Then the random field $d\beta(t)$ for a given number of terms, m, can be evaluated by using equation (3.3.8) in conjunction with the standard Wiener increments.

3.4 The Dispersion and Travel Length Fluctuations

In this section, we explore the dispersion term in the SSTM to understand its behaviour in simplified settings; as we have seen previously, $S(C(x,t)d\beta(t))$ can be expanded by using the spectral expansion,

$$S(C(x,t)\,d\beta(t)) = S(C(x,t)\sum_{j=1}^{m} f_j(x)\sqrt{\lambda_j}\,db_j(t)). \tag{3.4.1}$$

Let us assume that we use the same realisation of the standard Wiener process across all x in evaluating this term, i.e., db_j is independent of x. Then we can express the dispersion term for this situation,

$$S(C(x,t)\,d\beta(t)) = \sum_{j=1}^{m} S(C(x,t)\,f_j\sqrt{\lambda_j})\,db_j(t). \tag{3.4.2}$$

Now for the sake of argument, let us assume that all the Wiener increments are the same, a situation which is never encountered in computation. Then we can take db_j s out of the summation, and explore the behaviour of the term, $\sum_{j=1}^{m} S(C(x,t) f_j(x) \sqrt{\lambda_j})$.

By substituting for the differential operator S and for $f_j(x)$, after symbolic manipulations, we can express,

$$\sum_{j=1}^{m} S(C(x,t) f_j(x) \sqrt{\lambda_j}) = \left[\sum_{j=1}^{m} \Phi_j\right] \frac{\partial^2 C}{\partial x^2} + \left[\sum_{j=1}^{m} \Psi_j\right] \frac{\partial C}{\partial x} + \left[\sum_{j=1}^{m} \Gamma_j\right] C \,, \tag{3.4.3}$$

where,

$$\Phi_j = \frac{1}{2}\sqrt{\frac{\lambda_j}{N}} \left(-h_x \, Sin(\omega_j x) - \frac{h_x \omega_j}{\theta} Cos(\omega_j x) \right); \tag{3.4.3a}$$

$$\Psi_j = \sqrt{\frac{\lambda_j}{N}} \left(\left(\frac{\omega_j h_x}{\theta} - 1 \right) Sin(\omega_j x) - \left(\omega_j h_x + \frac{\omega_j}{\theta} \right) Cos(\omega_j x) \right); \text{ and} \tag{3.4.3b}$$

$$\Gamma_j = \sqrt{\frac{\lambda_j}{N}} \left(\left(\frac{\omega_j^{\,2} h_x}{2} + \frac{\omega_j^{\,2}}{\theta} \right) Sin(\omega_j x) - \left(\frac{\omega_j^{\,3} h_x}{2\theta} + \omega_j \right) Cos(\omega_j x) \right). \tag{3.4.3c}$$

This exercise shows the complexity of the dispersion term and it is related to the local spatial gradient of the concentration, $\frac{\partial C}{\partial x}$, the second order spatial derivative $\frac{\partial^2 C}{\partial x^2}$, as well as $C(x,t)$. Φ_j , Ψ_j and Γ_j coefficients are also dependent on λ_i s, h_x, and ωi s, and they are analogous to Fourier series, i.e., the coefficients themselves have the form of spectral expansions. Recall that h_x is a scale length introduced to retain the second order derivative of fluxes in the SSTM. We can assume that $h_x \approx O(\Delta x)$, i.e., $\underset{\Delta x \to 0}{Lim} \left(\frac{h_x}{\Delta x} \right) \to 0$ where Δx is the discretization length in x direction. If we make h_x=0 then Φ_j =0, which eliminates the second order spatial derivative of the concentration in equation (3.4.3). Even in a simplified setting of having a common standard Wiener increment ($db_i(t)$) for all x, and assuming a differentiable function for $C(x,t)$, this analysis shows that the dispersion term depicts much more complicated behaviours that cannot be modelled by scale independent simplifying assumptions. Therefore, one should note that as time increases the effects of Wiener process increases, and it would be worthwhile to investigate the dependence of the random field $d\beta(t)$ for different scales of experiments. In other words, from equation (3.2.7), we see that,

$$dX(t) = \bar{V}(x,t)\,dt + \xi(x,t)\,dt \,, \tag{3.4.4}$$

where $X(t)$ is travel distance of a solute particle in x-direction.

As seen in equation (3.3.7), $d\beta(t) \approx \xi(x,t)dt$ is the fluctuating part of the travel length of a solute within the time dt at a given x. Assuming $\bar{V}(x,t)$ is a constant across the domain [0,a], by investigating the nature of the fluctuating component of the travel length as the scale of the experiments changes, we can understand the scale dependency of the dispersion term better.

$d\beta(t)$ is an irregular stochastic function of time with spatial component appearing in normalised eigen functions, $f_j(x)$ .A standard Wiener process increment (db_i) corresponding to a time increment, Δt_i , should be generated for each eigenfunction. $d\beta(t)$ can be approximated as a summation of M terms:

$$d\beta(t) \approx \sum_{i=1}^{M} \sqrt{\lambda_i}\, f_i(x) db_i(t) .$$

Let $\eta_i(x) = \sqrt{\lambda_i}\, f_i(x)$, then $d\beta \approx \sum_{i=1}^{M} \eta_i(x) db_i(t)$.

Recall $db_i(t)$ s are independent zero-mean Gaussian increments with Δt_i variance.

Therefore, $E[d\beta] = E\left[\sum \eta_i(x) db_i\right] = 0$, for a given x.

For a given x, $Var[d\beta] = \sum \eta_i^2(x)\, Var[db_i]$.

If we discretise the time axis equidistantly, $\Delta t = \Delta t_1 = \Delta t_2 = ... = \Delta t_i ...$, then,

$$Var[d\beta] = \left(\sum \eta_i^2(x)\right)\Delta t . \tag{3.4.5}$$

As seen from equation (3.4.5), $d\beta$ is a summation of independent Gaussian processes for a given x value making it a Gaussian process zero mean and variance proportional to Δt. We will make use of equation (3.4.5) in the approximate numerical solution of SSTM for large scale experiments in chapter 4.

3.5 Numerical Solutions of the 1-D SSTM and Their Behaviours

We solve equation (3.2.14) for strong solutions using a finite difference scheme which is based on Euler solution of Ito integral. One dimensional domain is discretised, and the basis of numerical solutions to SPDEs as given by Gaines and Lyons (1997) is adopted. A constant mean velocity is assumed in the scheme. The differential operator S in equation (3.2.11) was expressed as a differential operator using a backward difference scheme. One dimensional spatial length, a ($0 \le x \le a$) on x axis was divided into (k-1) equidistant intervals of small lengths of Δx . The total model time, T, was divided into (n-1) equidistant small intervals of Δt . The space-time grid for the explicit difference scheme that can be used to independently calculate the concentration value at time level t_{n+1} from the concentration values at time intervals t_n , thus preserving the non-antipating nature of Ito integral.

Figure 3.2 show the space-time grid for the explicit difference scheme that can be used to independently calculate the value at time t_{n+1} (denoted by ●) from the values at time t_n (denoted by ○).

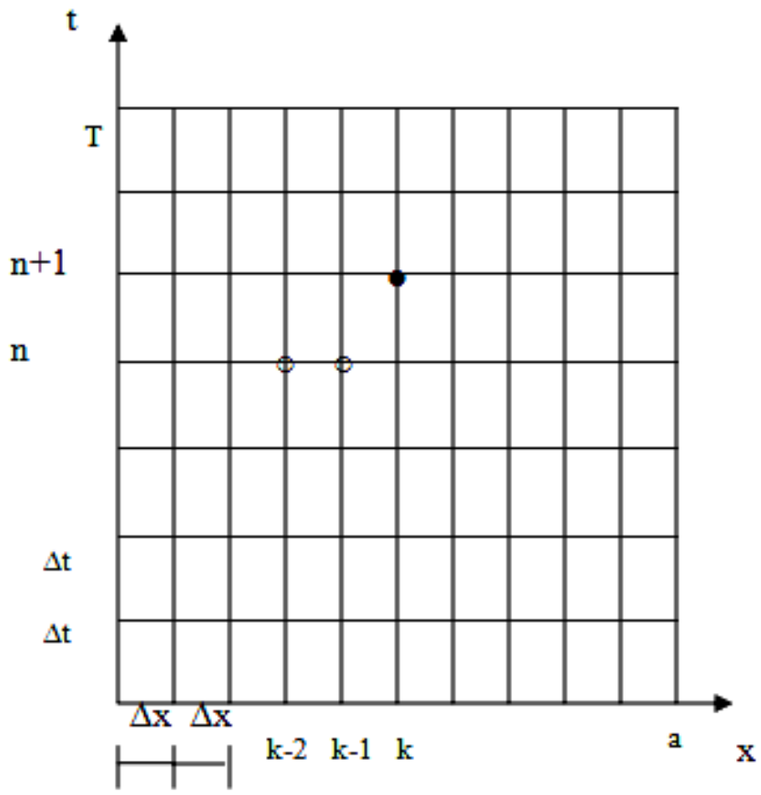


Figure 3.2. An explicit space-time scheme used for the computational solution.

The first derivative of a variable *U* can be described as (Morton and Mayers, 1994),

$$\left(\frac{\partial U}{\partial x}\right)_k^n = \frac{U_k^n - U_{k-1}^n}{\Delta x}, \tag{3.5.1}$$

where U_k^n = value of *U* at the grid point (*k*, *n*).

The second derivative can be given by,

$$\left(\frac{\partial^2 U}{\partial x^2}\right)_k^n = \frac{U_k^n - 2U_{k-1}^n + U_{k-2}^n}{\Delta x^2}. \tag{3.5.2}$$

The operator *S* can be written as,

$$SU = -\left(\frac{\partial U}{\partial x} + \frac{h_x}{2}\frac{\partial^2 U}{\partial x^2}\right),$$

Therefore, we can express the operator in the difference form

$$(SU)_k^n = -\left(\left(\frac{\partial U}{\partial x} \right)_k^n + \frac{h_x}{2} \left(\frac{\partial^2 U}{\partial x^2} \right)_k^n \right). \tag{3.5.3}$$

Substituting the backward difference schemes from (3.5.1) and (3.5.3) and taking $h_x = \Delta x$,

$$(SU)_k^n = -\left(\frac{1}{2\Delta x} \right) \left[3U_k^n - 4U_{k-1}^n + U_{k-2}^n \right]. \tag{3.5.4}$$

The first derivative of U with respect to time can be expressed using a forward difference scheme,

$$\frac{\partial U}{\partial t} = \frac{U_k^{n+1} - U_k^n}{\Delta t} \ . \tag{3.5.5}$$

Applying equation (3.5.1) and (3.5.3) to (3.5.5) and considering the mean velocity for the region as a constant, v, we can obtain the following scheme:

$$\begin{aligned} C_k^{n+1} = C_k^n & - \left(\frac{\Delta t.v}{2\Delta x} \right) * \left[3C_k^n - 4C_{k-1}^n + C_{k-2}^n \right] \\ & - \left(\frac{1}{2\Delta x} \right) * \left[3C_k^n d\beta(t)_k^n - 4C_{k-1}^n d\beta(t)_{k-1}^n + C_{k-2}^n d\beta(t)_{k-2}^n \right], \end{aligned} \tag{3.5.6}$$

where $d\beta(t)$ = Wiener process increments in Hilbert space for a given x.

The explicit difference model (3.5.6) gives the future values of a stochastic variable in terms of the past values preserving the properties of Ito definition of integration with respect to time. This scheme is stable and gives strong solutions to equation (3.2.14), as in many SPDEs, if $\Delta t \leq 10^{-4}$. For example, if a=1000 meters, and if we simulate the solute transport for at least 1000 days, taking Δx =0.01 m and Δt =0.0001 days for stability reasons, we need a grid of $10^5 \, x \, 10^7$. In addition, the evaluation of $d\beta(t)$ for each x involves the summation of a large number of $\eta_i(x)\, db_i$ terms. Because of the computational time it requires to solve the SPDE, we use the scheme given in equation (3.5.6) when a< 10 m, and for larger a values we approximate $d\beta(t)$ term as described later.

We can now investigate the behaviour of the stochastic solute transport model (SSTM) in one-dimension. The main parameters of the SSTM are correlation length, b and variance, σ^2 . As the statistical nature of the computational solution changes with different b and σ^2 , we would like to understand the effect of these parameters on the solution of the model. Furthermore, we attempt to understand these parameters in relation to the hydrodynamic dispersion.

The finite difference numerical schemes are used in the investigation taking the numerical convergence and stability into account for the domain [0, 1]. First we solve equation (3.3.20) for the given values of b and σ^2 **.** It is necessary to find and appropriate number of roots for equation (3.3.20) to produce the desired accuracy of the numerical solution. For the domain

[0, 1], the first 30 values of the roots (ω_i) are generally sufficient. We generate the standard Wiener process increments for 0.001 day time intervals for total of three days. Then the eigenvalues λ_n are computed for the given σ^2 using equation (3.3.21). With these roots, ω and λ_n, we calculate the basis functions using equation (3.3.23). These values are used to compute $d(t)$, the Hilbert space valued Wiener incremental processes, using the *KL* expansion (equation (3.3.7)). Then we calculate the concentration profile for the discretised values of spatial-temporal development for the mean velocity of 0.5 m/day. The numerical solution is implemented in a mathematical software package, Mathematica® (Wolfram Research, 1999).

We use a spatial grid length of 0.1 m for the numerical calculation. The initial concentration distribution profile of 1.0 unit at $x = 0$ is considered and it exponentially decreases through the rest of the domain according to the function, $e^{-5k\Delta x}$, where k = 1, 2, …,10 and Δx = grid size. We begin the numerical scheme with very small numerical concentration values, rather than zero concentrations, to reduce the numerical errors at the beginning of the scheme. The concentration of 1.0 unit is maintained at the boundary of $x = 0$ for the whole time period of the solution to mimic a continuous point source.

To investigate the general behaviour of the SSTM, we obtain the temporal development of the concentration profiles at the mid point of the domain, x = 0.5 m, for various parameter combinations of b and σ^2. The same realisation of the standard Wiener process increments and constant mean velocity of 0.5 m/day are used for all the experiments, so that we would not bias our comparisons.

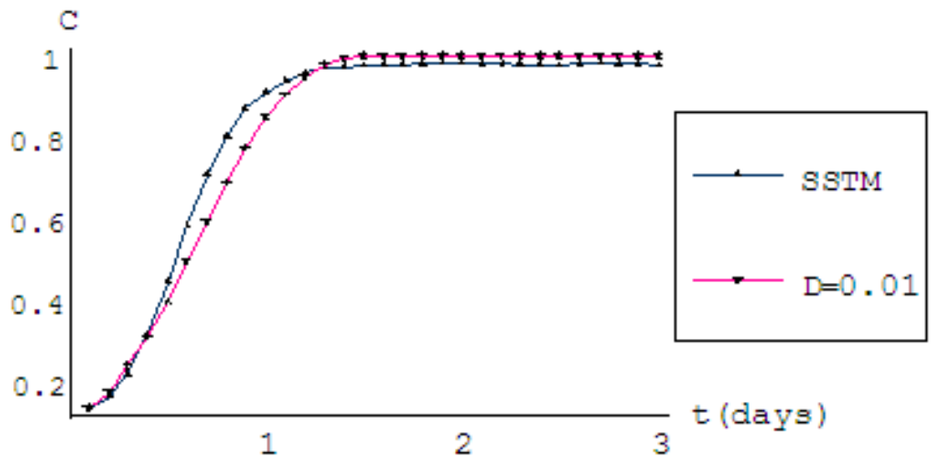


Figure 3.3. Comparison of deterministic advection-dispersion (D = 0.01) and stochastic (σ^2 = 0.001 and b = 0.0001) model concentration profiles. An explicit space-time scheme used for the computational solution.

Figure 3.3 shows that the stochastic model can mimic the solution of the advection-dispersion equation with reasonable accuracy. The concentration breakthrough curves (the time history of concentration at a fixed x) for the SSTM for σ^2 = 0.001 and b = 0.0001, and the deterministic curve for the advection-dispersion equation for dispersion coefficient (D)

of 0.01 m²/day are overlaid in Figure 3.3. We can always find a solution for the SSTM that reasonably represents the deterministic break through curve for any given dispersion coefficient using appropriate values for the parameters, σ^2 and b.

To study the influences of b and σ^2 on the solution of the problem, we keep one parameter constant and change the other within a reasonable range to examine the behavioural change of the concentration breakthrough curves. Figure 3.4 shows the concentration profile at x = 0.5 m for a small value of the variance, σ^2 = 0.0001, when the correlation length, b varies from 0.0001m to 0.25m. Although the range of b varies from 0.0001 to 0.25m (a change of 2500 times) the change of stochasticity (noise level) is negligible and the solutions of the SSTM are independent of b and behave like those of a deterministic model.

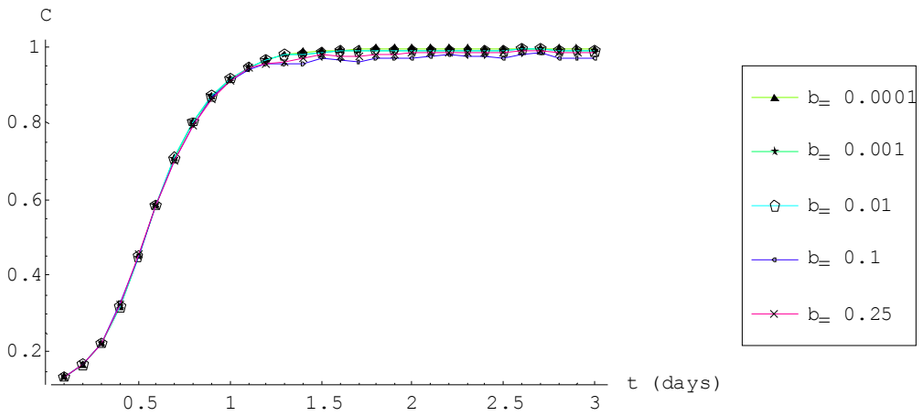


Figure 3.4. Concentration profiles at x = 0.5m for σ^2 = 0.0001.

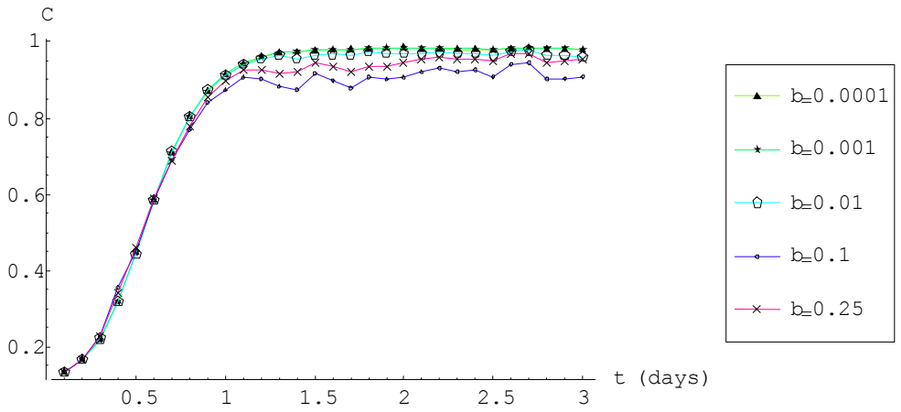


Figure 3.5. Concentration profiles at x = 0.5 m for σ^2 = 0.001.

We gradually increase σ^2 and obtain the concentration profiles for the same regime of b (0.0001 m to 0.25 m) to examine the effect of σ^2. With an increase in σ^2 by 10 times, Figure 3.5 shows that two types of changes have occurred in the concentration profiles;

individual concentration profiles have worse fluctuatiing stochasticity, and there are significant differences between concentration values for different b values at a given time. The high values of the variance not only directly increase the unpredictable nature of the flow but also influence the ways in which b affects the flow. We also observe that with high b values the asymptotic values (sills) of the concentration profiles are lower than the deterministic sill.

In the note that when b is very small, the concentration profile is smooth, but when b is 0.1 m it lowers the sill. By increasing σ^2 by 10 times, to 0.01, and by using the same standard Wiener process increments, we obtain the break through curves as shown in Figure 3.6. The flow tends to be significantly unsteady for larger correlation lengths and still shows smaller fluctuations smaller b values. Furthermore larger values of σ^2 intensities the fluctuation the effect of b significantly.

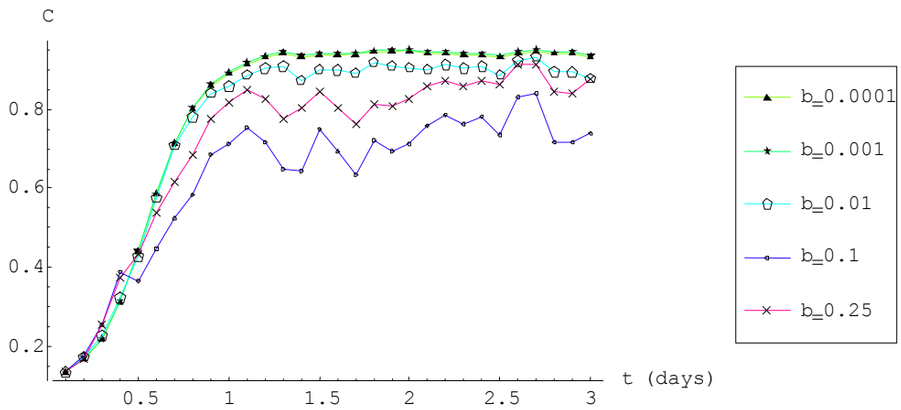


Figure 3.6. Concentration profiles at x = 0.5 m for σ^2 = 0.01.

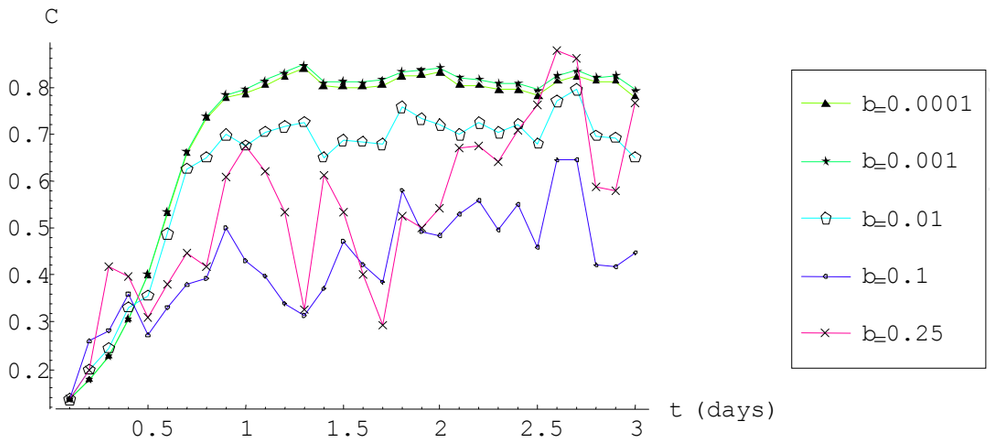


Figure 3.7. Concentration profiles at x = 0.5 m for σ^2 = 0.1.

The concentration profiles for higher b regimes for the increased value of σ^2 (=0.1) are highly random as shown in Figure 3.7. With higher values of b, the concentration profiles become highly irregular making the numerical scheme unstable. Therefore, limit our experiments to smaller b values, that are less than 0.01 m. Figure 4.8 shows that the fluctuation invariably increase with the high σ^2 values and the behaviour of the model continues with the same trend that we noticed earlier, but with enhanced effects.

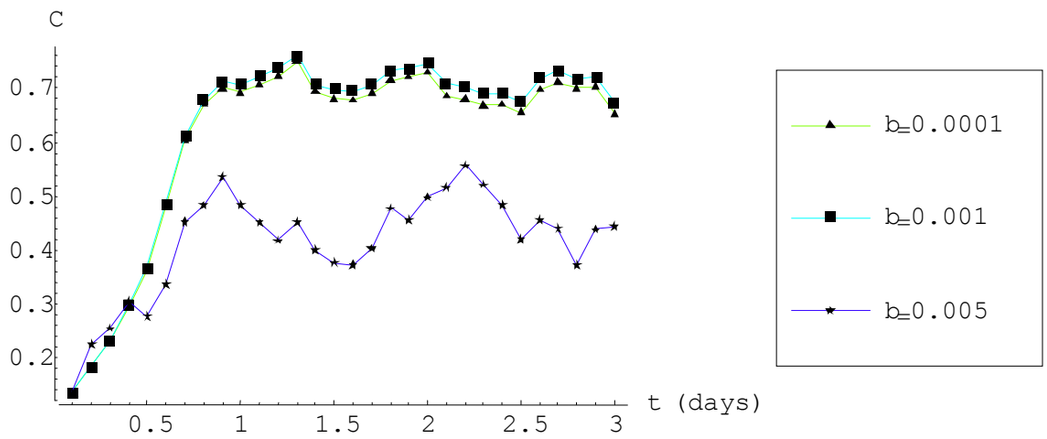


Figure 3.8. Concentration profiles at x = 0.5 m for σ^2 = 0.25.

Figure 3.9 shows the concentration profiles at b = 0.0001 for the range of σ^2 that varies from 0.0001 to 0.25. By comparing in Figure 3.4, when σ^2 was very small, the fluctuation are not distinguishable even for very high b values; however, in Figure 3.9, irrespective of smaller b, σ^2 influenced the behaviour of the flow. It is not possible to differentiate the concentration profiles for very small σ^2, such as 0.0001 and 0.001. With the increase of σ^2 stochasticity increases rapidly, and σ^2 influences the behaviour of the flow more than b does.

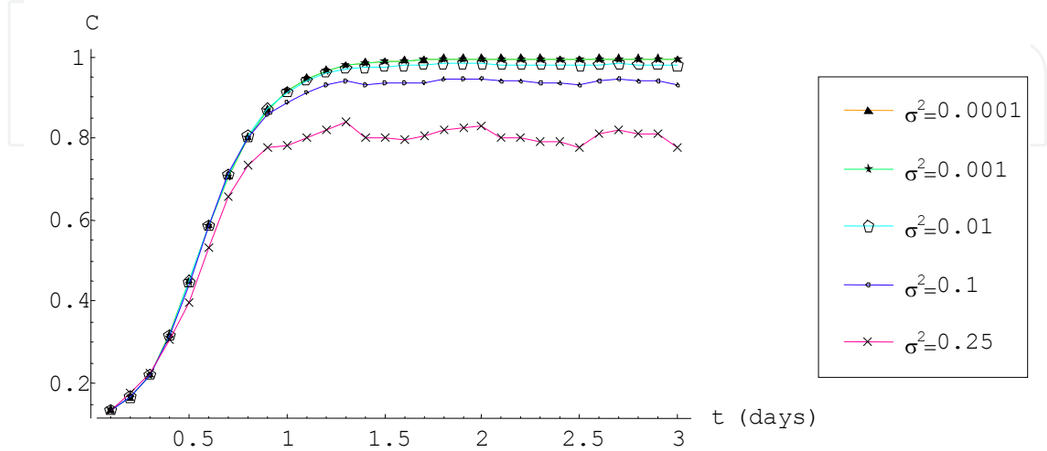


Figure 3.9. Concentration profiles at x = 0.5 m for b = 0.0001.

We increase b by 10 times to obtain Figure 3.10, which shows considerable changes in the breakthrough curves.

To understand the effects of different Wiener realizations on the concentration profiles by using 50 different Wiener realisations to calculate the 95% confidence intervals. They show that, for smaller values of parameters (for example, $\sigma^2 = 0.001$, $b = 0.01$), the variations in the concentration profile are negligible, but for larger values (for example, $\sigma^2 = 0.1$, $b = 0.1$) the fluctuation regimes increase but the solutions remain stable. Obviously, the confidence intervals widen with the larger values of the parameters.

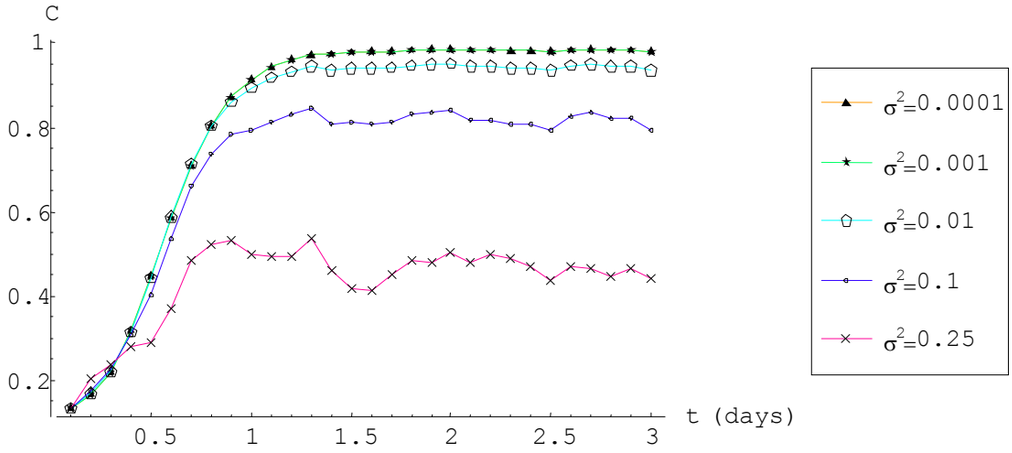


Figure 3.10. Concentration profiles at $x = 0.5$ m for $b = 0.001$.

We can now extend the investigation to a 10 m flow length. Similar to the 1 m investigation, a simple setting of the one-dimensional case is used to explore the behaviour of the model for different model parameters of the correlation length, b and the variance, σ^2. To obtain meaningful concentration profiles for the extended spatial domain, we also increase the time interval of the solution. The stochastic model is computationally solved for an average linear velocity of 0.5 m/day for 30 days, which allows sufficient time for the solute to travel the entire spatial domain. However, due to increases in the time length, a different realisation of the standard Wiener process increments was used for the 10 m length. This may raise the question of the validity in comparing of the model for two different spatial lengths. However, as we have shown earlier the model is reasonably stable for different Wiener process increments and therefore, it is reasonable to assume that the solutions of the model are not significantly affected by the different Wiener processes. A single realisation of the standard Wiener process is used throughout the investigation of the 10 m spatial domain.

Similar to 1 m case, the 10 m scale shows that a SSTM could mimic the solution of the advection – dispersion equation for the same distance (Figure 3.11)

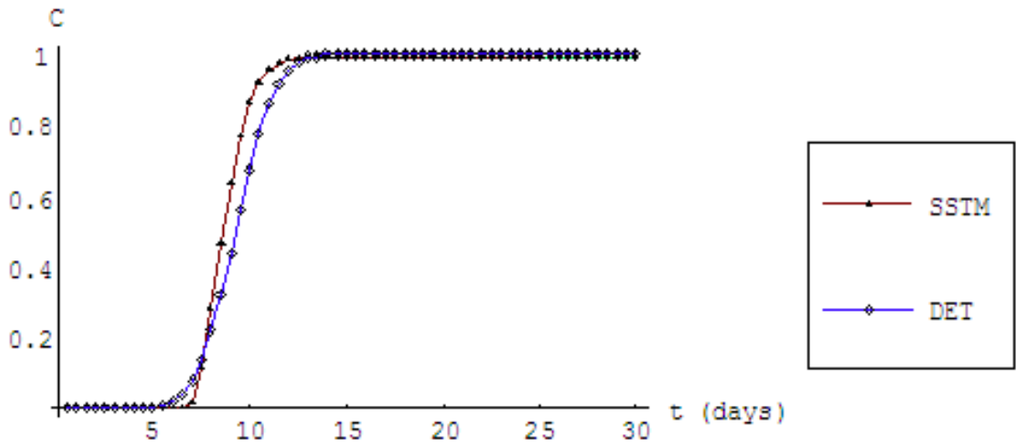


Figure 3.11. Comparison of deterministic advection-dispersion (D = 0.035 m²/day) and the SSTM (σ^2 = 0.001 and b = 0.0001) concentration profiles at x = 5m for 10 m domain.

The behaviours of the SSTM model for the 10-m flow length are quite similar to those for the 1-m flow length. The influence of the parameter b can be seen in Figure 3.12 for the fixed values of σ^2. High b values increase the propensity of the flow to be more stochastic and decrease the asymptotic values of the concentration. For a given b, the effects of increasing σ^2 are more profound in comparison to those associated with increasing b for a fixed σ^2.

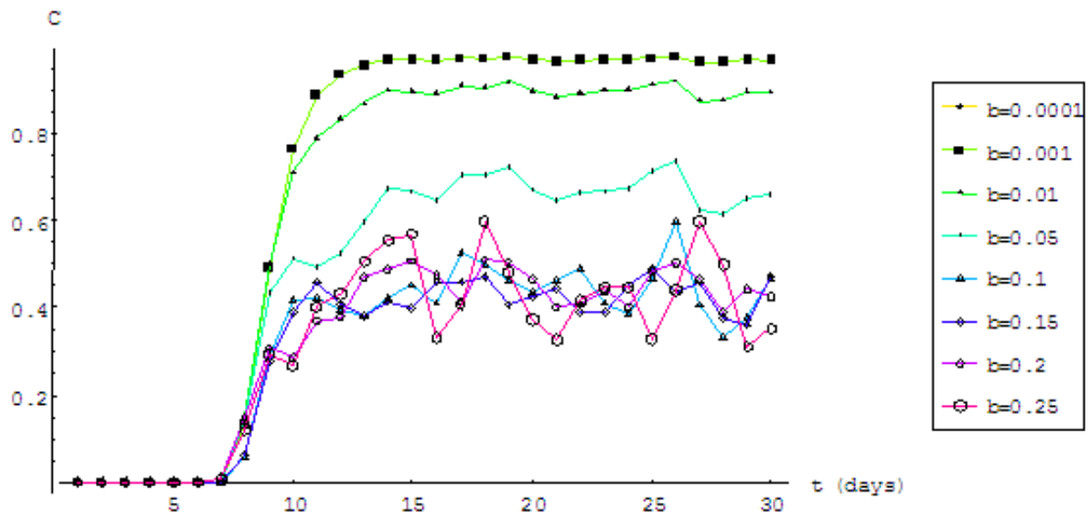


Figure 3.12. Comparison of deterministic advection-dispersion (D = 0.035 m²/day) and the SSTM (σ^2 = 0.001 and b = 0.0001) concentration profiles at x = 5m for 10 m domain.

How do we relate the parameter, σ^2 and b, to the physical porous structure? The relationship need to be understood through the influences on the concentration profiles. b is the correlation length of the velocity Kernel $q(x_1, x_2)$ (see equation 3.3.9), and higher the b slower the rate at which $q(x_1, x_2)$ decays. This means that pore structure contains larger

pores; and b is indicative of the size of pores, and, may be, geometric shapes of pores. σ^2 affects the profiles more dramatically, especially depressing asymptotic of the profiles, indicating that solute mass is dissolved in a larger volume of water. This alludes again to pore geometry (shapes and interconnecting paths) and if pore structure is heterogeneous with high porosity, one could expect σ^2 as well as b to be high. σ^2 and b allow us more flexibility of defining the nature of solute dispersion, and the complex interaction between σ^2 and b would help us to characterise the pore structures for a given velocity kernel.

3.6 A Comparison of the SSTM with the Experiments Data

The Lincoln University aquifer is 9.49 m long, 4.66 m wide and 2.6 m deep. As shown in Figure 3.13, constant head tanks are the boundaries of the aquifer at its upstream and downstream ends. A porous wall provides the hydraulic connection between the aquifer and the head tanks. A weir controls the water surface elevation in each head tank, and each weir can be adjusted to provide different hydraulic gradients. However, a uniform hydraulic gradient of 0.0018 (head loss along the aquifer / flow length = 0.017 m / 9.49 m) is maintained during the entire experiment along the longitudinal direction of the tank.

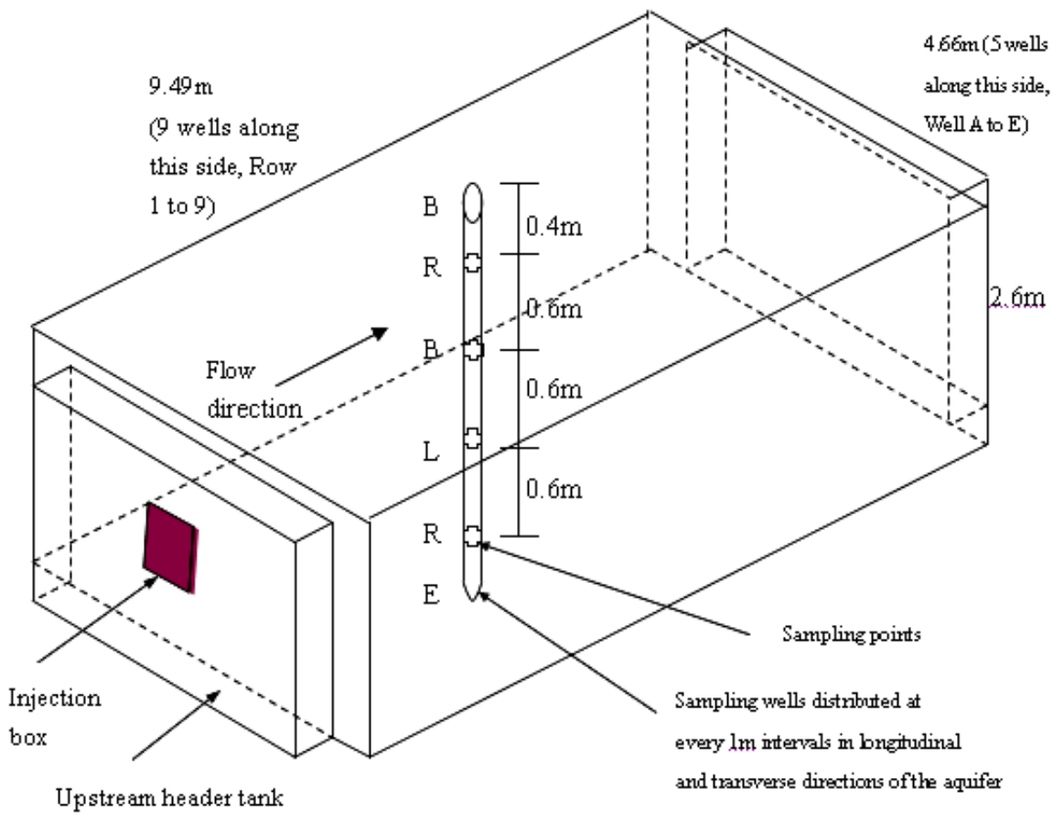


Figure 3.13. Schematic diagram of the artificial aquifer at Lincoln University, New Zealand (Courtesy of Dr.John Bright, Lincoln Ventures Ltd).

Multi-port monitoring wells are laid out on a 1m x 1m grid. The computer controlled peristaltic pumps enable fully automated, simultaneous solute water samples to be collected from sample points that are uniformly distributed throughout the aquifer (four sample points for each grid point at 0.4m, 1.0m, 1.6m and 2.2m depth from the top surface of the aquifer). The tracer used is Rhodamine WT (RWT) dye with an initial concentration of 200 parts per million and then allowed to decrease exponentially. Tracer is injected at the middle of the header tank using an injection box (dimensions of 50 cm length, 10 cm width and 20 cm depth). This tracer is rapidly mixed in the upstream header tank and, thus, infiltrates across the whole of the upstream face of the aquifer. This particular experiment described here lasted 432 hours, and two samples were taken at four-hour intervals from the wells.

Since, STTM described in this chapter is a one-dimensional model, we experiment in directly relating the one-dimension solute concentration profiles of the aquifer. However, as one can assume, the actual aquifer is subjected to transverse dispersion, and consideration of only a one-dimensional flow is not sufficiently accurate. Hence, we employ the following methodology to approximate the aquifer parameters.

Solute concentration values for the artificial aquifer are available for a large number of spatial points at different temporal intervals. The data are available mainly for header tank, row 1, row 3, row 5, row 7 and row 9 (see Figure 3.13) at all levels. Initially, we select a few spatial coordinates at row 5 of well A – level YE. We then develop a two-dimensional deterministic advection-dispersion transport model and obtain corresponding concentration values of the model that are similar to the selected spatial locations of the aquifer. As past studies show, we approximate that the transverse dispersion coefficient is 10% of the longitudinal dispersion (Fetter, 1999). The mean velocity is 0.5 m/day. The profiles of both the aquifer and the deterministic model are plotted in one axis system, to compare their similarities. This trial-and-error curve fitting technique is carried out to determine the most accurate dispersion coefficients of the deterministic model. In this procedure concentration values of the aquifer are normalised (i.e., the values vary from 0 to 1).

By trial and error, we find that the closest fit is given by the longitudinal dispersion coefficient of 0.15 m^2/day, i.e., transverse dispersion is 0.015 m^2/day, (Figure 3.14) for the aquifer.

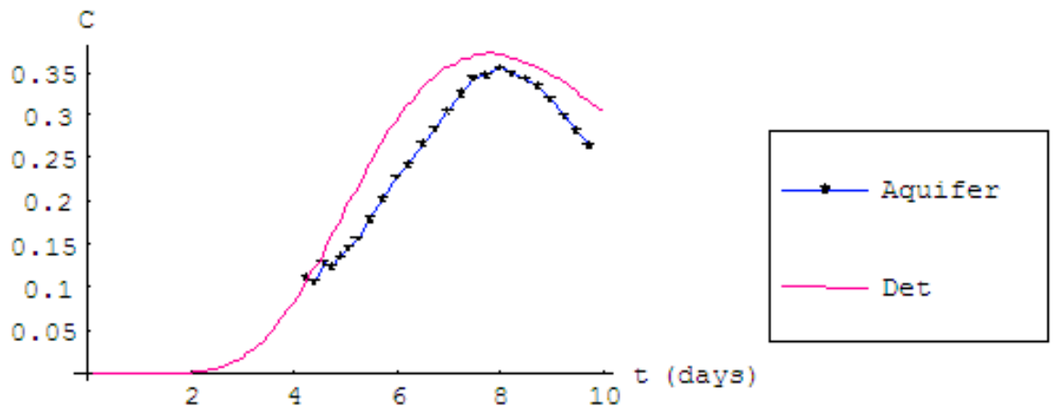


Figure 3.14. Concentration profile of trial and error curve fit for D = 0.15 m^2/day of the advection dispersion model with row 5 of the aquifer data.

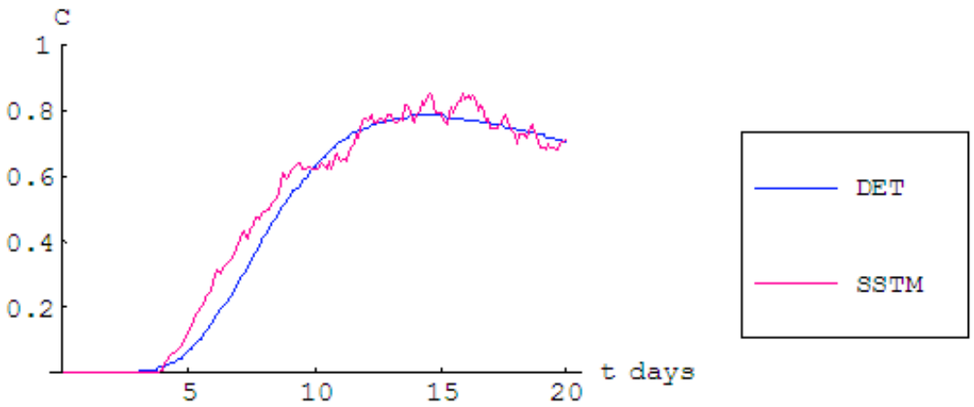


Figure 3.15. Concentration profiles of deterministic advection-dispersion model with D = 0.15 m²/day and SSTM with σ^2 = 0.01 and b = 0.01.

Subsequently, we develop a one-dimensional deterministic advection-dispersion model using the D obtained from a two-dimensional comparison. We then use a similar curve fitting technique with the 1-D deterministic model and 1-D stochastic model to find the most suitable σ^2 and b for the SSTM.

Figures 3.15 show the best fitting curves for both the SSTM and the 1-D advection-dispersion model. Having determined the appropriate parameters of the SSTM (σ^2 = 0.01 and b = 0.01) that simulate the Lincoln University aquifer at the spatial location under consideration (Row 5 – well A) we investigated the robustness of the model for different Wiener processes and found that the SSTM is reasonably stable for the seven different Wiener realisations tested.

Even though the results show, as mentioned above, that the parameter combination of σ^2 = 0.01 and b = 0.01 is a fairly accurate representation of the experimental aquifer for the given spatial point, we need to test these parameters other spatial locations. Figure 3.16 shows that the 2-D deterministic advection-dispersion model with the longitudinal dispersion coefficient of 0.15 m²/day fits the aquifer data reasonably well for the similar locations.

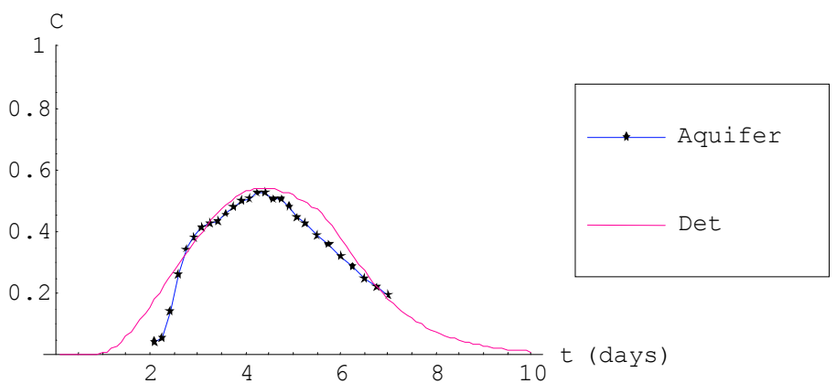


Figure 3.16. Concentration profiles of 2-D deterministic advection-dispersion model (D = 0.15 m²/day) and the experimental aquifer.

We compare the curves of the one-dimensional deterministic advection-dispersion model with D= 0.15 and of the SSTM with σ^2 = 0.01 and b = 0.01. Figure 3.17 illustrates that the curves for the deterministic model and the SSTM are in agreement.

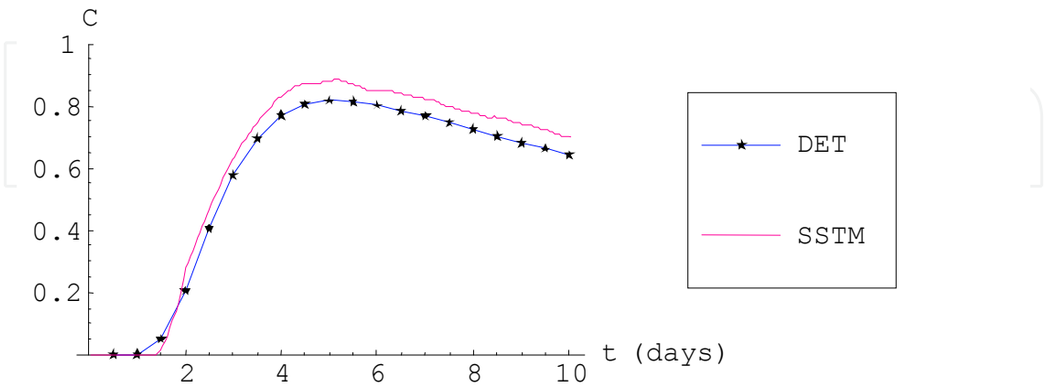


Figure 3.17. Concentration profiles of the deterministic advection-dispersion model (D = 0.15 m²/day) and SSTM with σ^2 = 0.01 and b = 0.01 for row 3 well A.

Figures 3.18 and 3.19 show that the same parameter combination is valid for the data from well A of row 7.

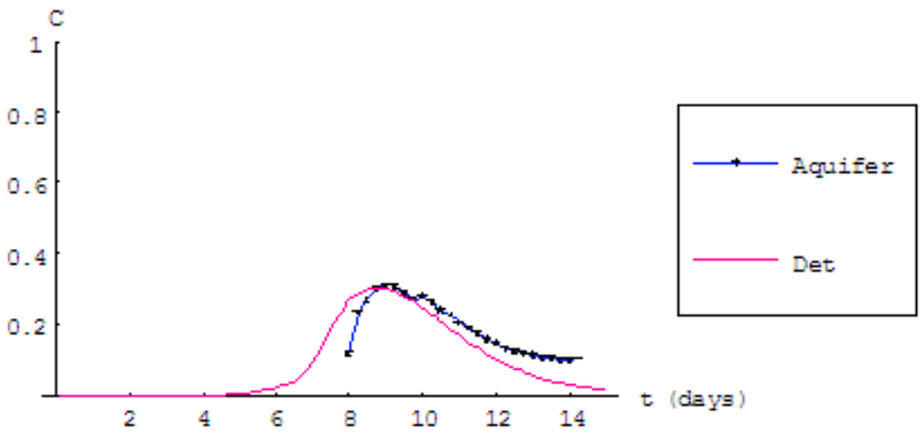


Figure 3.18. Concentration profiles of the 2-D deterministic advection-dispersion model (D = 0.15 m²/day) and the experimental aquifer at row 7.

The comparative study of the SSTM and the experimental data described here is to show that the SSTM would faithfully mimic the concentration breakthrough curves in a real experiment situation. However, we do not attempt to validate the SSTM using the trial-and-error curve fitting techniques for all the data sets as if is too labour intensive. As the experimental aquifer consists of sand of similar particle size, one would expect that one set of σ^2 and b values should be able to characterise the whole aquifer. It appears that σ^2 =0.01 and b=0.01 would be an acceptable set of parameters. However, we can not say

much about the uniqueness of these set of parameters. A correlation length (b) of 0.01 m seems to be reasonable for a porous medium containing sand. But this has not been ascertained through the experiments in a laboratory. Therefore, one of the ways of determining the uniqueness of the parameters set is to develop the methodologies to estimate them using all the observations we have. We have used the maximum likelihood estimation (MLE) to determine the parameters of the deterministic advection and dispersion model for the aquifer data. However, MLE can not be used to estimate parameters of SSTM as the SSTM contains the stochastic integral for dispersion which can not be built into likelihood functions. Therefore, we make use of ANN to estimate the parameters of the SSTM for the experimental aquifer in section 3.6.

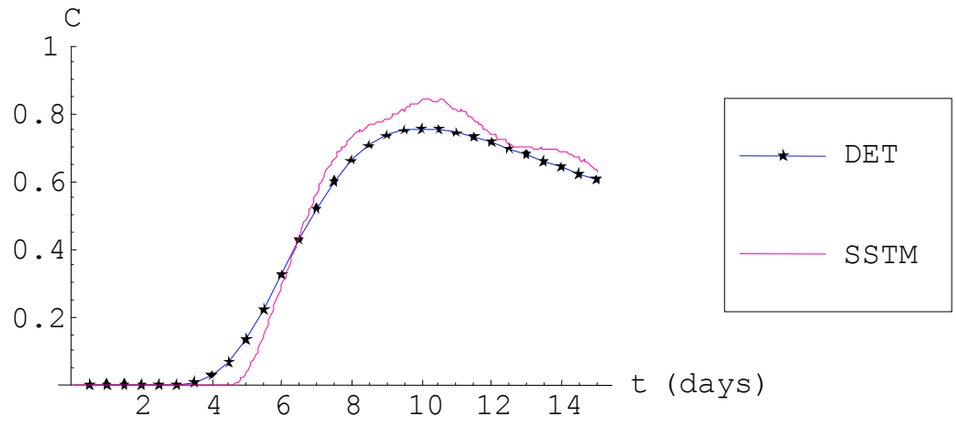


Figure 3.19. Concentration profiles of the deterministic advection-dispersion model (D = 0.15 m²/day) and SSTM with $\sigma^2 = 0.01$ and $b = 0.01$ for row 7 well A.

3.7 Parameter Estimation using Maximum Likelihood Method

In chapter 1, we discuss the parameter estimation of the partial differential equations when the additive noise terms are present. The development of the likelihood functions are based on the theory of parameter estimation for stochastic process (Kutoyants, 1984; Lipster and Shirayer, 1977; Basawa and Prakasa Rao, 1980) and we use this theory to estimate parameters.

Let us consider a one-dimensional stochastic advection-dispersion equation with additive fluctuations:

$$\frac{\partial C}{\partial t} = D_L\left(\frac{\partial^2 C}{\partial x^2}\right) - v_x\left(\frac{\partial C}{\partial x}\right) + \xi(x,t), \tag{3.7.1}$$

where D_L is the longitudinal hydrodynamic dispersion coefficient, in m²/day, and v_x is the velocity of water flow, in m/day.

The two parameters to be estimated are D_L and v_x, and we can express the right hand side of equation (3.7.1) as a linear function of the parameters (see section 1.4)

$$f(t,C,\theta) = a_0(C,\ t)+\theta_1 a_1(C,t)+\theta_2 a_2(C,\ t). \tag{3.7.2}$$

where

$$a_0(C_i,t)=0;\quad a_1(C_i,t) = \left(\frac{\partial^2 C}{\partial x^2}\right)_i;\quad a_2(C_i,t) = -\left(\frac{\partial C}{\partial x}\right)_i;\ \theta_1 = D_L;\quad \theta_2 = v_x .$$

Then the log-likelihood function can be written as,

$$\begin{aligned} l(\theta_1,\theta_2) = &\sum_{i=1}^{M}\int_0^T \{a_0(C_i,t)+\theta_1 a_1(C_i,t)+\theta_2 a_2(C_i,t)\}\, dC_i(\mathrm{t}) \\ &-\frac{1}{2}\sum_{i=1}^{M}\int_0^T \{a_0(C_i,t)+\theta_1 a_1(C_i,t)+\theta_2 a_2(C_i,t)\}^2\, dt \end{aligned} \tag{3.7.3}$$

Differentiating equation (3.7.3) with respect to θ_1 and θ_2, respectively, we obtain the following two simultaneous equations:

$$\begin{aligned} &\sum_{i=1}^{M}\int_0^T \{a_1(C_i,t)\}\, dC_i(t) - \sum_{i=1}^{M}\int_0^T \{a_0(C_i,t)+\theta_1 a_1(C_i,t)+\theta_2 a_2(C_i,t)\}\{a_1(C_i,t)\}\, dt = 0 . \\ &\sum_{i=1}^{M}\int_0^T \{a_2(C_i,t)\} dC_i(t) - \sum_{i=1}^{M}\int_0^T \{a_0(C_i,t)+\theta_1 a_1(C_i,t)+\theta_2 a_2(C_i,t)\}\ \{a_2(C_i,t)\}\, dt = 0 . \end{aligned} \tag{3.7.4}$$

Suppose we have the observations of solute concentration, C_i at M independent space coordinates along the x-axis, where $1 \le i \le M$, at different time intervals, t (where $0 \le t \le T$, and T is an integer that represents the last reading taken at unit intervals on t-axis). In other words, we have M number of C_i observations for each time step. Hence, there are, altogether, $((T+1)M)$ C_i observations. We use these observations to estimate the parameter θ_1 and θ_2,

We substitute $a_0(C_i,t)$, $a_1(C_i,t)$, $a_2(C_i,t)$, θ_1 and θ_2 in equations (3.7.4) to obtain the following set of equations,

$$\begin{aligned} &\sum_{i=1}^{M}\sum_{t=0}^{T}\left\{\frac{\partial^2 C_i}{\partial x^2}\right\}\Delta C_i - D_L\sum_{i=1}^{M}\sum_{t=0}^{T}\left\{\frac{\partial^2 C_i}{\partial x^2}\right\}^2 \Delta t - v_x\sum_{i=1}^{M}\sum_{t=0}^{T}\left\{\frac{\partial C_i}{\partial x}\right\}\left\{\frac{\partial^2 C_i}{\partial x^2}\right\}\Delta t = 0 \\ &\sum_{i=1}^{M}\sum_{t=0}^{T}\left\{\frac{\partial C_i}{\partial x}\right\}\Delta C_i - D_L\sum_{i=1}^{M}\sum_{t=0}^{T}\left\{\frac{\partial C_i}{\partial x}\right\}\left\{\frac{\partial^2 C_i}{\partial x^2}\right\}\Delta t - v_x\sum_{i=1}^{M}\sum_{t=0}^{T}\left\{\frac{\partial C_i}{\partial x}\right\}^2 \Delta t = 0. \end{aligned} \tag{3.7.5}$$

Therefore, D_L and v_x values can be obtained by solving these two simultaneous equations.

Once v_x is known, the hydraulic conductivity (K) can be obtained from,

$$v_x = \frac{K}{n_e}\left(\frac{dh}{dl}\right) \text{ = average linear velocity, m/day,} \tag{3.7.6}$$

C = solute concentration, mg/l,

K = hydraulic conductivity, m/day,

$\frac{dh}{dl}$ = hydraulic gradient, m/m, and

n_e = effective porosity.

We have used the experimental data from the artificial aquifer (Figure 3.13) to estimate D_L and v_x. Table 3.1 shows the estimated and experimental values for D_L and K (calculated from equation (3.7.6)) at different depths of the aquifer from the top surface.

Depth (m)	Hydraulic conductivity, K (m/day)		Longitudinal hydrodynamic dispersion, D_L (m²/day)	
	Estimated	Experimental	Estimated	Experimental
0.4	203.2	137	0.167	0.1596
1.0	210.6	137	0.143	0.1596
1.6	208.9	137	0.134	0.1596
2.2	262.3	137	0.242	0.1596

Table 3.1. Estimated and experimental parameters hydraulic conductivity, K (m/day) and longitudinal hydrodynamic dispersion, D_L (m²/day).

The results show the accuracy of D_L estimates is better than that of K. We estimate the v_x by using equation (3.7.6) for and for simplicity, we assume that the hydraulic gradient, $\frac{dh}{dl}$ and the effective porosity, n_e to be constants, or in other words, their spatial distributions are homogeneous. However, distribution of these values in the aquifer may be slightly heterogeneous. This assumption may have influenced the accuracy of less precise K estimates. However, the D_L values are estimated directly and not affected by such assumptions, and the maximum likelihood estimates are closer to the experimental.

Another possible phenomenon that can be present in solute transport is adsorption and the occurrence of short circuits which are, assumed to be included in the random component, $\xi(x,t)$, of the governing equation, equation (3.7.1). However, we assumed that in the experiments the tracer is mixed in the upstream header tank, hence, adsorption in the aquifer could be neglected.

As mentioned earlier, in the process of calibration to obtain parameters, it is assumed that the experimental aquifer is homogeneous. However, Figures 3.20, 3.21 and 3.22 show that the aquifer did not behave as expected (please refer to Figure 3.13 for notation). Figure 3.10 shows that the concentration values at a well which is closer to the middle of the artificial aquifer (Row 5 – Well B). This well is approximately 3 m away from the header tank as shown in Figure 3.13. Other Figures, 3.21 and 3.22, show that the concentration profiles of

wells at Well B at Row 3 and Well A at Row 7, respectively. These figures demonstrate that the concentration values are not the same at all the depths and, hence, the behaviour of the aquifer is not similar throughout. The plots of other wells also exhibit heterogeneous behaviour. Therefore, we can state that the aquifer is not behaving homogeneously, meaning that the aquifer parameters, such as hydraulic gradient and effective porosity are not uniformly distributed throughout the system. The variables used to calibrate the aquifer parameters are subjected to randomness and the accuracy of the results could be affected, considerably.

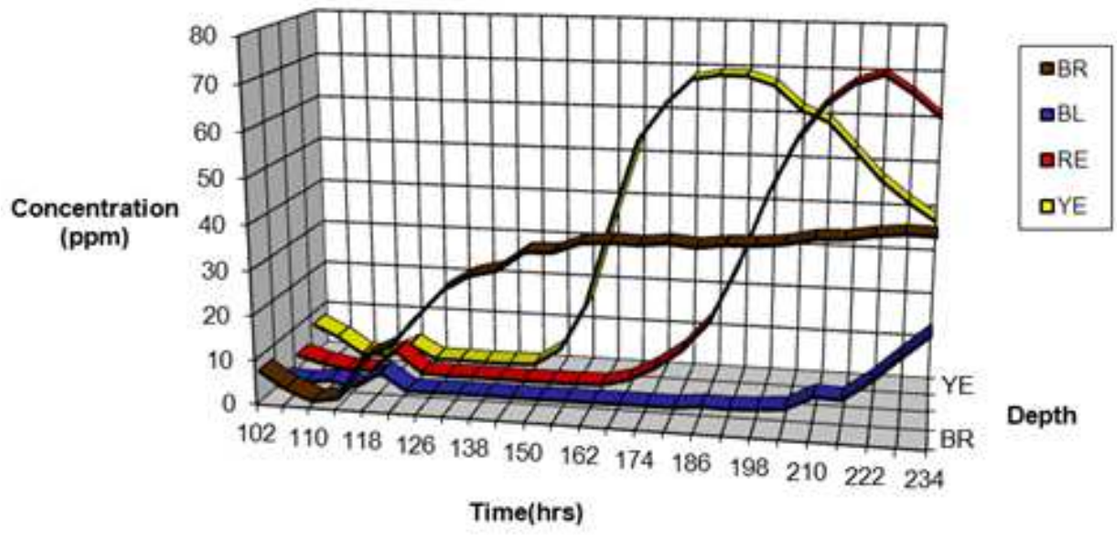


Figure 3.20. Concentration profiles at Row 5 – Well B.

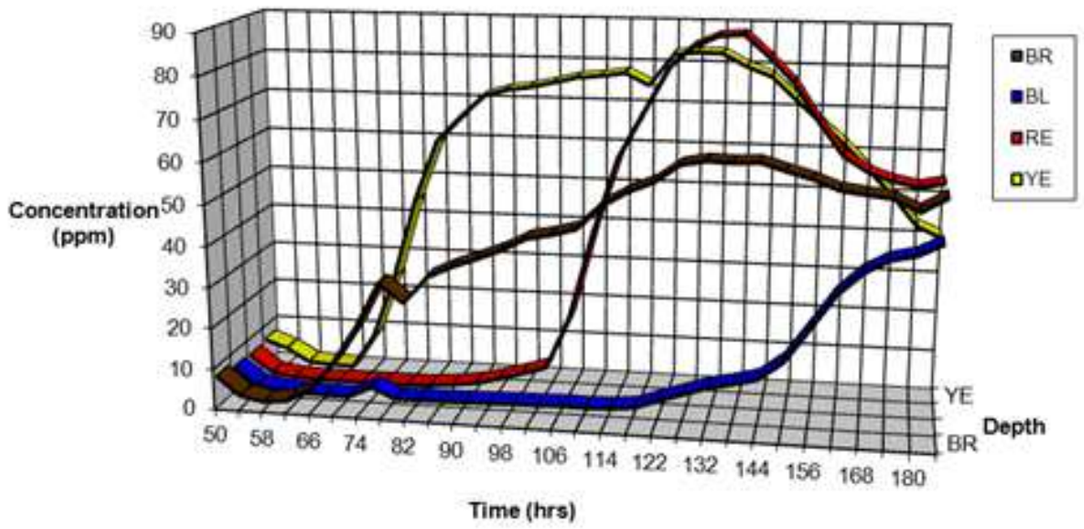


Figure 3.21. Concentrations profiles at Row 3 – Well B.

The reason that the artificial aquifer does not behave homogeneously may be due to the method of construction. The aquifer was constructed using sand blocks that were laid layer by layer. We assume that even though material used in the aquifer is uniform, joints in the blocks can create diverse flow patterns and different flow lengths. Besides, due to the high pressure on the bottom layers (from the top layers), they may be more compacted and, therefore, behave differently.

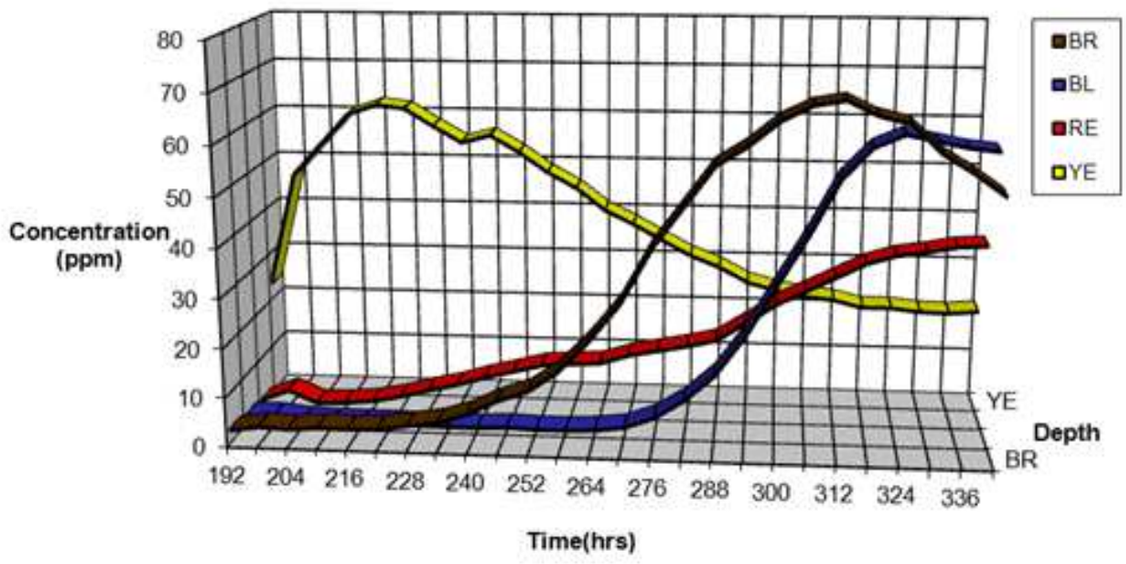


Figure 3.22. Concentrations profiles at Row 7 – Well A.

We can extend the parameter estimation procedure to determine parameters of a two-dimensional groundwater problem. In the two-dimensional case, an advancing solute front will also tend to spread in the directions normal to the direction of flow because at the pore scale the flow paths can diverge. This results in mixing in the directions normal to the flow path, which is called transverse dispersion. Considering the transverse dispersion, the two-dimensional advection-dispersion equation can be written as (Fetter, 1999),

$$\partial C = \left\{ D_L \left(\frac{\partial^2 C}{\partial x^2} \right) + D_T \left(\frac{\partial^2 C}{\partial y^2} \right) \right\} dt - v_x \left(\frac{\partial C}{\partial x} \right) dt \,, \tag{3.7.7}$$

where C = solute concentration (mg/l),

t = time (day),

D_L = hydrodynamic dispersion coefficient parallel to the principal direction of flow (longitudinal) (m^2/day),

D_T = hydrodynamic dispersion coefficient perpendicular to the principal direction of flow (transverse) (m^2/day), and

v_x = average linear velocity (m/day).

The randomness of heterogeneous groundwater systems can be accounted for by adding a noise component to equation (3.7.7), and it can be given by

$$\partial C = \left\{ D_L \left(\frac{\partial^2 C}{\partial x^2} \right) + D_T \left(\frac{\partial^2 C}{\partial y^2} \right) \right\} dt - v_x \left(\frac{\partial C}{\partial x} \right) dt + \xi(x,t)\, dt \,, \tag{3.7.8}$$

where $\xi(x,t)$ is assumed to be a zero-mean stochastic process.

We multiply equation (3.6.8) by dt throughout and replace $\xi(x,t)dt$ by $\sigma^2 dB(t)$ (Jazwinski, 1970), where σ^2 is the amplitude of the Wiener increments, $dB(t)$, to obtain equation (3.7.9). We can now obtain the stochastic partial differential equation as follows,

$$\partial C = \left\{ D_L \left(\frac{\partial^2 C}{\partial x^2} \right) + D_T \left(\frac{\partial^2 C}{\partial y^2} \right) \right\} dt - v_x \left(\frac{\partial C}{\partial x} \right) dt + \sigma^2 dB(t) \,. \tag{3.7.9}$$

As we described in equation (3.7.6) average linear velocity can be expressed by

$$v_x = \frac{K}{n_e} \left(\frac{dh}{dl} \right),$$

Hence, equation (3.7.9) becomes,

$$\partial C = \left\{ D_L \left(\frac{\partial^2 C}{\partial x^2} \right) + D_T \left(\frac{\partial^2 C}{\partial y^2} \right) \right\} dt - \frac{K}{n_e} \left(\frac{dh}{dl} \right) \left(\frac{\partial C}{\partial x} \right) dt + \sigma^2 dB(t) \,. \tag{3.7.10}$$

We assume that transverse dispersion (D_T) can be approximated to 10% of the longitudinal dispersion D_L , i.e., D_T = 0.1 D_L (Felter,1999).

Then equation (3.7.10) becomes,

$$\partial C = D_L \left\{ \left(\frac{\partial^2 C}{\partial x^2} \right) + 0.1 \left(\frac{\partial^2 C}{\partial y^2} \right) \right\} dt - \frac{K}{n_e} \left(\frac{dh}{dl} \right) \left(\frac{\partial C}{\partial x} \right) dt + \sigma^2 dB(t) \,. \tag{3.7.11}$$

Equation (3.7.11) can be written in the following form:

$$f(t, C, \theta) = a_0(C,t) + \theta_1 a_1(C,t) + \theta_2 a_2(C,t),$$

where,

$$a_0(C_i,\ t) = 0; \quad a_1(C,t) = \left\{ \frac{\partial^2 C}{\partial x^2} + 0.1 \left(\frac{\partial^2 C}{\partial y^2} \right) \right\}; \quad a_2(C_i,t) = -\left(\frac{\partial C}{\partial x} \right)_i ;$$

$$\theta_1 = D_L; \quad \theta_2 = v_x = \frac{K}{n_e} \left(\frac{dh}{dl} \right).$$

We use the observations for $C(x, y, t)$ at M discrete points in (x, y) coordinate space for a period of time t (where $0 \le t \le T$). Then we obtain the estimates for two unknown parameters as the solution to the following simultaneous equations:

We can simplify equation (3.7.11) to

$$\sum_{i=1}^{M}\int_{0}^{T}\left\{\frac{\partial^2 C}{\partial x^2}+0.1\left(\frac{\partial^2 C}{\partial y^2}\right)\right\}dC_i$$

$$-\sum_{i=1}^{M}\int_{0}^{T}\left\{D_L\left\{\frac{\partial^2 C}{\partial x^2}+0.1\left(\frac{\partial^2 C}{\partial y^2}\right)\right\}+\frac{K}{n_e}\left(\frac{dh}{dl}\right)\left(\frac{\partial C}{\partial x}\right)\right\}\left\{\frac{\partial^2 C}{\partial x^2}+0.1\left(\frac{\partial^2 C}{\partial y^2}\right)\right\}dt=0$$

$$\sum_{i=1}^{M}\int_{0}^{T}\left\{\frac{\partial C}{\partial x}\right\}dC_i-\sum_{i=1}^{M}\int_{0}^{T}\left\{D_L\left\{\frac{\partial^2 C}{\partial x^2}+0.1\left(\frac{\partial^2 C}{\partial y^2}\right)\right\}+\frac{K}{n_e}\left(\frac{dh}{dl}\right)\left(\frac{\partial C}{\partial x}\right)\right\}\left\{\frac{\partial C}{\partial x}\right\}dt=0. \qquad (3.7.12)$$

We can rewrite equations (3.7.12) as

$$\begin{aligned} D_L k_1+\frac{K}{n_e}\left(\frac{dh}{dl}\right)l_1+m_1=0,\\ D_L k_2+\frac{K}{n_e}\left(\frac{dh}{dl}\right)l_2+m_2=0, \end{aligned} \qquad (3.7.13)$$

where

$$k_1=-\sum_{i=1}^{M}\sum_{j=0}^{n-1}\left\{\frac{\partial^2 C_i^j}{\partial x^2}+0.1\,\frac{\partial^2 C_i^j}{\partial y^2}\right\}^2\left(t_{j+1}-t_j\right),$$

$$l_1=-\sum_{i=1}^{M}\sum_{j=0}^{n-1}\left\{\frac{\partial C_i^j}{\partial x}\right\}\left\{\frac{\partial^2 C_i^j}{\partial x^2}+0.1\,\frac{\partial^2 C_i^j}{\partial y^2}\right\}\left(t_{j+1}-t_j\right),$$

$$m_1=\sum_{i=1}^{M}\sum_{j=0}^{n-1}\left\{\frac{\partial^2 C_i^j}{\partial x^2}+0.1\,\frac{\partial^2 C_i^j}{\partial y^2}\right\}\left(C_i^{j+1}-C_i^j\right),$$

$$k_2=-\sum_{i=1}^{M}\sum_{j=0}^{n-1}\left\{\frac{\partial C_i^j}{\partial x}\right\}\left\{\frac{\partial^2 C_i^j}{\partial x^2}+\frac{\partial^2 C_i^j}{\partial y^2}\right\}\left(t_{j+1}-t_j\right),$$

$$l_2=-\sum_{i=1}^{M}\sum_{j=0}^{n-1}\left\{\frac{\partial C_i^j}{\partial x}\right\}^2\left(t_{j+1}-t_j\right), \text{ and}$$

$$m_2=\sum_{i=1}^{M}\sum_{j=0}^{n-1}\left\{\frac{\partial C_i^j}{\partial x}\right\}\left(C_i^{j+1}-C_i^j\right).$$

We can develop the finite difference approximations for $\frac{\partial C_i}{\partial x_i}$, $\frac{\partial^2 C_i}{\partial x_i^2}$ and $\frac{\partial^2 C_i}{\partial y_i^2}$, to obtain values of k_1, l_1, m_1, k_2, l_2, and m_2. Furthermore, by substituting average values for the hydraulic gradient, $\left(\frac{dh}{dl}\right)$, and the effective porosity, n_e, two simultaneous equations (3.7.13) can be solved to obtain D_L and K, for a two-dimensional groundwater system.

We estimate the parameters of the artificial aquifer at Lincoln University using the procedure developed. The observations of solute concentration are obtained at 1 m grid intervals at four different levels (0.4, 1.0, 1.6, and 2.2 m from the surface of the aquifer). Hence, we rearrange the dataset to be four two-dimensional datasets for each level.

Table 3.2 shows the estimated parameters for the two-dimensional aquifer and the estimates have come closer to the experimental values. However, as the experimental values may not represent the real values within the aquifer, further computational experiments by changing the lateral dispersion would not improve the estimates.

Depth (m)	Hydraulic conductivity, K (m/day)		Longitudinal hydrodynamic dispersion, D_L (m²/day)	
	Estimated	Experimental	Estimated	Experimental
0.4	198.2	137	0.165	0.1596
1.0	166.7	137	0.162	0.1596
1.6	171.5	137	0.143	0.1596
2.2	231.1	137	0.197	0.1596

Table 3.2. Estimated and experimental parameters, hydraulic conductivity, K (m/day), and longitudinal hydrodynamic dispersion, D_L (m²/day), for the aquifer.

3.8 Parameter Estimation using ANN

In estimating parameters with ANN, first we employ a deterministic 2-D advection-dispersion transport numerical model to generate synthetic data. Afterwards, ANN are trained to learn the complex excitation and response relationship of the generated data. This is done by training the network sufficiently to minimise the error between the actual and network response while retaining the generalising capabilities of the network. We then estimate the associated parameters using noisy concentration data that represent real world aquifer systems. We also test the ability of the model to estimate hydraulic conductivity of an artificial experimental aquifer.

The two-dimensional deterministic advection–dispersion equation (Fetter, 1999), is used as the governing equation for this section. It is important to mention that other possible phenomenon that can be present in the solute transport, such as adsorption and the occurrence of short circuits, are neglected in the governing equation on the assumption that the introduction of noise into the solute concentration values used to estimate the parameters would compensate for them.

The deterministic solute concentration values are generated for a 10 m x 5 m 2-D aquifer using equation (3.6.7). Eight hundred data examples (patterns) are generated for different hydraulic conductivity, *K,* values that ranged from 40 to 240 m/day. It is assumed that all other parameters, control variables and subsidiary conditions are fixed. An initial concentration value of 100 ppm is considered as a point source at the middle of the header boundary of the aquifer and the same source is maintained at the boundary throughout the 10 day period considered. Exponentially distributed concentration values of the point source (at the middle of the header boundary) are considered along with the longitudinal and lateral directions as the initial conditions for the other spatial coordinates. We gather 50 input values for each example. These input values represent solute concentration values at 10 spatial locations (Figure 6.1) at five different time intervals; t = 1, t = 3, t = 5, t = 7, t = 10 day. We examine the possibility of amalgamating the time as an independent variable into the concentration input data. However, it is difficult to meaningfully integrate them into presently available ANN architectures and innovative model structures need to be developed.

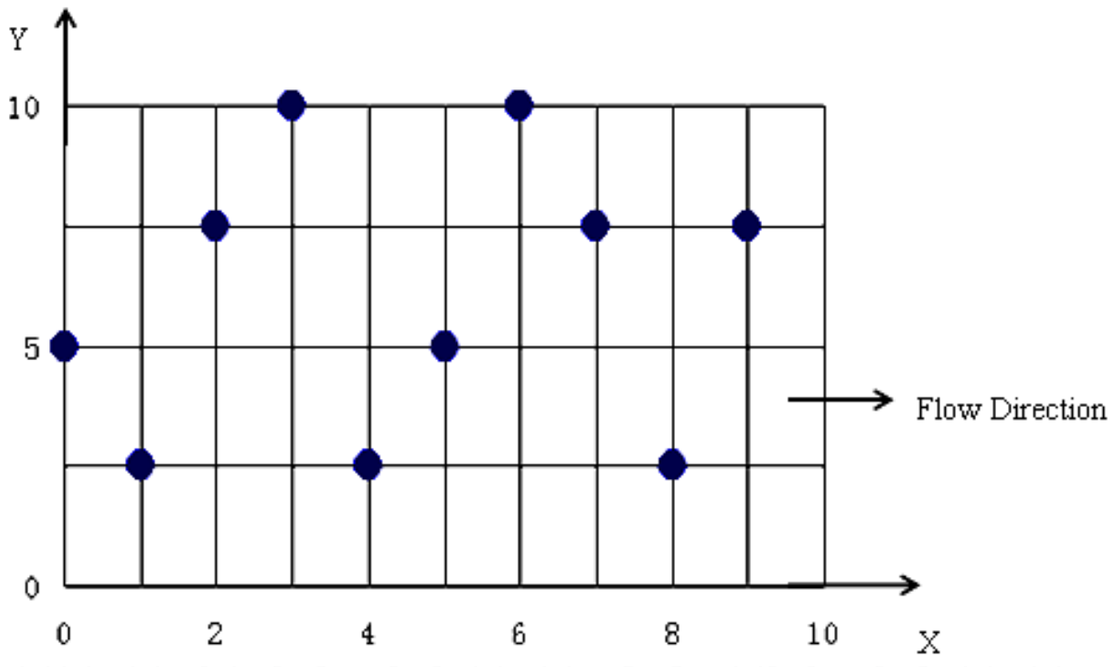


Figure 3.23. Spatial coordinates of concentration observations of the 2-D aquifer.

It is noted, in many studies, that determination of an appropriate network architecture is one of the most important but also one of the most difficult tasks in the model building process (Maier and Dandy, 2000). The network architecture determines the number of connection weights (free parameters) and the way information flows through the network. It is common practice to fix the number of hidden layers in the network and then to choose the number of nodes in each of these layers. Initially, it may be worthwhile to consider a network with simple properties. Such networks have higher processing speeds and can be implemented on hardware more economically (Towell et al., 1991; Bebis and Georgiopoulos,

1994; Castellano et al., 1997). Cybenko (1989) showed that only one hidden layer was required to approximate any continuous function, given that sufficient degrees of freedom (i.e. connection weights) are provided. However, in practice many functions are difficult to approximate with one hidden layer, requiring a prohibitive number of hidden layer nodes (Flood and Kartam, 1994). The use of more than one hidden layer provides greater flexibility and enables the approximation of complex functions with fewer connection weights in many situations (Sarle, 1994).

Despite numerous studies, no systematic approach has been developed for the selection of an optimal network architecture and its geometry. Hence, we employ a trial and error exercise to determine the network architecture (number of hidden layers and number of nodes for each layer). Initially, a simple three layer MLP network (only one hidden layer) is used to train the network to build the complex relationship of output, *K*, and the associated concentration values. Hecht-Nielsen (1987) suggested that an optimum number of hidden nodes for a single hidden layer network can be selected from following relationship, Number if hidden nodes = 2 (Number of input nodes) + 1.

This suggests we should have 101 hidden nodes for 50 inputs. Nevertheless, after a number of trial and error tests, it was found that the optimum results can be achieved by 20 hidden neurons. As reported elsewhere, Abrahart et al. (1999) presented a method based on genetic algorithms to identify the best number of suitable hidden nodes. Chakilam (1998) used principal component analysis to determine the optimal structure of the multi-layered feed forward neural network for a time series forecasting problems, thus reducing the generalisation error and overcoming the over fitting problems.

The dataset is divided into two categories with a random selection of 80% used for training and the rest for testing. The maximum and minimum values of the training network are set by selecting the values from both training (and testing) and estimating dataset, to prevent the ANN from extrapolating beyond its range. We apply scale functions of none, logistic and logistic for input, hidden and output layers, respectively. The default network parameters of NeuroShell2 (neural computing software package) are used; learning rate = 0.1, momentum = 0.1, initial weight = 0.3. The network reached the stopping criterion of average error on test set, fixed at 0.000002, in less than two minutes in a 1GHz personal computer with performance measurements of the coefficient of multiple determination, R^2 = 0.9999 and the square of the correlation coefficient, r^2 = 0.9999. The network that produces the best results on the test set is the one most capable of generalising, so this is saved as the best network.

Having completed the successful training, another dataset is employed to test the performance of the trained network. We make use of the same model to generate 800 new data values, however, the initial concentration is randomly changed by up to ±5%, and up to ±5% noise is arbitrarily added to all concentration input values. The reason for adding the noise is to simulate the real world problem of erratic behaviour of aquifers. The estimation error of each *K* value is given in Figure 3.4.4, which shows that the error increases with *K*.

Table 3.3 illustrates that the statistical measurements of error with mean square error (MSE) of 45.25, an average absolute percentage error (AAPE) of 5.63%, and a maximum error of 22.45 m/day. Such high error values may not be acceptable in the most practical cases. Since the objective range of parameters is fairly large (40 –240 m/day), the accuracy of the approximation tends to decrease (Figure 3.24). Therefore, we conduct the same estimation procedure with four

smaller permissible parameter regimes of *K*; (i) 40-90, (ii) 90-140, (iii) 140-190, and (iv) 190-240 m/day. Table 3.3 shows that accuracy of the estimates has improved considerably.

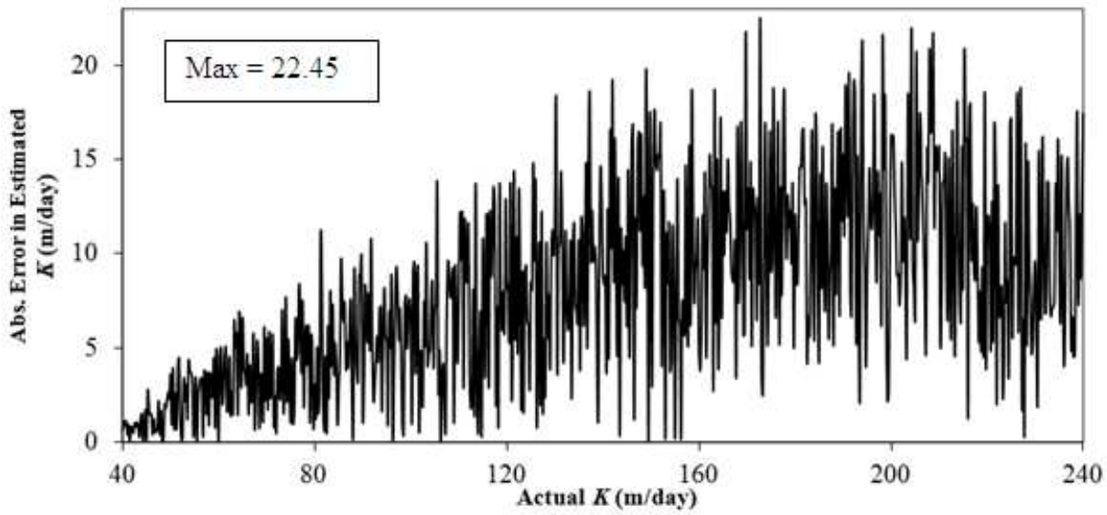


Figure 3.24. Absolute error of estimated parameter, *K*, when considering the whole range, 40–240 m/day.

The maximum error in the 190-240 range has been reduced by about 90% (Figure 3.25 and Table 3.3). Therefore, it is reasonable to assume that if we can gather prior information about the system under consideration, it is possible to obtain more accurate estimates. However, in the real world problems the prior knowledge of the system is limited. Later, we discuss a method to identify the range of parameters by using Self-Organising Maps (SOM). However, before using SOM, we explore the robustness of the ANN estimation models.

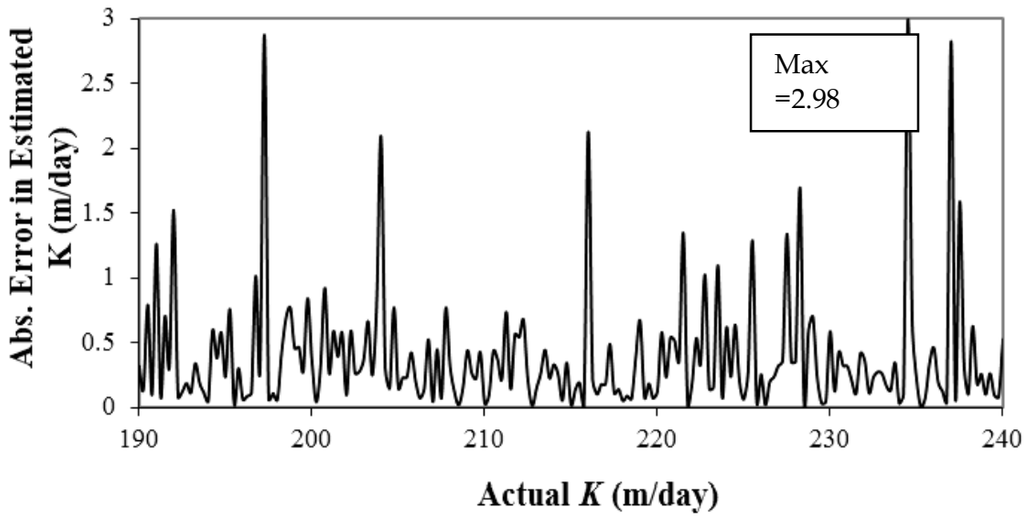


Figure 3.25. Absolute Error of estimated parameter *K*, only for the range 190 – 240 m/day.

Error	*K* Range (m/day)				
	40-240	40–90	90–140	140–190	190-240
Max	22.45	1.88	2.23	2.99	2.98
Mean	8.03	0.27	0.38	0.36	0.39
StDev	5.15	0.32	0.41	0.49	0.47
MSE	45.25	0.11	0.14	0.20	0.19
AAPE(%)	5.63	0.11	0.12	0.18	0.18

Table 3.3. Statistics of estimated error for different ranges of *K* with up to a ±5% difference in initial value and up to a ±5% noise in observations.

Initial condition		Error of *K* (m/day)				
Point C value	Noise	Maximum	Mean	Stdev	MSE	AAPE
50	-50%	9.72	3.48	1.68	12.66	0.95
60	-40%	7.55	2.06	1.57	3.36	0.94
70	-30%	6.18	2.01	1.42	3.01	0.92
80	-20%	4.36	1.59	0.97	2.58	0.77
90	-10%	2.84	0.97	0.72	1.46	0.43
Trained value (100)	0%	1.64	0.34	0.44	0.17	0.17
110	+10%	2.92	1.04	0.86	1.46	0.44
120	+20%	4.57	1.68	1.05	2.95	0.84
130	+30%	6.49	2.11	1.48	3.26	1.06
140	+40%	7.58	2.08	1.57	3.47	1.07
150	+50%	10.14	3.67	1.73	13.81	1.07

Table 3.4. Statistics of estimated error for different initial values with up to a ±5% error for the *K* range, 190-240.

As discussed in chapter 1, real world aquifer systems are subject to numerous random effects. One of them may be an initial value problem. First, we investigate in the range of *K* between 190 – 240 m/day for the stability of the model for different initial values. The point source value of the initial concentration are changed from –50% to 50%. Since, as explained above, an exponentially distributed pattern was used to determine the initial concentration values of other spatial points, change of initial value at the point source also resulted in changing the initial values of every spatial location. Furthermore, to illustrate the heterogeneity of the aquifers, up to ±5% extra noise is added to the generated concentration

values of all time intervals. Table 3.4 shows the statistics of the estimates. The estimates exhibit a direct relationship to the noise; however, the most of the results are dependable even at higher noise levels.

Random boundary conditions and an irregular porous structure can result in an erratic distribution of flow paths. Therefore, solute concentration spreads could be highly stochastic. We address this issue by extending the investigation of the robustness by adding different level of randomness to the concentration values. First, the data is generated using the deterministic solutions of equation (3.6.7) for each case and then noise is added randomly to each deterministic concentration value to generate a noisy dataset. For example, to generate up to a ±10% noise component to a deterministic value, d, two random functions are used as follows,

random function 1 → generate a random number between 0-1 (say n)

random function 2 → generate either + or -.

Therefore, noisy data = $d\,(1 \pm 10\% * n)$.

Table 3.5 demonstrates the statistics of the estimates obtained for noisy concentration data. Estimates show that the ANN model is stable even for highly stochastic systems.

± %	Error of Estimate K (m/day)				
added noise	Max	Mean	StDev	MSE	AAPE
10	2.29	0.85	1.10	1.98	0.68
20	3.54	1.05	1.19	2.04	0.76
30	5.46	1.74	1.22	2.25	0.81
40	5.88	1.96	1.35	2.31	0.92
50	6.00	1.99	1.39	2.39	0.93

Table 3.5. Statistics of estimates for noisy data for the K range, 190-240 m/day.

The ANN model gives more accurate estimates when the target parameter range is small. However, in real world heterogeneous aquifers, it may not be a trivial task to identify the accurate parameter range without reliable prior information. Self Organising Maps (SOM) has the ability to sort items into categories of similar objects by nonlinearly projecting the data onto a lower dimensional display and by clustering the data (Kohonen, 1990). We use this power of SOM and develop a method to identify the parameter range for given solute concentration values. We employ SOM to cluster an 800 x 50 dimension noisy dataset used before with a parameter range of 40 –240 m/day into four different categories. The "Supervised Kohonen" network architecture of NeroShell2 successfully categorised four different groups with 201, 200, 197 and 202 data patterns in each cluster, respectively. The SOM put data into categories with a high accuracy, with few exceptions which can be expected with noisy data, at the boundaries of the parameter ranges. To test the accuracy of the prediction capability of the trained model, we then create and feed 12 different test

datasets with the same number of input variables (50) into the trained SOM and it accurately identified the correct parameter range for all the datasets.

We extend the hybrid methodology to solve the groundwater inverse problem in the case of two unknown system parameters. We simulate the same aquifer as used above. The two parameters to be estimated are hydraulic conductivity, K (m/day) and longitudinal dispersion coefficient, D_L (m²/day). In line with earlier work, we fed 50 concentration values and two actual outputs (K and D_L) to train the network.

We use a simple three layer network and produce $R^2 = 0.9999$ and $r^2 = 0.9999$ for both outputs in 2 min and 50 sec. We then feed a different dataset, which has not been seen by the trained network before. The new dataset consisted of randomly varying (up to ±5%) initial conditions and added noise to replicate a natural system. We explore two different levels of noise; up to ±5% and ±50%. The parameter ranges are; K between 190 – 240 m/day, D_L between 0.03 – 0.08 m²/day. The ANN model produces reasonable estimates for both parameters and the summary of estimates is given in Table 3.6.

Parameter	Actual Range	± % noise	Error of Estimate			
			Max	Mean	MSE	AAPE
K	190-240	5	2.48	0.99	2.65	0.81
		50	6.78	2.35	3.18	1.12
D_L	0.03-0.08	5	0.00341	0.00092	0.0014	0.0005
		50	0.00875	0.00247	0.0029	0.0010

Table 3.6. Statistics of estimates for two parameter case

We apply the hybrid ANN inverse approach presented in the above sections to estimate parameters of the artificial aquifer at Lincoln University. Although, initial conditions, other parameters and the subsidiary conditions of the aquifer are somewhat known, we have to conduct a fairly tiresome, "trial and error" exercise to replicate the aquifer. Eight hundred data patterns are generated for the hydraulic conductivity range of 80 to 280 m/day. Each pattern consists of 100 concentration input variables for 10 distinct spatial locations for 10 different time intervals. We then use Kohonen's SOM (80% data for training and 20% for testing) to classify the input values into four clusters. Next we feed the actual aquifer data into the trained network and the selected subrange; it is established that the aquifer parameter should be within the second cluster (130 – 180 m/day). Based on this information, we generate a separate dataset for the specified range and train an MLP network with the associated K values.

The estimate of K given by the trained ANN is 152.86 m/day. The experimental value of hydraulic conductivity, K, is found to be 137 m/day, which is calculated by calibration tests conducted by aquifer testing staff. In these experiments they have assumed that the aquifer is homogeneous. The difference between two estimates is only 10.37 %. Considering the assumptions of homogeneity made by the aquifer researchers and other possible human

errors, it is fair to state that the estimate obtained from the ANN model is reasonable and acceptable.

Our investigation emphasise that the importance of modelling a sufficiently true representation of the physical system and subsidiary conditions to obtain accurate parameter values. As Minns et al. (1996) pointed out, the ANN is susceptible to becoming "a prisoner of its training data". Therefore, prior information, such as type of contaminant source, boundary conditions and subsidiary conditions is crucial to modelling the system accurately. If we could gain such prior information and model the system with ANN, it would be capable of solving the inverse problem with greater accuracy, even with highly noisy data, as well as different system input values.

3.9 Parameter Estimation of SSTM

In this section, we apply the ANN hybrid approach presented in the previous section to estimate parameters of the Stochastic Solute Transport Model (SSTM) developed in this chapter. Since SSTM consists of two parameters ; variance (σ^2) and correlation length (b), we estimate both parameters simultaneously.

In section 3.8 we showed that the accuracy of the estimates was inversely proportional to the size of the objective range of the output (maximum error of the estimate of parameter K was reduced by about 90% for smaller output ranges). In addition, the accuracy of the estimates may reduce when two parameters are estimated simultaneously. To limit the diminution of accuracy of the results imposed by the above mentioned performance characteristics of the method and, as this research is conducted in a general personal computer, we select a smaller permissible output ranges. Additionally, it has shown that the higher parameter values of SSTM (especially σ^2 around 0.25) represent greater heterogeneous flow systems. Thus, we limit the parameter range for both parameters, σ^2 and b, to be between 0.0001 and 0.2.

The main objective of this section is to estimate parameters of SSTM using a hybrid ANN approach developed in the above section. We then extend the exercise using a case study to validate the method. As used in the previous section, the results from the artificial aquifer is used for the validation. However, we intend to achieve an auxiliary objective in the validation process by comparing the artificial aquifer parameters with the estimates obtained for the same aquifer using a curve fitting technique.

As the first step of the implementation process, we use SSTM to simulate a one-dimensional aquifer of 10 m in length. Eight hundred data patterns for different combinations of σ^2 and b are generated. Each parameter ranges from 0.0001 and 0.2. Every data pattern consists of 200 inputs for 10 various spatial locations of the aquifer for 20 distinct time intervals. The same standard Wiener process is used for generating all data sets. The initial condition of the concentration value of unit 1.0 at x = 0.0 and exponentially distributed values for other spatial locations are considered. Throughout the simulation the same concentration (unit 1.0) is maintained at the upper end boundary. It is assumed that the mean velocity of the solute is 0.5 m/day. As mentioned above, it is very important to limit the objective range of the parameters for smaller regimes to attain accurate approximations. Thus, Kohonen's Self Organising Map (SOM) is employed to cluster the data set into four categories.

Since the data set represents the stochastic behaviour of the flow, the time needed to classify the data into separate groups is much more than for the similar case, in the deterministic advection – dispersion data, in the previous section. Randomly selected 80% of data were used for training the network and the rest for the validation. Notwithstanding the random nature of the current dataset, the SOM has clustered the data to an adequate degree of accuracy that may be sufficient for the problem at hand.

To test the performance of the trained network model, we use eight different data patterns. Two datasets are generated for each group of clustered network. Moreover, each dataset is produced using different standard Wiener process. New data patterns are fed into the trained model and it accurately identified the parameter range it fitted into.

Having substantiated that the SOM could be successfully used to cluster large dataset that represent heterogeneous data, such as given by SSTM, we create four separate specialised network models for smaller parameter regimes of both parameters in the following ranges;

$$0.0001 – 0.05; 0.05 – 0.1; 0.1 – 0.15; \text{ and } 0.15 – 0.2.$$

Each dataset comprises 441 data patterns of different combinations of σ^2 and b at intervals of 0.0025. For instance, for the range of 0.05 – 0.1, there are 21 different values of each parameter; 0.05, 0.0525, 0.055, 0.0575, … , 0.1 (21 x 21 = 441 data patterns). Same SSTM model used above for SOM cluster distribution is also used for data generation. Nevertheless, in this case each data pattern contains not only 200 inputs but also two corresponding output parameters. The number of training patterns has considerable influence on the performance of the ANN model (Flood and Kartam, 1994). Increasing the number of data patterns provides more information about the shape of the solution surface and, thus, improves the accuracy of the model prediction. However, in most real world applications, numerous logistical issues impose limitations on the amount of data available and, consequently, the size of the training set. Hence, in developing a method for practical applications, it is important to test the robustness of the method for such data limitations. For that reason, we limit the parameter range of each model to 21 values (between 0.0001 and 0.2 at intervals of 0.0025).

Maier and Dandy (2000) showed that it was important to select a suitable network architecture and model validation method in the development of ANN models to achieve optimum results. In addition, it may be necessary to select the most suitable model for handling highly random data such as SSTM data. Therefore, we conduct a few trial and error exercises to choose the appropriate model structure, and training and testing procedure. The following are the MLP ANN models that are considered for different combinations of hidden layers and various grouping of activation functions: three layer standard connections; four layer standard connections; five layer standard connections; one hidden layer with two parallel slabs with different activation functions; one hidden layer with three hidden slabs with different activation functions; one hidden layer of two parallel slabs, different activation functions and jump connection; three layer jump connections; four layer jump connections; and five layer jump connections.

After numerous attempts, it is found that the network model of five layer standard connections could produce the best trained model in the least time. In the selected model, each hidden layer consists of 30 neurons. Activation functions of linear <0, 1>, logistic, tanh,

Gaussian and logistic are used for layers of input, hidden (3 layers) and output, respectively. The default network parameters (NeuroShell2) are employed; learning rate = 0.1, momentum = 0.1, initial weight = 0.3. The stopping criterion is set to a minimum error of 0.000001. All four networks reach the stopping condition in about 30 minutes in a 1 GHz personal computer with the performance measurements shown in Table 3.7.

After the completion of successful training for each model, separate datasets are generated to test the prediction capability of each model. The same SSTM is employed to produce another dataset of 441 data patterns for each parameter range. However, different standard Wiener process increments are used. In addition, initial and boundary conditions are adjusted up to ±5% by adding random values. Input data values of each dataset are then fed into the corresponding trained network and processed to obtain model predictions.

Parameter range	Coefficient of multiple determination, R^2		Square of the correlation coefficient, r^2		Mean absolute error	
	σ^2	b	σ^2	b	σ^2	b
0.0001 – 0.05	0.9912	0.9876	0. 9911	0.9876	0.0	0.0
0.05 – 0.1	0.9911	0.9876	0. 9899	0.9876	0.0	0.0
0.1 – 0.15	0.9898	0.9870	0. 9872	0.9868	0.0	0.0
0.15 – 0.2	0.9728	0.9774	0. 9721	0.9661	0.0001	0.0001

Table 3.7. Performance measurements of trained ANN model for four different parameter ranges.

Figure 3.26 illustrates the absolute error of estimated parameter σ^2, for the range of 0.0001 – 0.05. It shows that the ANN model prediction is extremely satisfactory and that the average absolute error is approximately 0.04%. Figure 3.27 shows that prediction for the other parameter, *b* also met with similar accuracy for the same range.

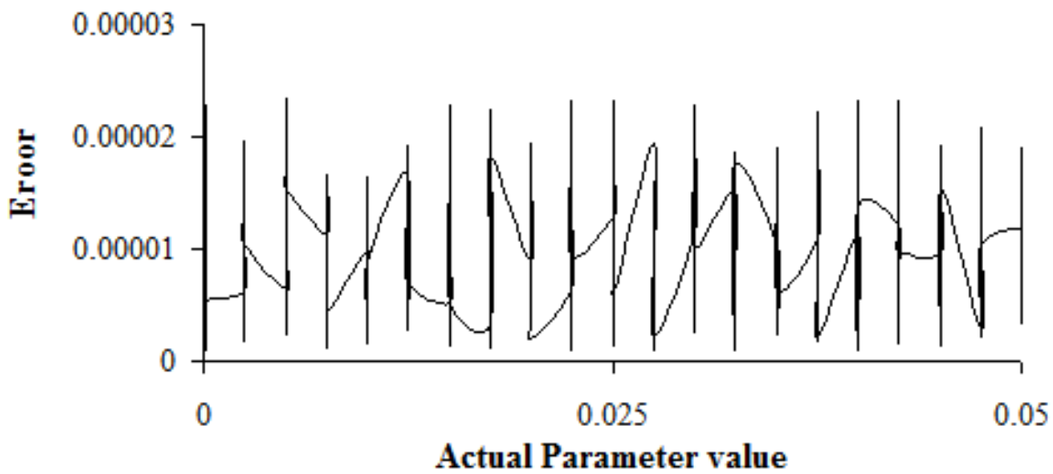


Figure 3.26. Absolute error of estimated parameter σ^2, for the range of 0.0001 – 0.05.

A similar approach is also applied to other parameter ranges. The precision of the estimates given by ANN models shrinks with highly heterogeneous data. As larger values of parameters indicate excessive stochastic flows, we can expect the accuracy of the prediction to diminish for highly stochastic flows. Nonetheless, the average absolute error for the estimates for a range of 0.15 to 0.2 is approximately 5.5% (for example, Figure 3.28 illustrates the error of estimated parameter σ^2 for parameter range of 0.15 to 0.2), which may be acceptable for the most of practical applications.

The above prediction accuracy analyses of the ANN models are based on similar ranges for both parameters. However, in real world applications we may have to deal with extremely different values for two parameters. Therefore, the robustness of the ANN method for different values of parameter regimes for two parameters need to be assessed.

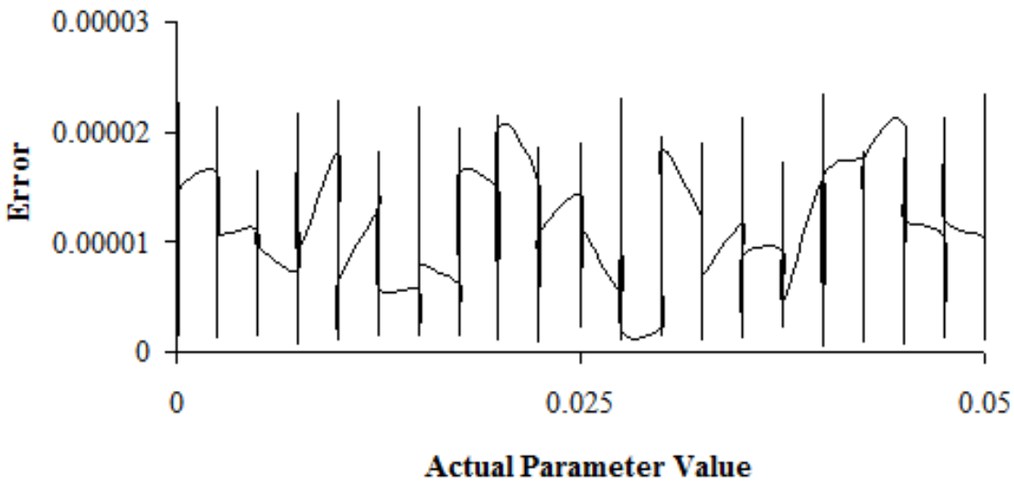


Figure 3.27. Absolute error of estimated parameter *b*, for the range of 0.0001 – 0.05.

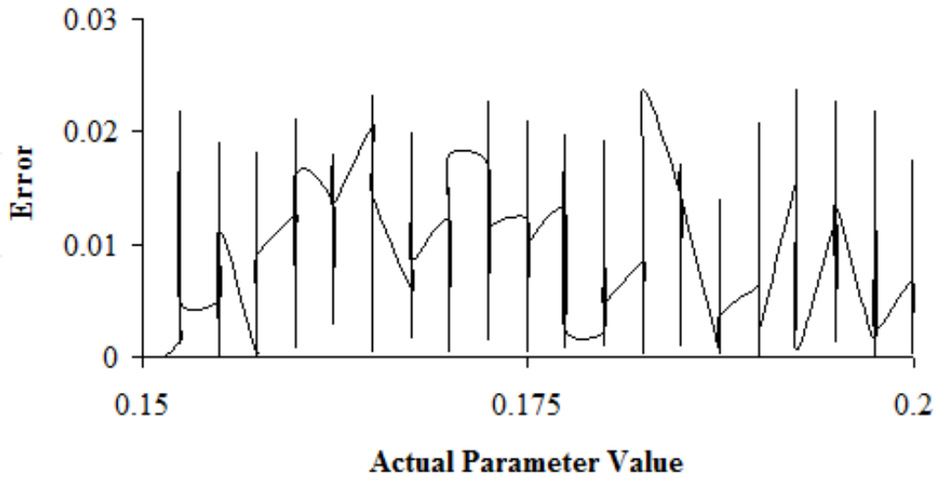


Figure 3.28. Absolute error of estimated parameter σ^2, for the range of 0.15– 0.2.

We generate two separate datasets for two extreme cases:

where σ^2 is smaller and b is higher - σ^2 ranges of 0.0001 – 0.05 and b ranges of 0.15 – 0.2; and where σ^2 is higher and b is smaller - σ^2 ranges of 0.15 – 0.2 and b ranges of 0.0001 – 0.05.

A similar method to that which was used for earlier investigations is employed to gauge the capability of the ANN model. Figures 3.29 and 3.30 reveal that the trained network has predicted the estimates with reasonable precision. In both cases the percentage average absolute error is approximately 4%.

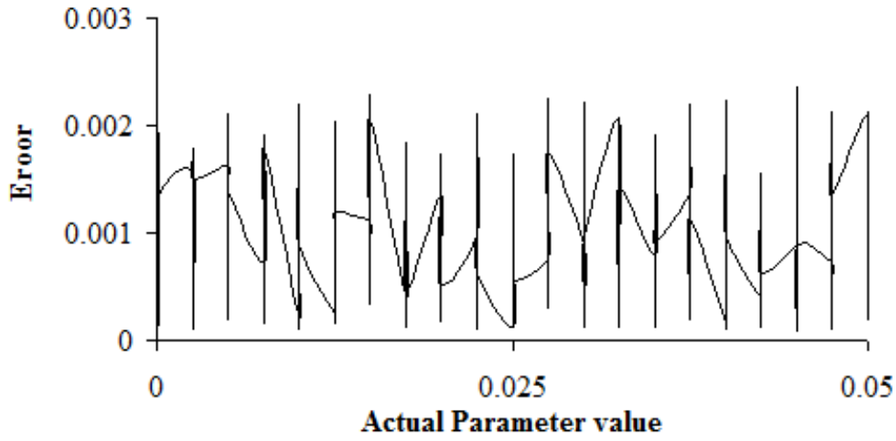


Figure 3.29. Absolute error of estimated parameter σ^2, for the σ^2 range of 0.0001 – 0.05 and b range of 0.15 – 0.2.

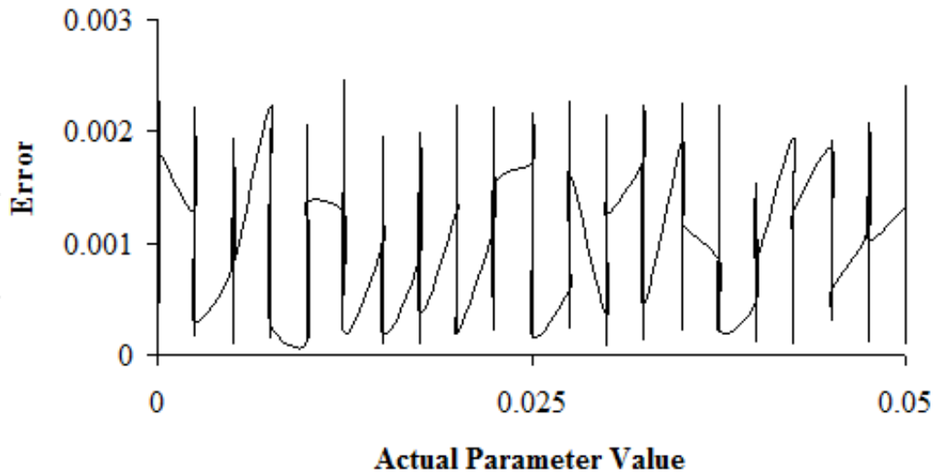


Figure 3.30. Absolute error of estimated parameter b, for the σ^2 range of 0.15 – 0.2 and b range of 0.0001 – 0.05.

We use the artificial aquifer data to validate the developed hybrid ANN method. We make use of the same aquifer dataset in this section to estimate the parameters of SSTM for the

aquifer. Additionally, we use the same data to partially validate the stochastic model (SSTM) previously using a curve fitting technique to approximate the aquifer parameters. It was found that the approximate artificial aquifer parameters are variance, σ^2 = 0.01 and correlation length, b = 0.01.

In the present validation, first, we used the known conditions such as initial and boundary conditions, and hydraulic gradient to simulate the aquifer using the SSTM. The initial concentration at x = 0 is 1.0 unit and it is reduced exponentially with time. Initial values of other spatial points are considered as zero. 1681 data patterns are generated for different combinations of parameters, σ^2 and b. Single standard Wiener process increments are retained for every simulation run. Both parameters varied between 0.0001 and 0.2.

The 1681 generated data patterns are fed into Kohonen's self-organising map architecture to cluster them into four different groups. After classifying them with reasonable accuracy the aquifer dataset is fed into the SOM model to identify relevant groups that the data resemble the most. In this case, we make an assumption that the effect of the transverse dispersion of the flow is reflected in the stochastic flow described by SSTM. As shown in Figure 3.13, the data have been collected along five wells at four levels (A to E wells at levels BR, BL, RE and YE). We consider the data collected along each well at a certain level as a one-dimensional flow path. Hence, we reproduce 20 different one-dimensional datasets.

Concentration values of the aquifer are normalised to enable to weigh them against the normalised SSTM data.

Initially, a dataset closer to the middle of the aquifer, well C – level RE is chosen for the estimation of parameters. An aquifer dataset has to be interpolated to produce missing data and fabricate uniform spatial and temporal grids. Having constructed an exact number of data for similar spatial and time intervals as for the original ANN model, the aquifer dataset is fed into the trained model. The selected dataset is then separated from the larger set and trained for the smaller range selected. Based on the findings described previously, a five layer standard connections is used with the same activation functions, initial weights, momentum and learning rates.

After completing sufficient training, the artificial aquifer dataset is fed into the ANN model to estimate parameters. The estimates that are produced by the model are σ^2 = 0.01364 and b = 0.01665. The estimates produced by the ANN model show close resemblance to the values given by the curve fitting technique. Since we avoid considering lateral dispersions for estimation by ANN model, the results may be subject to slight errors.

We extend the estimation procedure to determine the parameters of the Lincoln University aquifer for other flow lines. Since the earlier dataset is closer to the middle of the aquifer, we choose the next dataset that is nearer to the boundary of the aquifer. Estimates obtained using well A – level YE are σ^2 = 0.01483 and b = 0.00912. These estimates are similar to earlier values.

3.10 Dispersivity Based on the SSTM

To estimate the parameter of the SSTM, we develop a procedure consisting of the following steps: (1) we generate a large number of realizations of concentration (usually 100) for a particular set of values of σ^2 and b using the SSTM; (2) estimate the diffusion coefficient

(*D*) using the maximum likelihood estimation procedure for the 1-D advection-dispersion equation for a given velocity using each of the realization; and (3) take the mean of the estimates of *D* as the dispersivity for the given set of parameters

We have demonstrated that the SSTM can be used to characterise an experimental homogeneous sand aquifer of 5 m width x 10 m length x 2.7 height using a single set of values of (σ^2 = 0.01 and *b* = 0.01) quite satisfactorily. It has also been shown that the numerical solutions of SSTM, in conjunction with the parameter estimation methods such as maximum likelihood method and artificial neural networks, can be used to estimate reliable effective dispersion coefficients, therefore, effective dispersivity, for different scale experiments up to 10 meters.

We use this procedure, which we call stochastic inversive method (SIM) to estimate *D* for other combinations of σ^2 and *b* for different flow lengths. Table 3.8 exhibits the estimated dispersion coefficients (*D*) for the range of scales; 1, 10, 20, 30, 50 and 100 m.

We need to remember that the SSTM is based on the velocity covariance kernel, $\sigma^2 e^{\frac{-|x_1 - x_2|}{b}}$, and a different kernel would give different D values. In addition, SIM is not accurate for very noisy realizations because the theory of estimation is strictly valid when noise is small and Gaussian (Kutoyants, 1984).

b	σ^2	Estimated *D*					
		1 m	10 m	20 m	30 m	50 m	100 m
0.0001	0.0001	0.01634	0.03502	0.05804	0.07328	0.09447	0.12493
0.0001	0.001	0.01942	0.03738	0.06126	0.07526	0.09930	0.13006
0.0001	0.01	0.03844	0.05287	0.08469	0.10502	0.13591	0.18294
0.0001	0.05	0.04758	0.07799	0.12986	0.16064	0.20645	0.27009
0.0001	0.1	0.06786	0.08690	0.14214	0.17914	0.23342	0.31205
0.0001	0.15	0.06968	0.09060	0.15044	0.18790	0.24022	0.31327
0.0001	0.2	0.07047	0.09259	0.15717	0.20209	0.25756	0.33937
0.0001	0.25	0.07188	0.09382	0.15905	0.19747	0.25542	0.33289
0.0001	0.3	0.07258	0.09466	0.16131	0.19895	0.26026	0.34742
0.001	0.0001	0.01917	0.03738	0.05519	0.07807	0.09628	0.12984
0.001	0.001	0.03749	0.05289	0.08020	0.11085	0.13335	0.18273
0.001	0.01	0.06739	0.08698	0.13530	0.18702	0.23005	0.31770
0.001	0.05	0.07424	0.09650	0.15306	0.21656	0.26236	0.36391
0.001	0.1	0.08492	0.09910	0.15522	0.21291	0.25431	0.35222

0.001	0.15	0.08671	0.10782	0.17443	0.24632	0.29578	0.40749
0.001	0.2	0.09482	0.11886	0.18858	0.26626	0.32559	0.44440
0.001	0.25	0.10701	0.12933	0.20057	0.28264	0.33817	0.46265
0.001	0.3	0.11008	0.13860	0.21898	0.30438	0.36830	0.50826
0.01	0.0001	0.02541	0.05302	0.08035	0.11327	0.14376	0.20657
0.01	0.001	0.05697	0.08773	0.13444	0.19613	0.24574	0.34753
0.01	0.01	0.07906	0.10003	0.15332	0.22578	0.28589	0.40472
0.01	0.05	0.09457	0.16952	0.26730	0.38284	0.48728	0.69328
0.01	0.1	0.11483	0.21334	0.33450	0.48658	0.61670	0.87920
0.01	0.15	0.13745	0.24065	0.37173	0.54195	0.67787	0.96664
0.01	0.2	0.15574	0.25994	0.40478	0.58780	0.73802	1.05102
0.01	0.25	0.18468	0.27120	0.41747	0.59966	0.75083	1.06491
0.01	0.3	0.18994	0.27751	0.43013	0.62369	0.78586	1.11591
0.05	0.0001	0.01874	0.07828	0.12467	0.17543	0.22998	0.34209
0.05	0.001	0.03559	0.10128	0.16245	0.22834	0.29946	0.44681
0.05	0.01	0.06957	0.15994	0.25211	0.34326	0.45352	0.68013
0.05	0.05	0.07651	0.26271	0.41319	0.56701	0.75557	1.13054
0.05	0.1	0.08450	0.29209	0.45872	0.63193	0.83988	1.26490
0.05	0.15	0.08725	0.29659	0.47462	0.66040	0.88362	1.33137
0.05	0.2	0.08987	0.29700	0.46840	0.64492	0.85205	1.27735
0.05	0.25	0.09219	0.29480	0.46693	0.65051	0.87047	1.30991
0.05	0.3	0.09294	0.29137	0.45748	0.63153	0.83935	1.26428
0.1	0.0001	0.01797	0.08546	0.14242	0.19805	0.26901	0.42102
0.1	0.001	0.02971	0.10536	0.16856	0.23555	0.31622	0.48560
0.1	0.01	0.05925	0.17795	0.29101	0.41507	0.56815	0.88488
0.1	0.05	0.06452	0.28179	0.45334	0.64813	0.87743	1.35846
0.1	0.1	0.07232	0.29698	0.48594	0.68641	0.93357	1.45105
0.1	0.15	0.07485	0.29776	0.47662	0.67992	0.91941	1.42114
0.1	0.2	0.07587	0.29876	0.48105	0.68887	0.93379	1.44814
0.1	0.25	0.07608	0.29928	0.48535	0.68808	0.93290	1.45029

0.1	0.3	0.07678	0.29791	0.48068	0.67575	0.91363	1.41871
0.15	0.0001	0.01731	0.08784	0.14149	0.20423	0.28678	0.45299
0.15	0.001	0.02570	0.10921	0.18047	0.25898	0.35905	0.56336
0.15	0.01	0.05454	0.16218	0.26117	0.37775	0.52802	0.82934
0.15	0.05	0.06574	0.24498	0.40463	0.57946	0.81625	1.27853
0.15	0.1	0.07627	0.24830	0.40532	0.58494	0.82098	1.28586
0.15	0.15	0.07722	0.24896	0.40327	0.58331	0.82088	1.29136
0.15	0.2	0.07785	0.25016	0.40435	0.58084	0.81027	1.27489
0.15	0.25	0.07807	0.25278	0.41587	0.59926	0.83683	1.32106
0.2	0.3	0.07527	0.27394	0.44577	0.65306	0.93236	1.47560
0.25	0.0001	0.01694	0.08642	0.14037	0.20553	0.29769	0.47547
0.25	0.001	0.02292	0.11779	0.19345	0.28363	0.41296	0.65983
0.25	0.01	0.04829	0.13433	0.22093	0.32259	0.46790	0.74928
0.25	0.05	0.06187	0.18828	0.30661	0.45919	0.67190	1.08123
0.25	0.1	0.07470	0.23226	0.37711	0.55761	0.81495	1.30630
0.25	0.15	0.07537	0.25020	0.41219	0.60827	0.88655	1.42797
025	0.2	0.07559	0.26583	0.42875	0.63902	0.92617	1.48670
0.25	0.25	0.07567	0.28102	0.46074	0.68281	0.99764	1.60438
0.25	0.3	0.07593	0.29162	0.46948	0.69110	1.00010	1.61016
0.3	0.0001	0.01785	0.08229	0.13906	0.21167	0.31537	0.51395
0.3	0.001	0.02169	0.11016	0.18535	0.28209	0.42416	0.69259
0.3	0.01	0.04787	0.14860	0.25288	0.38448	0.58112	0.95251
0.3	0.05	0.06316	0.22140	0.38149	0.57637	0.87271	1.42535
0.3	0.1	0.07527	0.25171	0.42778	0.64452	0.97782	1.58726
0.3	0.15	0.07728	0.27513	0.47773	0.73238	1.10681	1.80695
0.3	0.2	0.07799	0.28755	0.49003	0.73771	1.11325	1.81448
0.3	0.25	0.07824	0.29803	0.50666	0.77144	1.17052	1.91032
0.3	0.3	0.07911	0.30699	0.53747	0.80802	1.22297	1.98500

Table 3.8. Estimates of D obtained by using a stochastic inverse method for different combinations of parameters of SSTM for different flow lengths (velocity = 0.5 m/day).

We can summarise some of the estimates in the plots in Figures 3.31, 3.32 and 3.33 as functions of the scales of the experiments. What these plots show is that, for a given set of parameters, the SSTM would give the estimates of dispersivity increasing with the flow length.

As discussed in chapter 1, Pickens and Grisak (1981), and Lallemand-Barres and Peaudecerf (1978, cited in Fetter, 1999) showed that the scale dependency of α_L has a linear relationship of $\alpha_L = 0.1\ L$, where L is the mean travel distance. However, Pickens and Grisak (1981) recognised that the linear increase of dispersivity with the mean travel distance was unlikely for large travel distances. It was expected that tracer migration between aquifer layers could cause a reduction in the magnitude of the proportionality constants, since the transverse migration would tend to reduce the spreading effect caused by the stratification. Field measurements obtained by Gelhar (1986) illustrate that the scale dependence relationship between α_L and the flow length is non-linear (Figure 3.34).

To evaluate the comparative estimates of D obtained from the inverse method for the SSTM parameters and the field measurements observed by Gelhar (1986), we plot them on the same graph (Figure 3.35 – 3.37). Only reliable observations of Figure 3.34 (indicated by larger symbols) are considered. Since the parameter estimated from the inverse approach is D, α_L values of Figure 3.34 are converted to D ($D = \alpha_L v$). Furthermore, we plot the relationship of $\alpha_L = 0.1\ L$ in the same graph to assess our estimates. Three different ranges of b are chosen. Figure 3.35 shows the estimates for smaller b, 0.0001 m, for four values of σ^2 (0.0001, 0.05, 0.2 and 0.3). Figures 3.36 and 3.37 illustrate the similar σ^2 values for a mid range value b, 0.01 m, and larger b, 0.3, respectively.

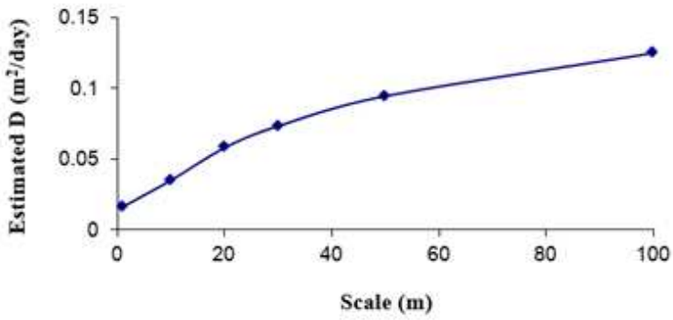


Figure 3.31. D for the parameter combination of b = 0.0001 and σ^2 = 0.0001.

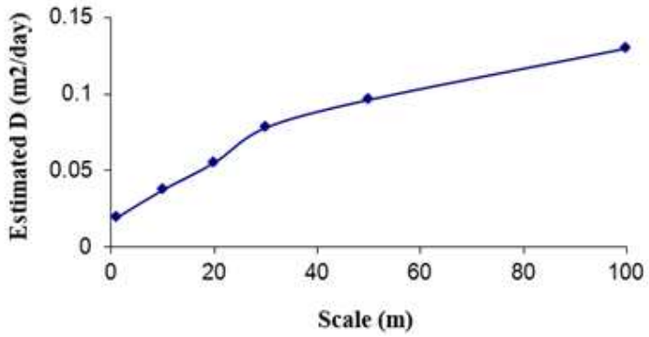


Figure 3.32 D for the parameter combination of b = 0.001 and σ^2 = 0.0001.

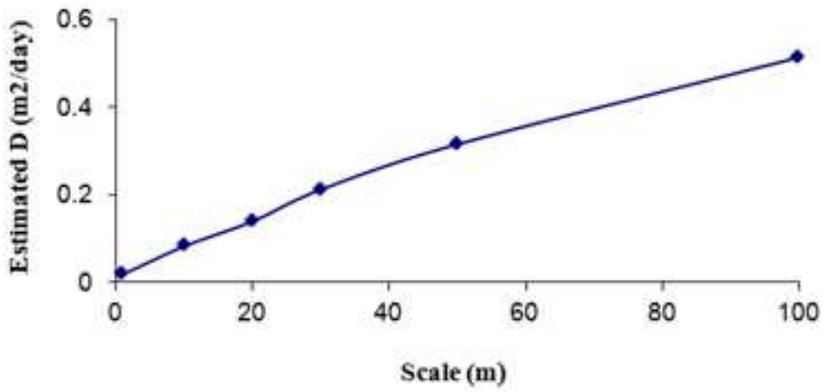


Figure 3.33. *D* for the combination of b = 0.3 and σ^2 = 0.0001.

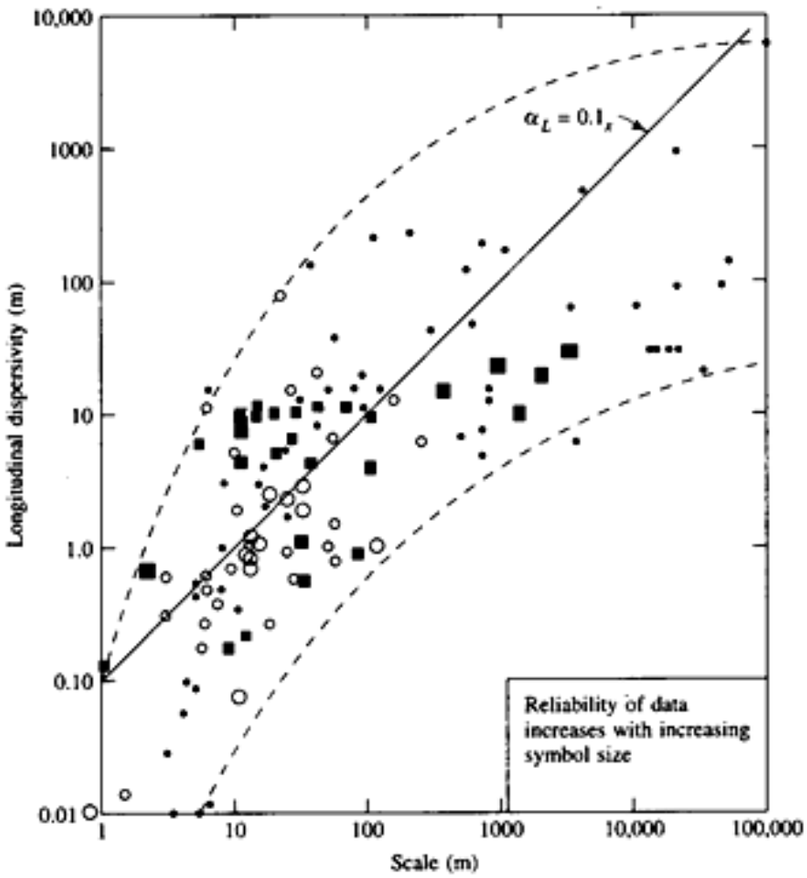


Figure 3.34. Field measured values of longitudinal dispersivity as a function of the scale of measurement. The largest circles represent the most reliable data. The estimated dispersivity from the SSTM are given by the squares. (Source: Gelhar (1986).)

Figures 3.35 – 3.37 demonstrate that corresponding *D* values obtained for SSTM parameters do not agree with the relationship of $\alpha_L = 0.1\ L$. However, for mid and larger ranges of *b*, estimated *D* s are in reasonable agreement with the most of the reliable field measurements

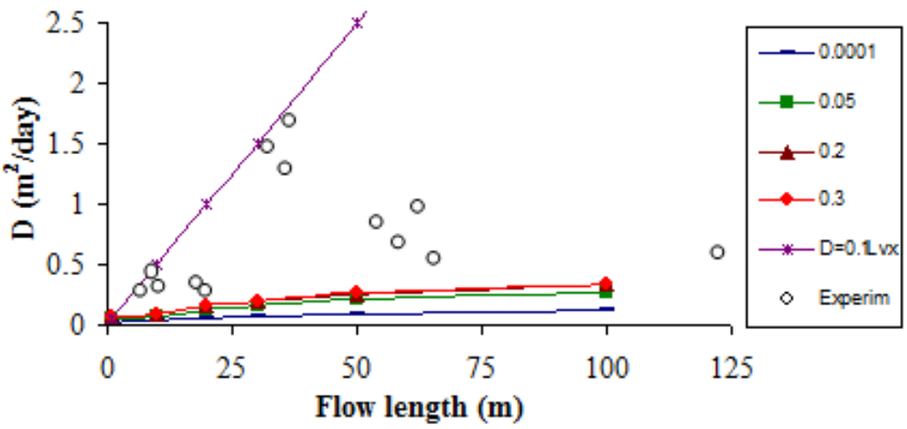


Figure 3.35. Estimated D values for b = 0.0001 m for range of σ^2 (0.0001 – 0.3), D= *0.1 L v,* and reliable field measurements observed by Gelhar (1986) for different flow lengths.

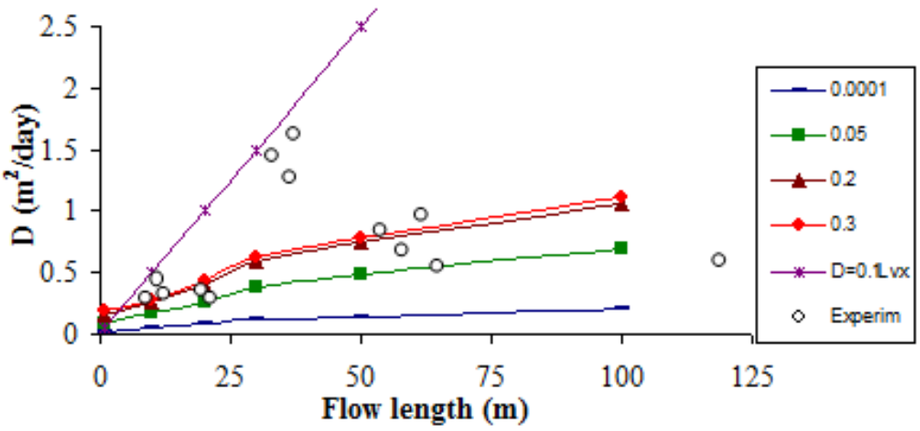


Figure 3.36. Estimated D values for b = 0.01 m for range of σ^2 (0.0001 – 0.3), D= *0.1Lv,* and reliable field measurements observed by Gelhar (1986) for different flow lengths.

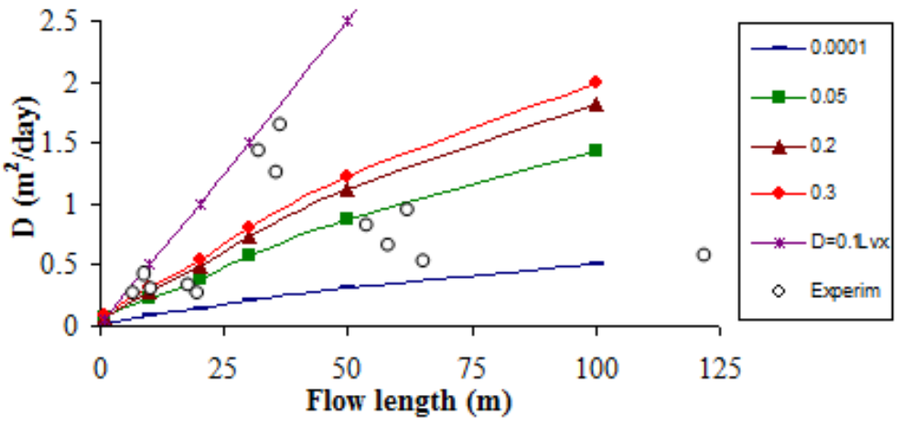


Figure 3.37 Estimated D values for b = 0.1 m for range of σ^2 (0.0001 – 0.3), D= *0.1Lv,* and reliable field measurements observed by Gelhar (1986) for different flow lengths.

observed by Gelhar (1986). Data in Figure 3.35 represent a solute transport system with very low stochasticity, which may not be realistic in real world aquifers. Figures 3.38 and 3.39 may be a better representation of an actual aquifer. Note that the α_L values of Figure 3.36 were obtained from many sites around the world. The estimated data that agrees with field measurements may be obtained from a variety of soil and heterogeneity.

To simplify the computational burden, we make use of the fact that the expected value of mean $d\beta(t)$ is zero and the variance is given by equation (3.4.5). Instead of calculating $d\beta(t)$ for each x within a, we take the mean of $Var[d\beta]$ over the length a,

$$E[Var(d\beta(t))]=\frac{\int_0^a\left(\sum_{i=1}^{M}(\eta_i^2)\,\Delta t\right)dx}{a}=\left(\frac{\int_0^a\sum_{i=1}^{M}(\lambda_i f_i^2(x))dx}{a}\right)\Delta t \tag{3.10.1}$$

$$\text{But by orthogonality, } \int_0^a f_i^2(x)dx=1\text{ , and,} \tag{3.10.2}$$

$$\text{it can be shown that approximately, } \underset{M\to\infty}{Limit}\sum_{i=1}^{M}\lambda_i=\sigma^2 a\ . \tag{3.10.3}$$

$$\text{Therefore, } E[Var(d\beta(t))]=\sigma^2\Delta t\ . \tag{3.10.4}$$

To illustrate this approximation, the number of eigenvalues required for each a (M) is determined by calculating a number of roots that are sufficient for the Mth eigenvalue to reach a value ε, where ε was set to be 5% of the value of the first eigenvalue. The relationship between a and the number of roots required is approximately 150 * a, for a = 1 to 10 000 metres. For example, when σ^2 =1.0 and b=0.01, M = {150,750,1500,7500,15000,75000,150000,300000,450000}, then for a = {1,5,10,50,100,500,1000,2000,3000}. $\sum_{i=1}^{M}\lambda_i$ = {0.8762, 4.2920, 8.5443, 42.5518, 85, 04312, 425.0663, 850.1131, 1710.6345, and 2684.8313}. This shows the extent of the computational problem without the approximation for $d\beta(t)$. We have computed ($E[Var(d\beta(t))]/\Delta t$) using equation (3.10.1) for different values of a with b=0.01, Δt =0.001, $\Delta x=0.01$, σ^2 =1.0, and 150* a number of eigen values. For each a , the computation was done using a set of routines written in the Python™ language with the *Numeric* extension for fast array operations. Computing ($E[Var(d\beta(t))]/\Delta t$) for larger values of a is a computationally expensive operation with the computing times increasing exponentially with a. As an example, for a = 1000 meters, it takes approximately 200 minutes, with a = 2000 taking approximately 760 minutes on a 1.8 GHZ computer. For a ={1,5,10,50,100,500,1000,2000,3000}, the corresponding values for $E[Var(d\beta(t))]/\Delta t$ are {0.8653, 0.8647,0.8630,0.8615,0.8611,0.8610,0.8610,0.8655, 0.8983}, respectively, and we can expect that for an infinite number of eigen functions it will approach 1.0. Therefore, the

spatially averaged $d\beta(t)$ is a Gaussian random variable with zero-mean and $\sigma^2 \Delta t$ - variance, which can be readly incorporated into the numerical solution scheme for SSTM.

Because of the high computational times involved, instead of using the SIM procedure, we estimate the dispersivity for a computational experiment by limiting ourselves to the SSTM parameters obtained for experimental sand aquifer (i.e., σ^2 =0.01; b=0.01 m) in the following way:

(a) Set the initial and boundary conditions as,

$$C(x,0)=0, \quad x \geq 0$$
$$C(0,t)=1.0, \; t \geq 0 \quad \text{and}$$
$$C(\infty,t)=0, \quad t \geq 0$$

(b) Solve equation (21) assuming mean velocity to be 1.0 m/day,

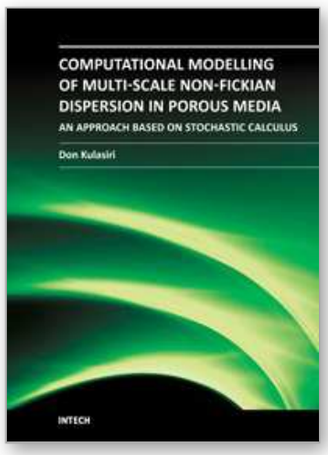
(c) Use a realization of $C(x,t)$ at x=a to estimate the dispersivity, α_L , using equation (3.10.5) (Fetter, 1999) given below using nonlinear regression: (Mathematica® was used for this purpose.)

$$C(x,t)=0.5\left(erfc\left(\frac{a-t}{2\sqrt{\alpha_L t}}\right)+\exp\left(\frac{a}{\alpha_L}\right)erfc\left(\frac{a+t}{2\sqrt{\alpha_L t}}\right)\right). \tag{3.10.5}$$

Equation (3.10.5) is the analytical solution for the one-dimensional advection-dispersion equation for the initial and boundary conditions given in (a) above. To get a reliable estimate of α_L for given a, we need to have x-axis length of 1.5 a meters and should have $C(x,t)$ realization upto 2 a days. For example, to obtain an estimate of α_L for a = 3000 meters, we need to run the simulation of a domain of 4500 meters for 6000 days. However, to reduce the computational time, one can use higher Δx and Δt values than ideally suitable in solving SPDEs thereby sacrificing the reliability. (Δx =0.01 m and Δt =0.00001 days would give very good solutions to SSTM.) However, this approximate procedure is only valid for the velocity covariance kernel used in this chapter. This procedure is not as reliable as the SIM.

We overlay some representative values of dispersivities from the SSTM on a graph of the field measurements obtained by Gelhar (1986) from different experiments in Figure 3.34, which shows that SSTM could model the multi-scale dispersion with a single set of parameters, σ^2 =0.01; b=0.01 m, that would give rise to similar non-linear scale dependency of "deterministic" dispersivity evaluated using the realisations of SSTM.

In Figure 3.34, the larger (hollow) circles depict the most reliable experimental data whereas the smaller (filled) circles give the data with lesser experimental accuracy. The dotted lines show the bounded region of experimental data. The estimated dispersivity values (filled squares) from SSTM are within the bounded region and follow similar trends to the most reliable experimental data. We estimated the dispersivities from SSTM using only a limited number of computational experiments, and each computational experiment produces a random realisation. Therefore, the estimated dispersivities are stochastic quantities just like the experimental values, and it is reasonable to expect discrepancies within the bounded region.



Computational Modelling of Multi-scale Solute Dispersion in Porous Media - An Approach Based on Stochastic Calculus

Edited by

ISBN 978-953-307-726-0
Hard cover, 234 pages
Publisher InTech
Published online 04, November, 2011
Published in print edition November, 2011

This research monograph presents a mathematical approach based on stochastic calculus which tackles the "cutting edge" in porous media science and engineering - prediction of dispersivity from covariance of hydraulic conductivity (velocity). The problem is of extreme importance for tracer analysis, for enhanced recovery by injection of miscible gases, etc. This book explains a generalised mathematical model and effective numerical methods that may highly impact the stochastic porous media hydrodynamics. The book starts with a general overview of the problem of scale dependence of the dispersion coefficient in porous media. Then a review of pertinent topics of stochastic calculus that would be useful in the modeling in the subsequent chapters is succinctly presented. The development of a generalised stochastic solute transport model for any given velocity covariance without resorting to Fickian assumptions from laboratory scale to field scale is discussed in detail. The mathematical approaches presented here may be useful for many other problems related to chemical dispersion in porous media.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Don Kulasiri (2011). A Stochastic Model for Hydrodynamic Dispersion, Computational Modelling of Multi-scale Solute Dispersion in Porous Media - An Approach Based on Stochastic Calculus, (Ed.), ISBN: 978-953-307-726-0, InTech, Available from: http://www.intechopen.com/books/computational-modelling-of-multi-scale-solute-dispersion-in-porous-media-an-approach-based-on-stochastic-calculus/a-stochastic-model-for-hydrodynamic-dispersion

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen