4

A Generalized Mathematical Model in One-Dimension

4.1 Introduction

In the previous chapter we derived a stochastic solute transport model (equation (3.2.14)); we developed the methods to estimate its parameters, and investigated its behaviour numerically. We see some promise to characterise the solute dispersion at different flow lengths, and there are some indications that equation (3.2.14) produce the behaviours that would be interpreted as capturing the scale-dependency of dispersivity. However, there are weaknesses in the model as evident from Chapter 3. These weaknesses, which are discussed in the next section, are stemming from the very assumptions we made in the development of the model. One could argue that by relaxing the Fickian assumptions, we are actually complicating the problem quite unnecessarily. But as we see in Chapter 3 and in this chapter, we develop a new mathematical and computational machinery at a more fundamental level for the hydrodynamic dispersion in saturated porous media.

We see that equation (3.2.14) is based on assuming a covariance kernel for the velocity fluctuations, and the solution is dependent on solving an integral equation (see equation (3.3.11)). In Chapter 3, the integral equation is solved analytically for the covariance kernel given by equation (3.3.10) to obtain the eigen values and eigen functions, but analytical solutions of integral equations can not be easily derived for any arbitrary covariance kernel. This limits the flexibility of the SSTM in employing a suitable covariance kernel independent of the ability to solve relevant integral equations. Further, we need to solve the SSTM in a much more computationally efficient manner, and estimating dispersivity by always relating to the deterministic advection-dispersion equation is not quite satisfactory. Therefore, we seek to develop a more general form of equation (3.2.14) in this chapter.

4.2 The Development of the Generalized Model

We restate equation (3.2.14) in the differential form:

$$dC = S\left(\overline{V}(x,t)\, C(x,t)\right) dt + S\left(C(x,t)\, d\beta_m(t)\right), \tag{4.2.1}$$

$$\text{where} \quad d\beta_m(t) = \sigma \sum_{j=1}^{m} \sqrt{\lambda_j} f_j db_j(t). \tag{4.2.2}$$

We use the same notations and symbols as in Chapter 3. In equation (4.2.2), $d\beta_m(t)$ is calculated by summing m terms of ($\sqrt{\lambda_j} f_j\, db_j(t)$), and for each eigen function, f_j, there is an associated independent Wiener process increment ($db_j(t)$).

This Wiener process increment is δ-correlated in time only, and for a particular point in time, t_i, the Wiener process $B(\omega, f)$ generates the Wiener increments $db_j(t_i)$ for each of the eigen functions. At another time, t_k, $B(\omega, f)$ generates $db_j(t_k)$ for each of the eigen function. However, as the Wiener increments are Gaussian with Δt variance (see Chapter 2), i.e., $db_j(t_i)$ and $db_j(t_k)$ are essentially the same stochastic variable.

The number of terms that need to be summed up, *m*, has to be determined by considering the contribution each pair of eigen value λ_i and corresponding eigen function f_j make to approximate the covariance.

Instead of solving the integral equation to obtain eigen function, we assume the following function to be a generalized eigen function. (We will discuss the method to obtain this function for any given kernel in section 4.3.)

We define,

$$f_j(x) = g_{0j} + g_{1j}x + \sum_{k=2}^{p_j} g_{kj} e^{-r_{kj}\left(x - s_{kj}\right)^2}, \tag{4.2.3}$$

where $f_j(x)$ is the *j*th eigen function of a given covariance kernel, g_{0j}, g_{1j} and g_{kj} are numerical coefficients, r_{jk} and s_{jk} are the coefficients in the exponent where *k* is an integer index ranging from 2 to p_j. p_j is determined by the computational method in section 4.3.

The differential operator is defined by,

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

and the second term on the right hand side of equation (4.2.1) can be written as,

$$S\left(C(x,t)\sum_{j=1}^{m}\sqrt{\lambda_j} f_j db_j(t)\right) = \sigma\sum_{j=1}^{m}\sqrt{\lambda_j} S\left(C(x,t) f_j\right) db_j(t), \tag{4.2.4}$$

after substituting equation (4.2.2) into the second terms of equation (4.2.1) and taking $\sqrt{\lambda_j}$ and $db_j(t)$ out of the operand. We can now expand the differential,

$$S\left(C(x,t) f_j\right) = S\left(C(x,t)\left(g_{0j} + g_{1j}x + \sum_{k=2}^{p_j} g_{kj} e^{-r_{kj}\left(x - s_{kj}\right)^2}\right)\right). \tag{4.2.5}$$

Let us evaluate the derivatives separately;

$$\frac{\partial}{\partial x}\left(C(x,t)\left(g_{0j}+g_{1j}x+\sum_{k=2}^{p_j}g_{kj}e^{-r_{kj}\left(x-s_{kj}\right)^2}\right)\right)$$
$$=\frac{\partial}{\partial x}\left[g_{0j}C(x,t)+g_{1j}xC(x,t)+\sum_{k=2}^{p_j}g_{kj}C(x,t)e^{-r_{kj}\left(x-s_{kj}\right)^2}\right], \tag{4.2.6}$$
$$=g_{0j}\frac{\partial C(x,t)}{\partial x}+g_{1j}x\frac{\partial C(x,t)}{\partial x}+g_{1j}C(x,t)+\sum_{k=2}^{p_j}g_{kj}\frac{\partial}{\partial x}\left(C(x,t)e^{-r_{kj}\left(x-s_{kj}\right)^2}\right).$$

Now the derivative within the summation in equation (4.2.6) is evaluated.

$$\frac{\partial}{\partial x}\left(C(x,t)e^{-r_{kj}\left(x-s_{kj}\right)^2}\right)=C(x,t)\left(-2r_{kj}\left(x-s_{kj}\right)e^{-r_{kj}\left(x-s_{kj}\right)^2}\right)+e^{-r_{kj}\left(x-s_{kj}\right)^2}\frac{\partial C(x,t)}{\partial x}$$
$$=\left[-2r_{kj}C(x,t)\left(x-s_{kj}\right)+\frac{\partial C(x,t)}{\partial x}\right]e^{-r_{kj}\left(x-s_{kj}\right)^2}.$$

Substituting this derivation back to equation (4.2.6) and defining, $\psi_j(x,t)$ for eigen function *j*,

$$\psi_j(x,t)\equiv\frac{\partial}{\partial x}\left(C(x,t)\left(g_{0j}+g_{1j}x+\sum_{k=2}^{p_j}g_{kj}e^{-r_{kj}\left(x-s_{kj}\right)^2}\right)\right),$$
$$=g_{0j}\frac{\partial C(x,t)}{\partial x}+g_{1j}x\frac{\partial C(x,t)}{\partial x}+g_{ij}C(x,t)+\sum_{k=2}^{p_j}g_{kj}\left[\frac{\partial C(x,t)}{\partial x}-2r_{kj}\left(x-s_{kj}\right)C(x,t)\right]e^{-r_{kj}\left(x-s_{kj}\right)}, \tag{4.2.7}$$
$$=\left\{g_{1j}-2\sum_{k=2}^{p_j}g_{kj}\,r_{kj}\left(x-s_{kj}\right)e^{-r_{kj}\left(x-s_{kj}\right)^2}\right\}C(x,t)+\left\{g_{0j}+g_{1j}\,x+\sum_{k=2}^{p_j}g_{kj}e^{-r_{kj}\left(x-s_{kj}\right)^2}\right\}\frac{\partial C(x,t)}{\partial x}.$$

Then taking the derivative of $\psi_j(x,t)$ with respect to x,

$$\frac{\partial\psi(x,t)}{\partial x}=\left\{g_{1j}-2\sum_{k=2}^{p_j}g_{kj}r_{kj}\left(x-s_{kj}\right)e^{-r_{kj}\left(x-s_{kj}\right)^2}\right\}\frac{\partial C(x,t)}{\partial x}$$
$$+\left\{4\sum_{k=2}^{p_j}g_{kj}r_{kj}^2\left(x-s_{kj}\right)^2e^{-r_{kj}\left(x-s_{kj}\right)^2}-2\sum_{k=2}^{p_j}g_{kj}r_{kj}e^{-r_{kj}\left(x-s_{kj}\right)^2}\right\}C(x,t)$$
$$+\left\{g_{0j}+g_{1j}x+\sum_{k=2}^{p_j}g_{kj}e^{-r_{kj}\left(x-s_{kj}\right)^2}\right\}\frac{\partial^2C(x,t)}{\partial x^2}+\left\{g_{1j}-2\sum_{k=2}^{p_j}g_{kj}r_{kj}\left(x-s_{kj}\right)e^{-r_{kj}\left(x-s_{kj}\right)^2}\right\}\frac{\partial C(x,t)}{\partial x},$$

$$
\begin{aligned}
&= 2\left\{ g_{1j} - 2\sum_{k=2}^{p_j} g_{kj} r_{kj} \left(x - s_{kj}\right) e^{-r_{kj}\left(x-s_{kj}\right)^2} \right\} \frac{\partial C(x,t)}{\partial x} + \left\{ g_{0j} + g_{1j} x + \sum_{k=2}^{p_j} g_{kj} e^{-r_{kj}\left(x-s_{kj}\right)^2} \right\} \frac{\partial^2 C(x,t)}{\partial x^2} \\
&+ \left\{ 4\sum_{k=2}^{p_j} g_{kj} r_{kj}^2 \left(x - s_{kj}\right)^2 e^{-r_{kj}\left(x-s_{kj}\right)^2} - 2\sum_{k=2}^{p_j} g_{kj} r_{kj} e^{-r_{kj}\left(x-s_{kj}\right)^2} \right\} C(x,t)
\end{aligned}
\tag{4.2.8}
$$

Therefore, equation (4.2.5) can be expressed as,

$$
\begin{aligned}
-S\left(C(x,t) f_j\right) &= \frac{h_x}{2}\left[\frac{\partial^2 \left(C(x,t) f_j\right)}{\partial x^2} \right] + \left[\frac{\partial \left(C(x,t) f_j\right)}{\partial x} \right], \\
&= \psi_j(x,t) + \frac{h_x}{2} \frac{\partial \psi_j(x,t)}{\partial x}, \\
&= P_{0j}(x) C(x,t) + P_{1j} \frac{\partial C(x,t)}{\partial x} + P_{2j} \frac{\partial^2 C(x,t)}{\partial x^2},
\end{aligned}
\tag{4.2.9}
$$

Where

$$
\begin{aligned}
P_{0j}(x) &= \left[g_{ij} - 2\sum_{k=2}^{p_j} g_{kj} r_{kj} \left(x - s_{kj}\right) e^{-r_{kj}\left(x-s_{kj}\right)^2} \right] \\
&+ \left(\frac{h_x}{2}\right)\left[4\sum_{k=2}^{p_j} g_{kj} r_{kj}^2 \left(x - s_{kj}\right)^2 e^{-r_{kj}\left(x-s_{kj}\right)^2} - 2\sum_{k=2}^{p_j} g_{kj} r_{kj} e^{-r_{kj}\left(x-s_{kj}\right)^2} \right],
\end{aligned}
\tag{4.2.10}
$$

$$
P_{1j}(x) = g_{0j} + g_{1j} x + \sum_{k=2}^{p_j} g_{kj} e^{-r_{kj}\left(x-s_{kj}\right)^2} + \left(\frac{h_x}{2}\right)\left[2\left(g_{ij} - 2\sum_{k=2}^{p_j} g_{kj} r_{kj} \left(x - s_{kj}\right) e^{-r_{kj}\left(x-s_{kj}\right)^2} \right) \right], \tag{4.2.11}
$$

And

$$
P_{2j}(x) = \left(\frac{h_x}{2}\right)\left[g_{0j} + g_{1j} x + \sum_{k=2}^{p_j} g_{kj} e^{-r_{kj}\left(x-s_{kj}\right)^2} \right]. \tag{4.2.12}
$$

We see that, in equation (4.2.9), the coefficients, $P_{ij}(x)\,(i = 0,1,2)$ are functions of x only.

If we recall the premise on which stochastic calculus is discussed in Chapter 2, $C(x,t)$ is a continuous, non-differentiable function, and equation (4.2.1) should be interpreted as an Ito stochastic differential, which essentially mean equation (4.2.1) has to be understood only in the following form:

$$
C(x,t) = \int A\left(C(x,t),t,x\right) dt + \sum_i \int B_i\left(C(x,t),t,x\right) dw_i, \tag{4.2.13}
$$

where $A\left(C(x,t),t,x\right)$ is the drift coefficient and $B_i\left(C(x,t),t,x\right)$ are diffusion coefficients of the Ito diffusion, equation (4.2.13); here $dw_i(t)$ are increments of the standard Wiener process. Ito diffusion are an interesting class of stochastic integrals and has many advantageous properties of practical importance (Klebaner, 1998).

The expression of stochastic partial differential equation (equation (4.2.1)) as an Ito diffusion in time would give us the mathematical justification in solving it using Ito calculus. In our development, the eigen functions, $f_j(x)$ are continuous, differentiable functions so are the coefficients, $P_{ij}(x)(i=0,1,2)$. Therefore, the Ito stochastic product rule is the same as the product rule in standard calculus (Klebaner, 1998), and we employ this fact in the previous derivations. Further, we assume that the mean velocity $\overline{V}(x,t)$ is a continuous, differentiable function which is a reasonable assumption given that the average velocity in aquifer situation is based on the hydraulic conductivity, porosity and the pressure gradient across a large enough domain within a much larger total flow length.

Therefore, we can write the drift term of equation (4.2.1) as follows:

$$\begin{aligned}
&-S\left(\overline{V}(x,t)C(x,t)\right)=\frac{\partial\overline{V}(x,t)}{\partial x}C(x,t)+\overline{V}(x,t)\frac{\partial C(x,t)}{\partial x}\\
&+\frac{h_x}{2}\left(\frac{\partial^2\overline{V}(x,t)}{\partial x^2}C(x,t)+2\frac{\partial C(x,t)}{\partial x}\frac{\partial\overline{V}(x,t)}{\partial x}+\overline{V}\frac{\partial^2 C(x,t)}{\partial x^2}\right)\\
&=\left(\frac{\partial\overline{V}(x,t)}{\partial x}+\frac{h_x}{2}\frac{\partial^2\overline{V}(x,t)}{\partial x^2}\right)C(x,t)+\left(\overline{V}(x,t)+\frac{h_x}{2}2\frac{\partial\overline{V}(x,t)}{\partial x}\right)\frac{\partial C(x,t)}{\partial x}\\
&+\frac{h_x}{2}\overline{V}(x,t)\frac{\partial^2 C(x,t)}{\partial x^2}.
\end{aligned}\tag{4.2.14}$$

By substituting equations (4.2.14) and (4.2.9) into equation (4.2.1),

$$\begin{aligned}
dC(x,t)=&-\left\{\begin{aligned}&\left(\frac{\partial\overline{V}(x,t)}{\partial x}+\frac{h_x}{2}\frac{\partial^2\overline{V}(x,t)}{\partial x^2}\right)C(x,t)+\left(\overline{V}(x,t)+\frac{h_x}{2}2\frac{\partial\overline{V}(x,t)}{\partial x}\right)\frac{\partial C(x,t)}{\partial x^2}\\&+\left(\frac{h_x}{2}\overline{V}(x,t)\right)\frac{\partial^2 C(x,t)}{\partial x^2}\end{aligned}\right\}\\
&-\sigma\sum_{j=1}^{m}\sqrt{\lambda_j}\left(P_{0j}C(x,t)+P_{1j}\frac{\partial C(x,t)}{\partial x}+P_{2j}\frac{\partial^2 C(x,t)}{\partial x^2}\right)db_j(t),
\end{aligned}$$

and then,

$$dC(x,t)=-C(x,t)dI_0-\frac{\partial C(x,t)}{\partial x}dI_1-\frac{\partial^2 C(x,t)}{\partial x^2}dI_2,\tag{4.2.15}$$

where,

$$dI_0(x,t)=\left(\frac{\partial\overline{V}(x,t)}{\partial x}+\frac{h_x}{2}\frac{\partial^2\overline{V}(x,t)}{\partial x^2}\right)dt+\sigma\sum_{j=1}^{m}\sqrt{\lambda_j}P_{0j}\,db_j(t),\tag{4.2.16}$$

$$dI_1(x,t)=\left(\bar{V}(x,t)+\frac{h_x}{2}\frac{2\partial\bar{V}(x,t)}{\partial x}\right)dt+\sigma\sum_{j=1}^{m}\sqrt{\lambda_j}P_{1j}\,db_j(t),\ \text{and} \tag{4.2.17}$$

$$dI_2(x,t)=\left(\frac{h_x}{2}\bar{V}(x,t)\right)dt+\sigma\sum_{j=1}^{m}\sqrt{\lambda_j}P_{2j}\,db_j(t). \tag{4.2.18}$$

In equation (4.2.15), dI_0, dI_1 and dI_2 are Ito stochastic differentials with respect to t as well, and we can rewrite a generalized SSTM given by equation (4.2.1) in terms of another set of stochastic differentials,

$$C(x,t)=-\int_t C(x,t)dI_0-\int_t\frac{\partial C(x,t)}{\partial x}dI_1-\int_t\frac{\partial^2 C(x,t)}{\partial x^2}dI_2+C(x,\phi). \tag{4.2.19}$$

We note that dI_0, dI_1 and dI_2 are only functions of the mean velocity, $\bar{V}(x,t)$, and $P_{ij}(x)(i=0,1,2)$; and $C(x,t)$ is separated out in equation (4.2.19), which can be interpreted as $C(x,t)$ and its spatial derivatives are modulating $dI_i(i=0,1,2)$s. I_i are Ito stochastic integrals of the form given by equation (4.2.13), and can be expressed as,

$$I_i(x,t)=\int F_i(x,t)dt+\sigma\sum_{j=1}^{m}\int G_{ij}(x,t)db_j(t),\ \ (i,0,1,2)\text{ and }(j,1,m), \tag{4.2.20}$$

where,

$$F_0(x,t)=\left(\frac{\partial\bar{V}(x,t)}{\partial x}+\frac{h_x}{2}\frac{\partial^2\bar{V}(x,t)}{\partial x^2}\right), \tag{4.2.21}$$

$$F_1(x,t)=\left(\bar{V}(x,t)+\frac{h_x}{2}2\frac{\partial\bar{V}(x,t)}{\partial x}\right), \tag{4.2.22}$$

$$F_2(x,t)=\left(\frac{h_x}{2}\bar{V}(x,t)\right), \tag{4.2.23}$$

$$G_{0j}=\sqrt{\lambda_j}P_{0j}(x), \tag{4.2.24}$$

$$G_{1j}=\sqrt{\lambda_j}P_{1j}(x), \tag{4.2.25}$$

$$\text{and}\ \ G_{2j}=\sqrt{\lambda_j}P_{2j}(x). \tag{4.2.26}$$

As the stochastic integrals, $I_i(x,t)$ are dependent only on the behaviours of $\bar{V}(x,t)$, the eigen values of the velocity covariance kernel and $P_{ij}(x)$ functions, i.e., $I_i(x,t)$s are only dependent on the velocity fields within the porous media. A corollary to that is if we know the velocity fields and characterize them as stochastic differentials, we can then

develop an empirical SSTM based on equation (4.2.19). We explore the behaviours of $I_i(x,t)$ in section 4.5. Next we discuss the derivation of the generalized eigen function, the form of which is given by equation (4.2.3).

4.3 A Computational Approach for Eigen Functions

We discuss the approach in this section to obtain the eigen functions for any given kernel in the form given by equation (4.2.3). We calculate the covariance kernel matrix (COV) for any given kernel function and COV can be decomposed in to eigen values and the corresponding eigen vectors using singular value decomposition method or principle component analysis. This can easily be done using mathematical software. Then we use the eigen vectors to develop eigen functions using neural networks.

Suppose that we already have an exponential covariance kernel as given by,

$$q(x_1, x_2) = \sigma^2 \; e^{-\frac{y}{b}} \;, \tag{4.3.1}$$

where $y = |x_1 - x_2|$,

b is the correlation length , and

σ^2 is the variance when $x_1 = x_2$.

In terms of x_1 and x_2 , both of them have the domain of [0,L]; we equally divide this range into (n) equidistant intervals of Δx for both variables. Thus, the particular position for x_1 and x_2 can be displayed as:

$$x_{1k} = k\,\Delta x \;, \quad \text{for } k = 0, 1, 2, \cdots, n \text{ , and} \tag{4.3.2}$$

$$x_{2j} = j\,\Delta x \;, \quad \text{for } j = 0, 1, 2, \cdots, n \;. \tag{4.3.3}$$

By substituting equations (4.3.2) and (4.3.3) into equation (4.3.1), we can obtain,

$$COV(k, j) = \sigma^2 \; e^{-\frac{|(k-j)\,\Delta x|}{b}} \quad \text{for } k, j \in [\,1,\; n-1\,]. \tag{4.3.4}$$

In equation (4.3.4), COV is the covariance matrix which contains all the variances and covariances, and it is a symmetric matrix with size $n \times n$ where n is the number of intervals. The diagonals of COV represent variances where x_1 is equal to x_2 and off-diagonals represent covariances between any two different discrete x_1 and x_2 .

After the covariance matrix is defined, we can transform the COV matrix into a new matrix with new scaled variables according to Karhunen-Loève (KL) theorem. In this new matrix, all the variables are independent of each other having their own variances. i.e, the covariance between any two new variables is zero. The new matrix can be represented by using the Karhunen-Loève theorem as,

$$COV1 = \sum_{j=1}^{n} \lambda_j \phi_j(x) \overline{\phi_j(x)} \;, \tag{4.3.5}$$

where n is the total number of variables in the new matrix; λ_j represents the variance of the j^{th} rescaled variable and λ_j is also the j^{th} eigen values of the $COV1$ matrix; and $\phi_j(x)$ is the eigen vectors of the $COV1$ matrix. The number of eigen vectors depends on the number of discrete intervals. This decomposition of the $COV1$ matrix which is called the singular value decomposition method can easily be done by using mathematical or statistical software.

Once we have the eigen vectors, the next step is to develop suitable neural networks to represent or mimic these eigen vectors. In fact, it is not necessary to simulate all the eigen vectors. The number of neural networks is decided by the number of eigen values which are significant in the KL representation. In some situations, for example, we may have 100 eigen values in the KL representation but only 4 significant eigen values. Then, we just create four networks to simulate these four eigenvectors which correspond to the most significant eigen values. The way we decide on the number of significant eigen values is based on the following equation:

$$R[i] = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} \quad \text{for} \quad i \in [1, n], \tag{4.3.6}$$

$$Th = \sum_{i}^{k} R[i], \tag{4.3.7}$$

where $R[i]$ represents the contribution of the i^{th} eigen value in capturing the total original variance, k represents the number of the significant eigen values, and Th is the contribution of all significant eigen values in capturing the total original variance. In this chapter, Th is chosen to vary between 0.95 and 1. This means that if the total number of eigen values is 100 from the KL expansion and the contribution of the first 4 eigen values takes up more than 95% of the original variance, there are only four individual neural networks that need to be developed.

The main factors that need to be decided in the development of neural networks are the number of neurons needed, the structure of neural networks and the learning algorithm. The number of neurons in neural networks is case-dependent. It is difficult to define the number of neurons before the learning stage. In general, the number of neurons is adjusted during training until the network output converges on the actual output based on least square error minimization. A neural network with an optimum number of neurons will reach the desired minimum error level more quickly than other networks with more complex structures. The proposed neural network is a Radial Basis Function (RBF) Network. The approximation function is the Gaussian Function given below:

$$G(s, r) = e^{-r\ (x-s)^2}, \tag{4.3.8}$$

where r and s are constants. In this symmetric function, s defines the centre of symmetry and r defines the sharpness of Gaussian function.

Based on numerical values of the significant eigen vectors, several RBF networks with one input (x) and one output (eigen vector) are developed to approximate each significant

eigenvector. Now let us have the following function the form of which is previously given in equation (4.2.3) to define the RBF network,

$$\phi(x) = g_0 + g_1\, x + \sum_{k=2}^{p+1} g_k\, e^{-r_k\,(x-s_k)2} = g_0 + g_1\, x + \sum_{k=2}^{p+1} g_k\, G(s_k, r_k) \tag{4.3.9}$$

where g_0 is the bias weight of the network, g_1 and g_k is the 1^{st} and k^{th} weight of the network, and p is the total number of neurons in this RBF network (the reason for using p+1 for the summation is that k starts at 2). Figure 4.1 displays the architecture of a neural network for the case of one-dimensional input and $\varphi(x)$ is used to represent the output of the neural network given by equation (4.3.9).

After we decide the input-output mapping and architecture of deterministic neural networks, the next step is to choose the learning algorithm, i.e., the method used to update weights and other parameters of networks. The backpropagation algorithm is used as the learning algorithm in this work. The backpropagation algorithm is used to minimize the network's global error between the actual network outputs and their corresponding desired outputs. The backpropagation leaning method is based on gradient descent that updates weights and other parameters through partial derivatives of the network's global error with respect to the weights and parameters. A stable approach is to change the weights and parameters in the network after the whole sample has been presented and to repeat this process iteratively until the desired minimum error level is reached. This is called batch (or epoch based) learning.

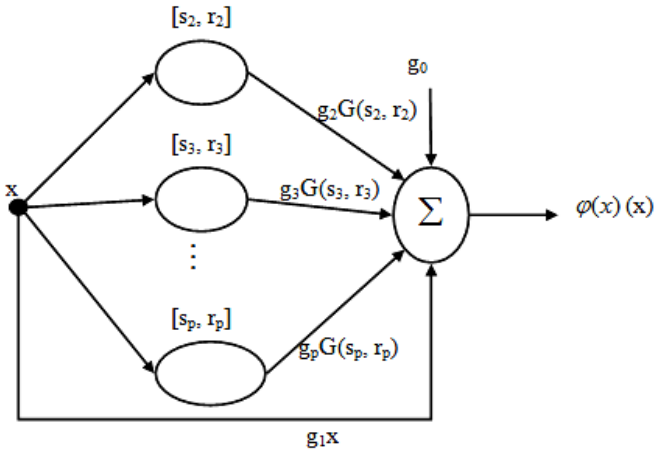


Figure 4.1. Architecture of the one-dimensional RBF network given by equation (4.3.9).

Their values are based on the summation over all training examples of the partial derivative of the network's global error with respect to the weights and parameters in the whole sample.

Now let us assume that the actual network output is T ($\varphi(x)$), the desired network output is Z ($\varphi(x)$). The network's global error between the network output and actual output is

$$E = \frac{1}{2N}\sum_{i}^{N}\left(T_i - Z_i\right)^2 , \tag{4.3.10}$$

where T_i and Z_i are the actual output and the network output for the i^{th} training pattern, and N is the total number of training patterns. The multiplication by $1/2$ is a mathematical convenience (Samarasinghe, 2006).

The method of modifying a weight or a parameter is the same for all weights and parameters so we show the change to an arbitrary weight as an example. The change to a single weight of a connection between neuron j and neuron i in the RBF network based on batch learning can be defined as,

$$\Delta g_{ji} = \eta \sum_{p=1}^{k}\left(\frac{dE}{dg_{ji}}\right)_p , \tag{4.3.11}$$

where η is called the learning rate with a constant value between 0 and 1. It controls the step size and the speed of weight adjustments. k is the total number of input vectors. The process that propagates the error information backwards into the network and updates weights and the parameters of network is repeated until the network minimizes the global error between the actual network outputs and their corresponding desired outputs. In the learning process, the weights and the parameters of the network converge on the optimal values.

To illustrate the computational approach, we give some examples here. The first covariance kernel is chosen to be

$$g(x_1, x_2) = \sigma^2 \, e^{-\frac{y}{b}} , \tag{4.3.12}$$

where $y = \left|x_1 - x_2\right|^2$,

b is constant , and

σ^2 is variance when $x_1 = x_2$.

This covariance kernel needs a relatively lower number of significant eigen values to capture 95% or more of the total variance; therefore we choose to work with this kernel and later we use two other forms of kernels: one discussed previously in Chapter 3 and other one is empirically based.

Figure 4.2 displays the covariance matrix based on the covariance kernel (equation (4.3.12)) when $\sigma^2 = 1$ and $b = 0.1$.

Table 4.1 reports all eigen values in the KL representation of the covariance matrix. The most of the eigen values is equal to zero and these eigen values can not affect the covariance matrix and just a few are significant and capture the total variance in the original data.

Therefore, we need to focus on the significant eigen values as well as their corresponding eigen functions. There are 6 significant eigen values whose contribution takes up 99.9035% of the original variance and table 4.1 shows the value of each significant eigen value and the proportion of variance captured by the corresponding eigen value.

Thus, the six eigenvectors corresponding to the significant eigenvalues are simulated by the individual RBF network. Although we use a different RBF network to approximate each of these six eigenvectors, the structure of RBF network is the same but the weights and parameters inside the individual RBF network are different.

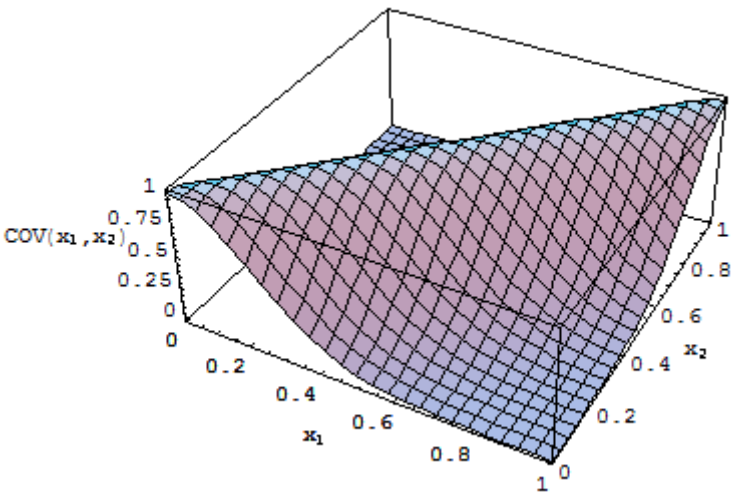


Figure 4.2. The covariance matrix calculated by the given covariance kernel (equation 4.3.12) when $\sigma^2 = 1 \text{ and } b = 0.1$.

The number of eigen values	Values	Contribution as a propotion
λ_1	48.128	0.477
λ_2	30.636	0.303
λ_3	14.692	0.145
λ_4	5.446	0.054
λ_5	1.61	0.016
λ_6	0.392	0.004

Table 4.1. Significant eigen values obtained from the KL expansion of the covariance kernel given by equation (4.3.12) (six significant eigen values take up 99.9035% of total variance)

Then each individual RBF network was trained to learn their related eigen vector, while the weights and all the parameters of Gaussian functions were updated until each network reaches global minimum error given by

$$\mathrm{E} = \frac{1}{N-1}\sum_{n=1}^{N}\left[\varphi_n^i(x) - S_n^i(x)\right]^2 . \tag{4.3.13}$$

where $S_n^i(x)$ is an approximation to $\varphi_n^i(x)$, i th eigen function; and N is the total number of eigen values.

Figure 4.3 displays all the six eigen functions and Table 4.2 gives their functional forms. Figure 4.3 shows the eigen functions given by the KL theory (dots), obtained by solving the corresponding integral equation, overlaid with the outputs from the neural networks (lines), and the approximation functions are the same as the theoretically derived functions.

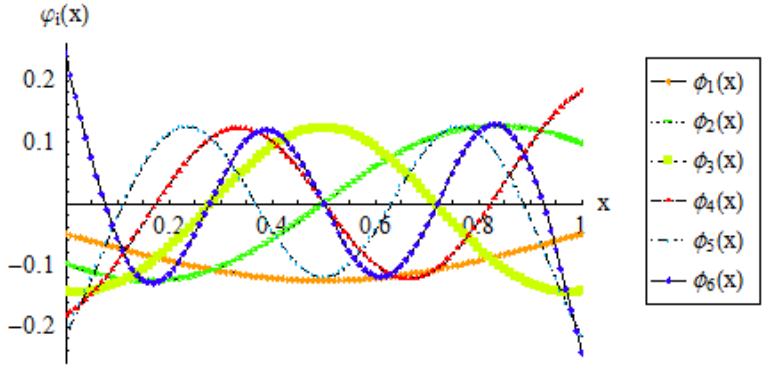


Figure 4.3. The approximated six eigen functions from BRF networks for equation (4.3.12) when $\sigma^2 = 1$ and b=0.1.

For the second example, we use the same covariance kernel with $b = 0.2$ and $\sigma^2 = 1$, and decomposed the matrix for the domain [0, 1]. Figure 4.4 shows the covariance matrix calculated by the given covariance kernel. Based on the standard of choosing the number of significant eigen values, it can be seen that five significant eigen values together capture 99.9582% of the original variance. Thus, there are five RBF networks to be developed for the corresponding eigenvectors. Table 4.3 gives the value of each significant eigen value and the proportion of variance captured by the corresponding eigen value; Table 4.4 provides the eigen functions obtained from the neural networks; the graphical forms of the eigen functions are given by Figure 4.5. As before, the analytically derived eigen function as an very well approximated by the approximations from the neural networks. (The theoretical eigen functions are not displayed in Figure 4.5).

Eigen values	Values	The analytical forms of eigen functions $(\phi_i(x))$ for $x \in [0,1]$
λ_1	48.128	$0.0456723 - 1.61025 \times 10^{-12}\ x - 0.170974\ e^{-2.34049\,(x-0.5)^2}$
λ_2	30.636	$0.689211 - 0.560102\ x - 0.360116\ e^{-5.37562\,(x-0.290915)^2}$ $-0.745057\ e^{-2.56631\,(x+0.335246)^2}$
λ_3	14.692	$-0.00132933 + 0.0185462\ x - 3.7566\ e^{-6.62157\,(x-0.703168)^2}$ $+3.7712\ e^{-6.6592\,(x-0.683601)^2} - 0.183884\ e^{-9.56556\,(x-0.0999343)^2}$
λ_4	5.446	$10.4236 - 5.90955\ x - 0.3771\ e^{-13.4883\,(x-0.68965)^2}$ $+0.282871\ e^{-14.6016\,(x-0.306358)^2} - 12.8992\ e^{-0.385895\,(x+0.699959)^2}$
λ_5	1.61	$0.0403405 - 3.13566 \times 10^{-14}\ x + 9.47326\ e^{-8.95446\,(x-0.727565)^2}$ $+19.0735\ e^{-8.41938\,(x-0.5)^2} - 31.1501\ e^{-5.67596\,(x-0.5)^2}$ $+9.47326\ e^{-8.95446\,(x-0.272435)^2}$
λ_6	0.392	$-0.318516 + 0.548883\ x + 127.699\ e^{-7.08944\,(x-0.658744)^2}$ $-429.099\ e^{-5.57986\,(x-0.511944)^2} + 237.989\ e^{-5.3708\,(x-0.46363)^2}$ $+88.8399\ e^{-8.18627\,(x-0.433595)^2}$

Table 4.2. The final formula from the developed RBF networks to approximate each significant eigen vector and their corresponding eigen values.

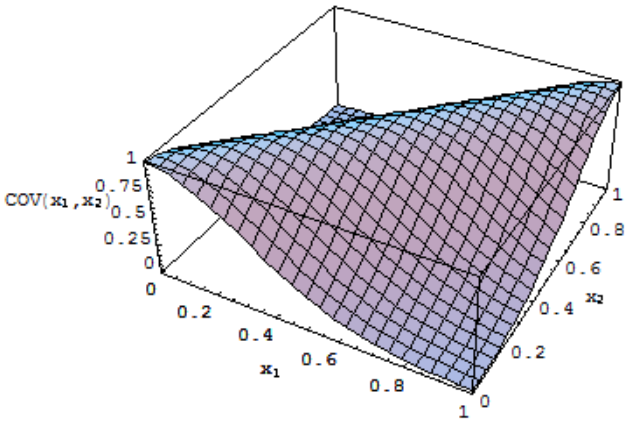


Figure 4.4. The covariance matrix of equation (4.3.12) when $\sigma^2 = 1$ and b = 0.2.

Eigen values	Values	Contribution as a proportion
λ_1	61.242	0.606
λ_2	28.820	0.285
λ_3	8.747	0.087
λ_4	1.853	0.018
λ_5	0.29597	0.00293

Table 4.3. Relative amounts of variance in the data captured by each significant eigenvalue obtained from the KL expansion of equation (4.3.12) when $\sigma^2 = 1$ and b=0.2.

Eigen values	Values	The analytical forms of eigen functions $(\phi_i(x))$ for $x \in [0,1]$
λ_1	61.242	$0.0179462 - 7.11262 \times 10^{-17}\, x - 0.137587\, e^{-2.18805\ (x-0.5)^2}$
λ_2	28.820	$0.000620869 - 0.00124174\, x - 0.147148\, e^{-4.39837\ (x-0.820696)^2}$ $-0.147148\, e^{-4.39837\,(x-0.179304)^2}$
λ_3	8.747	$-0.201764 - 3.67217 \times 10^{-8}\, x + 0.197208\, e^{-13.4824\ (x-0.625896)^2}$ $+0.197207\, e^{-13.4824\,(x-0.374103)^2}$
λ_4	1.853	$0.0357425 - 0.0604364\, x - 176.588\, e^{-3.81696\,(x--0.559606)^2}$ $+231.215\, e^{-3.70961\,(x-0.536422)^2} - 56.2021\, e^{-3.5887\,(x-0.459753)^2}$
λ_5	0.29597	$-0.195412 - 0.641926\, x + 0.834518\, e^{-11.2297\,(x-0.826432)^2}$ $+0.479846\, e^{-14.0892\,(x-0.212252)^2} - 0.649793\, e^{-14.9973\ (x+0.226353)^2}$

Table 4.4. The eigen functions from the developed RBF networks to approximate each significant eigen vector for the kernel given by equation (4.3.12) when $\sigma^2 = 1$ and b=0.2.

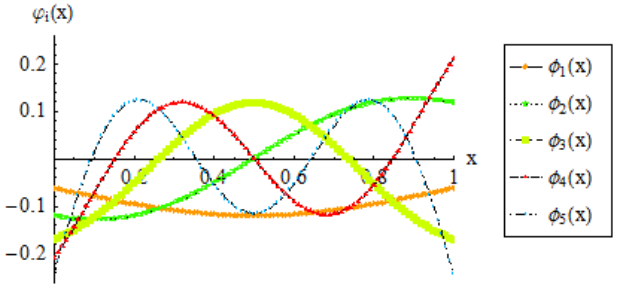


Figure 4.5. The approximated six eigen functions from BRF networks for equation (4.3.12) when $\sigma^2 = 1$ and $b = 0.2$.

From the previous two examples, we have seen that the covariance kernel given by equation (4.3.12) provides a relative small number of eigen functions and therefore one may say that the kernel given by equation (4.3.12) has fast convergence. This is quite a desirable property to have, especially in terms of computational efficiency of the algorithms. In the next example, we find the eigen values and the eigen functions of the covariance kernel we use in

the development of the SSTM in Chapter 3. In Chapter 3 the covariance kernel given by equation (4.3.14) - we reproduce the equation here- constitutes an integral equation which we solve analytically to obtain eigen values and eigen functions:

$$g(x_1, x_2) = \sigma^2 \, e^{-\frac{-|x_1 - x_2|}{b}}, \tag{4.3.14}$$

when σ^2 and b have the same meanings as before.

This covariance kernel is depicted graphically in Figure 4.6.

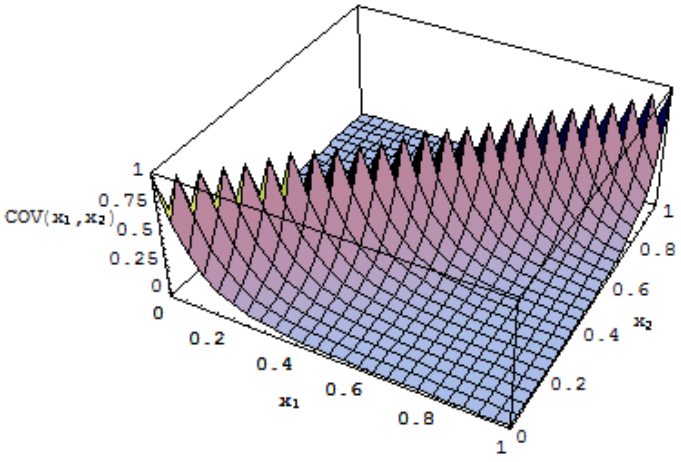


Figure 4.6. The covariance matrix calculated by the given equation (4.3.14) under the condition $\sigma^2 = 1$.

Eigen values	Values	Contribution as a proportion
λ_1	18.745	0.186
λ_2	15.681	0.155
λ_3	12.218	0.121
λ_4	9.241	0.091
λ_5	6.976	0.069
λ_6	5.332	0.053
λ_7	4.149	0.041
λ_8	3.293	0.033

λ_9	2.661	0.026
λ_{10}	2.188	0.022
λ_{11}	1.826	0.018
λ_{12}	1.545	0.015
λ_{13}	1.323	0.013
λ_{14}	1.145	0.011
λ_{15}	0.9999	0.0099
λ_{16}	0.881	0.0087
λ_{17}	0.782	0.0077
λ_{18}	0.699	0.0069
λ_{19}	0.628	0.0062
λ_{20}	0.568	0.0056
λ_{21}	0.516	0.0051
λ_{22}	0.471	0.0047
λ_{23}	0.432	0.0043
λ_{24}	0.398	0.0039
λ_{25}	0.367	0.0036
λ_{26}	0.341	0.0034
λ_{27}	0.317	0.0031
λ_{28}	0.295	0.0029
λ_{29}	0.276	0.0027
λ_{30}	0.259	0.0026
λ_{31}	0.244	0.0024
λ_{32}	0.229	0.0023

Table 4.5. The eigen values for the kernel given by equation (4.3.14) (32 significant eigen values which take up 94.29% of original variance)

In Table 4.5, we give the 32 eigen values which capture up to 94% of the only original variance; Table 4.6 gives only the first six eigen functions obtained from the networks for brevity. Figure 4.7 shows the first eight eigen functions.

Eigen values	Values	The analytical forms of eigen functions $(\phi_i(x))$ for $x \in [0,1]$
λ_1	18.745	$0.238941 + 7.92548 \times 10^{-16}\ x - 0.368191\ e^{-1.19444\ (x-0.5)^2}$
λ_2	15.681	$0.128953\ - 0.257905\ x + 0.23274\ e^{-5.79605\ (x-0.830952)^2}$ $-0.23274\ e^{-5.79605\ (x-0.169048)^2}$
λ_3	12.218	$0.297969\ + 0.071349\ x - 1.99033\ e^{-5.41338\ (x-0.640904)^2}$ $+1.74606\ e^{-7.13097\ (x-0.592419)^2}\ - 0.339719\ e^{-10.6256\ (x-0.0883508)^2}$
λ_4	9.241	$-0.520716\ + 0.848337\ x - 182.653\ e^{-5.89655\ (x-0.636135)^2}$ $+214.465\ e^{-5.66061\ (x-0.603547)^2}\ - 38.6999\ e^{-6.81302\ (x-0.448057)^2}$
λ_5	6.976	$-1.10289\ + 1.12328 \times 10^{10}\ x - 55.2051\ e^{-9.19718\ (x-0.611767)^2}$ $+99.6664\ e^{-7.12542\ (x-0.5)^2}\ - 55.2051\ e^{-9.19718\ (x-0.388233)^2}$
λ_6	5.332	$822.58\ - 496.01\ x - 6.74448\ e^{-17.354\ (x-0.590219)^2}$ $-190.785\ e^{-4.883\ (x-0.215442)^2}\ + 2539.25\ e^{-1.85592\ (x-0.0783536)^2}$ $-3181.28\ e^{-1.30206\ (x+0.0105042)^2}$
λ_7	4.149	$-2.16846\ + 0.148883\ x + 1451.69\ e^{-14.2086\ (x-0.592557)^2}$ $-2198.81\ e^{-13.3089\ (x-0.577832)^2}\ + 1013.85\ e^{-10.1574\ (x-0.488994)^2}$ $-358.254\ e^{-13.1355\ (x-0.350845)^2}$
λ_8	3.293	$-31.8294 - 108.195\ x + 6255.59\ e^{-9.37148\ (x-0.813402)^2}$ $-7941.31\ e^{-8.63616\ (x-0.807765)^2}\ + 1838.53\ e^{-5.03775\ (x-0.766933)^2}$ $-282.847\ e^{-13.5694\ (x-0.3499)^2}\ + 52.6999\ e^{-22.3578\ (x-0.307475)^2}$

Table 4.6. The final formula from the developed RBF networks to approximate the first 8 significant eigen vectors and their corresponding eigen values.

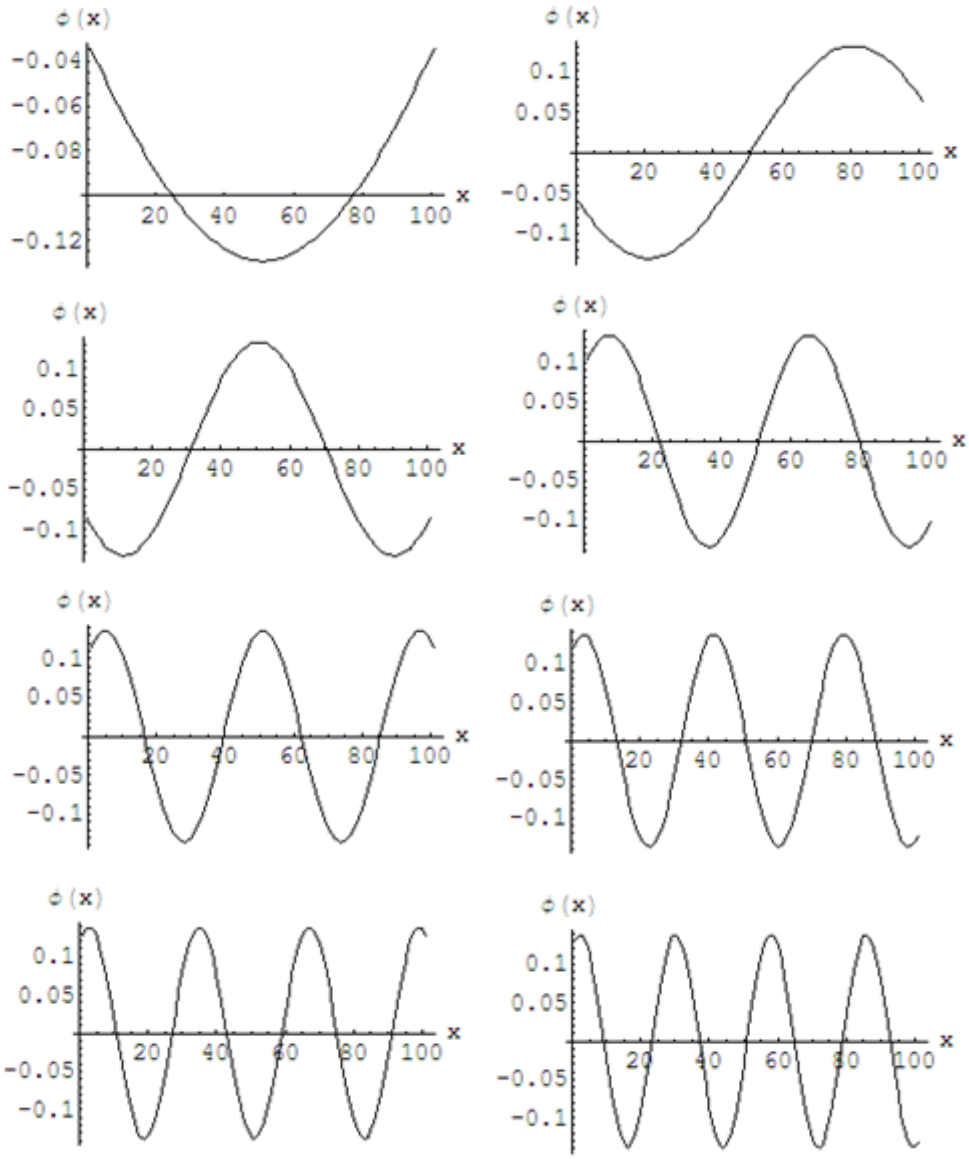


Figure 4.7. The approximated first eight eigen functions from RBF network for the equation (4.3.14) when $\sigma^2 = 1$ and $b = 0.1$.

We can also use an empirically derived covariance kernel. As an example, let us consider equation (4.3.15).

$$Cov(x_1,x_2)=\begin{cases} -|x_1-x_2|+1 & for \quad 0\le|x_1-x_2|\le 0.5 \\ -\frac{5}{3}(|x_1-x_2|-0.8) & for \quad 0.5<|x_1-x_2|\le 0.8 \\ 0 & for \quad 0.8<|x_1-x_2|\le 1 \end{cases}, \tag{4.3.15}$$

Equation (4.3.15) is depicted in Figure 4.8, and Figure 4.9 shows the corresponding covariance matrix.

Table 4.7 gives the most significant eigen values (the first nine values); Table 4.8 shows the functional forms of eigen functions and Figure 4.10 shows the graphical forms of eigen functions.

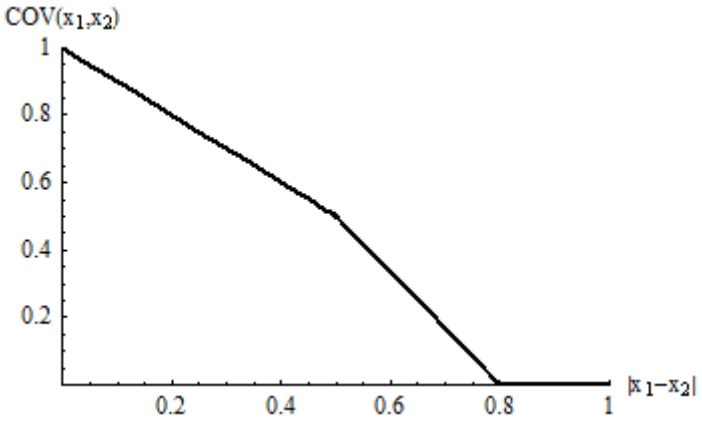


Figure 4.8. An empirical distribution.

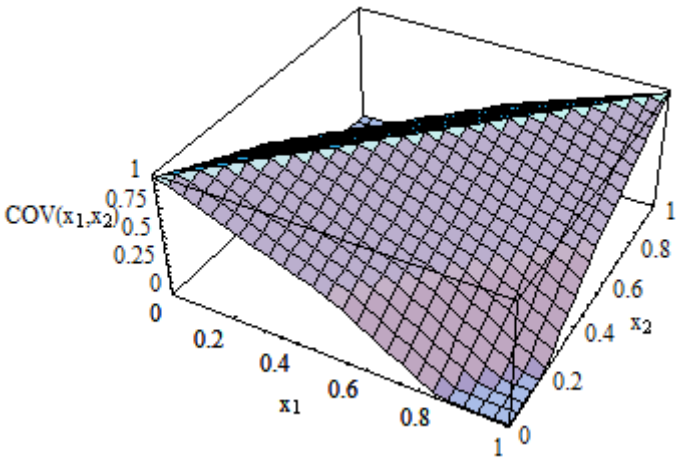


Figure 4.9. The covariance matrix given by the empirical distribution (equation (4.3.15)).

Eigen values	Values	Contribution as a proportion
λ_1	66.022	0.6537
λ_2	24.216	0.2398
λ_3	3.163	0.0313
λ_4	1.989	0.0197
λ_5	0.447	0.0189
λ_6	0.445	0.00443
λ_7	0.395	0.0044
λ_8	0.390	0.00391
λ_9	0.201	0.00386

Table 4.7. Relative amount of variance in the data captured by each significant eigen value for the kernel in equation (4.3.15). (9 significant eigen values capture 98% of original variance)

Eigen values	Values	The analytical forms of eigen functions $(\phi_i(x))$ for $x \in [0,1]$
λ_1	66.022	$0.0495323 + 3.28499\times10^{-16}\, x - 0.165441\, e^{-1.38223\,(x-0.5)^2}$
λ_2	24.216	$-0.0156028 + 0.0312062\, x + 0.127272\, e^{-8.33766\,(x-0.951692)^2}$ $-0.127273\, e^{-8.33761\,(x-0.0483079)^2}$
λ_3	3.163	$-672.733 + 332.866\, x + 1.19376\times10^{7}\, e^{-0.810094\,(x+0.374112)^2}$ $-1.60589\times10^{7}\, e^{-0.807196\,(x+0.375468)^2}$ $+4.12218\times10^{6}\, e^{-0.798664\,(x+0.379503)^2}$
λ_4	1.989	$93.5157 - 68.0722\, x - 813.179\, e^{-3.71071\,(x-0.0967842)^2}$ $+1577.84\, e^{-3.15124\,(x-0.0721231)^2} - 860.897\, e^{-2.33563\,(x-0.0182454)^2}$
λ_5	0.447	$4.82583 - 4.6287\, x - 1.32608\, e^{-14.504\,(x-0.624954)^2}$ $-4.73704\, e^{-4.31785\,(x+0.0230752)^2} - 0.196755\, e^{-174.457\,(x+0.074625)^2}$
λ_6	0.445	$-9.87408\times10^{7} + 3.76629\times10^{6}\, x$ $+1.39729\times10^{6}\, e^{-2.78874\,(x-1.40277)^2}$ $+4.53545\times10^{7}\, e^{-1.17605\,(x-0.688669)^2}$ $-3.41918\times10^{8}\, e^{-0.360088\,(x-0.448068)^2}$ $+3.84429\times10^{8}\, e^{-0.206739\,(x-0.405528)^2}$ $+1.13264\times10^{7}\, e^{-2.08493\,(x-0.139661)^2}$ $+1.06435\times10^{7}\, e^{-2.10254\,(x+0.336545)^2}$

λ_7	0.395	$-2.02737 - 7.9356\,x - 5263.75\,e^{-27.8868\,(x-0.860872)^2}$ $+1.03168\times10^{7}\,e^{-25.9199\,(x-0.765634)^2}$ $-7.83776\times10^{7}\,e^{-26.0007\,(x-0.764335)^2}$ $+6.80653\times10^{7}\,e^{-26.0126\,(x-0.764145)^2}$ $+4.26386\,e^{-4.50902\,(x-0.43218)^2} + 0.388207\,e^{-183.561\,(x-0.100811)^2}$
λ_8	0.390	$-386986 + 334630\,x + 1.13814\times10^{8}\,e^{-4.2926\,(x-0.649682)^2}$ $-1.52714\times10^{8}\,e^{-3.99793\,(x-0.630907)^2}$ $+8.57437\times10^{7}\,e^{-2.777\,(x-0.591124)^2} + 702189\,e^{-7.32322\,(x-0.57722)^2}$ $-5.31436\times10^{7}\,e^{-2.43237\,(x-0.572872)^2}$ $+6.4929\times10^{6}\,e^{-5.69446\,(x-0.396187)^2}$ $+1.62219\times10^{6}\,e^{-1.15913\,(x-0.0920272)^2}$
λ_9	0.201	$6589.2 + 102.051\,x - 277640\,e^{-8.1404\,(x-0.811728)^2}$ $-151130\,e^{-3.04348\,(x-0.528018)^2} + 1.6007\times10^{7}\,e^{-5.70678\,(x-0.480254)^2}$ $-1.67931\times10^{7}\,e^{-5.91597\,(x-0.47122)^2}$ $+2.8962\times10^{6}\,e^{-8.42978\,(x-0.395813)^2}$ $-1.79879\times10^{6}\,e^{-8.86148\,(x-0.388545)^2}$ $-25798.1\,e^{-3.56779\,(x-0.288557)^2}$

Table 4.8. The final formulas from the developed RBF networks to approximate each significant eigenvector and their corresponding eigen values.

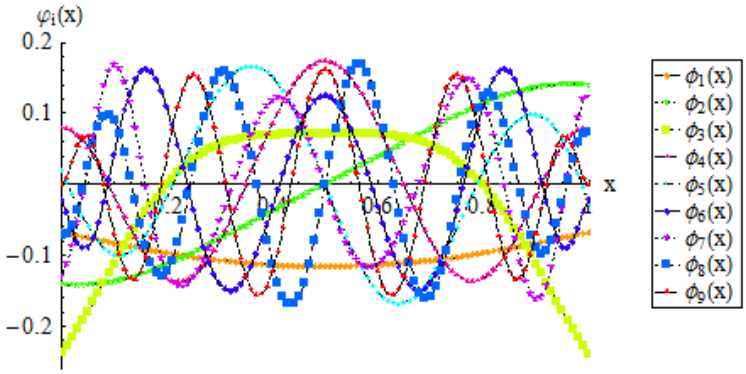


Figure 4.10. The approximated nine eigenfunctions from the BRF networks for the kernel given in equation (4.3.15).

We have seen that some covariance kernels provide relatively small number of significant eigen values for the given domain [0,1] and whereas for others, obtaining the significant

eigen functions could be tedious. The main point in this exercise is to show that any given covariance kernel can be used to obtain the eigen functions of the form given by equation (4.2.3). A corollary to this statement is that if we express the eigen functions in the form given by equation (4.2.3), then we can assume that these is an underlying covariance kernel responsible for these eigen functions. In deriving the SSTM in the form of equation (4.2.15), we assume that the form of equation (4.2.3) is given. But we see now that any covariance-kernel driven SSTM can be represented by equation (4.2.15).

4.4 Effects of Different Kernels and h_x

We have seen in sections 4.2 and 4.3 that the SSTM developed in Chapter 2 and 3 can be recasted so that we could employ any given velocity kernel. In fact, we can even use an empirical set of data for the velocity covariance. We can anticipate that the generalized SSTM would behave quite similar to the one developed in Chapter 2 given that same covariance kernel is used. We compute the 95% confidence intervals for the concentration breakthrough curves (concentration realizations) at x = 0.5 m when the flow length is 1 m to compare the differences that occur in using different kernels in the generalized SSTM. The mean velocity is kept constant at 0.5 m/day, and the covariance kernel given by equation (4.3.14) is used to obtain Figure 4.11, and Figure 4.12 is obtained by employing the kernel, $\sigma^2 e^{-\frac{-(x_1-x_2)^2}{b}}$. First, the confidence intervals shown in Figure 4.11 are very similar to those ones could obtain by using the SSTM developed in Chapter 2. Comparing the effects of the kernels on the behaviours of the generalized SSTM, we see that the confidence interval bandwidth in Figure 4.12 is almost non-existent. The reason is that the kernel used has a faster convergence when decomposed in the eigen vector space. For smaller values of σ, the randomness in the concentration realization are minimal but as σ^2 is increased, we see increased randomness in the realizations. This also allows us to use the kernel used in Figure 4.12 for larger scale computations. We conclude that the choice of the velocity covariance kernel has a significant effect on the behaviour of the generalized SSTM increasing the flexibility of the SSTM.

σ^2	b	The Kernel 1 $Cov(x_1,x_2)=\sigma^2 e^{-\frac{\lvert x_1-x_2\rvert}{b}}$		The Kernel 2 $Cov(x_1,x_2)=\sigma^2 e^{-\frac{(x_1-x_2)^2}{b}}$	
		D_L	α	D_L	α
0.0001	0.1	0.02305	0.04611	0.02460	0.04921
0.001	0.1	0.02699	0.05199	0.02513	0.05025
0.01	0.1	0.03059	0.06117	0.02782	0.05361
0.1	0.1	0.06852	0.13705	0.06407	0.12815

Table 4.9. Comparison of the dispersivity values for the two kernels.

We investigate the effects of the kernels on the dispersivity values; we compute them using the stochastic inverse method (SIM) discussed in Chapter 3. Table 4.9 shows the results. For all practical purposes they are essentially the same. The mechanic of dispersion is more influenced by σ^2 for a given b or if both σ^2 and b are allowed to vary, on both σ^2 and b. The mechanics of dispersion in general can also be assumed to be influenced by the mathematical form of the kernel. The both of these kernels are exponential decaying functions. Because of the case of computations, we continue to use kernel 2 in Table 4.9 in the most of the work discussed in this book.

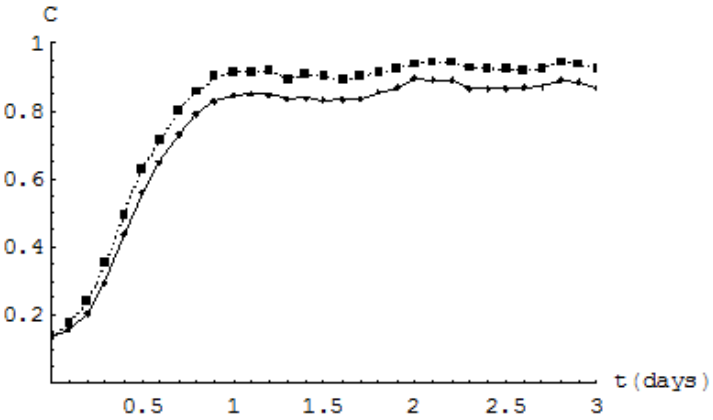


Figure 4.11. The generalized SSTM 95% confidence intervals for the concentration realization for the kernel, $\sigma^2 e^{-\frac{-|x_1-x_2|}{b}}$ when $\sigma^2 = 0.1$ and $b = 0.1$.

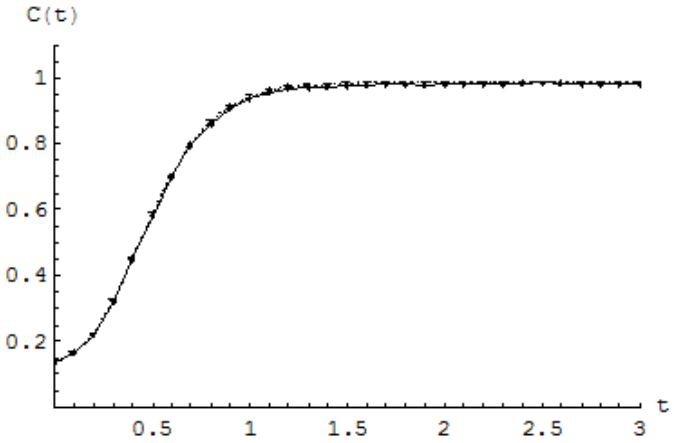


Figure 4.12. The generalized SSTM 95% confidence intervals for the concentration realization for the kernel, $\sigma^2 e^{-\frac{-(x_1-x_2)^2}{b}}$ when $\sigma^2 = 0.1$ and $b = 0.1$.

4.5 Analysis of Ito Diffusions $I_i(x,t)$

We have developed the Ito diffusions $I_i(x,t)$ in section 4.2 (see equation (4.2.20)), and we rewrite the diffusions in the differential form:

$$dI_i(x,t) = F_i(x,t)dt + \sigma\sum_{j=1}^{m} G_{ij}(x,t)db_j(t), \quad \text{for } (i=0,1,2) \text{ and } (j=1,...,m). \tag{4.5.1}$$

In equation (4.5.1), $F_i(x,t)$ are given by equations (4.2.21), (4.2.22) and (4.2.23), which can be considered as regular continuous and differentiable functions of x and t because we assumed the mean velocity $(\overline{V}(x,t))$ to be a continuous, differentiable function with finite variation in the development of the SSTM.

For many situations, we can assume $(\overline{V}(x,t))$ to be a function of x alone, and some regions of x it can be considered as a constant. G_{ij}'s are continuous differentiable functions of x only, with finite variation with respect to x (see equations (4.2.24) to (4.2.26)). Therefore, for a fixed value of x, we can write equation (4.5.1) as,

$$dI_{x,i}(t) = F_{x,i}(t)dt + \sigma\sum_{j=1}^{m} G_{x,ij}db_j(t), \quad (i=0,1,2) \text{ and } (j=1,...,m). \tag{4.5.2}$$

From the derivation of equation (4.2.19), we see that the $db_j(t)s$ are the same standard Wiener process increments for each $I_{x,i}$ and each $db_j(t)$: equation (4.5.2) is a diffusion in m dimensions, and we can write $\underline{I}_x(t)$ as a multidimensional stochastic differential equation (SDE) (Klebaner, 1998).

In coordinate form we can write,

$$dI_{x,i}(t) = F_{x,i}(t)dt + \sigma\begin{bmatrix} G_{x,i1} & G_{x,i2} & .... & G_{x,im} \end{bmatrix}\begin{bmatrix} db_1(t) \\ db_2(t) \\ . \\ . \\ db_m(t) \end{bmatrix}. \tag{4.5.3}$$

In matrix form, we can write,

$$d\underline{I}_x(t) = \underline{F}_x dt + \sigma\underline{G}_x dB(t) \tag{4.5.4}$$

when, $\underline{G}_x = \begin{bmatrix} G_{x,01} & G_{x,02} & .... & G_{x,0m} \\ G_{x,11} & G_{x,12} & .... & G_{x,1m} \\ G_{x,21} & G_{x,22} & .... & G_{x,2m} \end{bmatrix}_{3\times m}$,

$$d\underline{B} = \begin{bmatrix} db_1(t) \\ db_2(t) \\ . \\ . \\ db_m(t) \end{bmatrix}_{m\times 1},$$

$$\underline{F}_x = \begin{bmatrix} F_{x,0} \\ F_{x,1} \\ F_{x,2} \end{bmatrix}.$$

The drift and diffusion coefficients of the multi-dimensional SDEs are vector $\underline{F}_x$ and the matrix $\sigma\underline{G}_x$ is independent of t, and the drift coefficient $\underline{F}_x$ can also be assumed to be independent of t in many cases. Therefore, equation (4.2.4) is a linear multi-dimensional SDE. The associated with this SDE, the matrix $\underline{a}$ called the diffusion matrix can be defined,

$$\underline{a} = (\sigma\underline{G}_x)(\sigma\underline{G}_x)^T \tag{4.5.5}$$

when superscript T indicates the transposed matrix. Under the conditions such as the coefficients use locally Lipschitz, equation (4.5.4) has strong solutions (see Theorem 6.22 in Klebaner (1998)).

The diffusion matrix, $\underline{a}$, is important to obtain the covariation of $\underline{I}_x$:

$$d\left[I_i, I_j\right](t) = dI_i(t)\,dI_j(t) = a_{ij}dt, \tag{4.5.6}$$

when, $a_{ij} = \begin{cases} \sigma^2 \sum_{K=1}^{m} G_{x,ik}^2 & \text{if } i = j \\ \sigma^2 \sum_{K=1}^{m} G_{x,ik}G_{x,jk} & \text{if } i \neq j \end{cases}$, $\quad i = 0,1,2, \text{ and } j = 0,1,2.$

$\underline{a}$ is a symmetric matrix.

In obtaining equation (4.2.6), we employ the fact that independent Brownian motions have quadratic covariation. The most important use of quadratic covariation is that we can determine the movement of $\underline{I}_x(t)$ with respect to time using the following well known results for Ito diffusions, which are also continuous Markov processes. It can be shown that, for an infinitesimal time increment,

$$E\left(I_{x,i}(t+\Delta) - I_{x,i}(t)\right) = F_{x,i}\Delta + \mathrm{O}(\Delta), \tag{4.5.7}$$

$$E\left\{\left(I_{x,i}(t+\Delta) - I_{x,i}(t)\right)\left(I_{x,j}(t+\Delta) - I_{x,j}(t)\right)\middle| I_{x,i}(t)\right\} = a_{ij}\Delta + \mathrm{O}(\Delta), \tag{4.5.8}$$

and when $i = j$, equation (4.6.8) becomes,

$$E\left\{\left(I_{x,i}(t+\Delta)-I_{x,i}(t)\right)^2 \middle| I_{x,i}(t)\right\} = a_{ii}\Delta + \mathrm{O}(\Delta), \tag{4.5.9}$$

As the solution to the SDE equation (4.5.4) exists, from equation (4.5.8) it is seen that $\underline{F}_x$ is the form of $\underline{I}_x$ at time t, and $\underline{a}$ is the coefficients in the covariance of the infinitesimal displacement from $\underline{I}_x$. Using those results we can construct the realizations of $\underline{I}_x(t)$ by using the fact that $\underline{I}_x(t)$ are Gaussian processes. By dividing a given time interval into equidistant infinitesimal time interval, Δ, we can generate normally distributed $d\underline{I}_x$ increments for a given x, using the mean and variance obtained by equations (4.5.7) and (4.5.9). It should be noted that in generating the standard Wiener process increments, we use the zero-mean and Δ -variance Gaussian increments.

We take $\underline{I}_x(t) = \underline{0}$ when $t = 0$ because

$$\underline{I}_x(t) = \int_0^t \underline{F}_x(t)dt + \sigma\int_0^t \underline{G}_x(t)d\underline{B}(t). \tag{4.5.10}$$

Figure 4.13 shows some realizations of $\underline{I}_x(t)$ when $x = 0.5$.

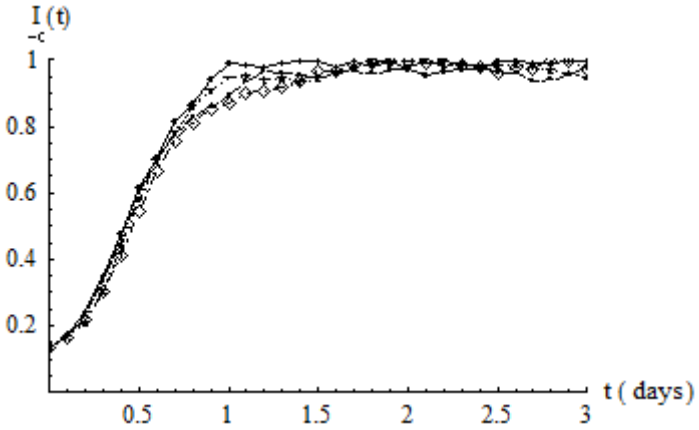


Figure 4.13. Some realizations of $\underline{I}_x(t)$ when $x = 0.5$.

The increments of $\underline{I}_x(t)$ are Gaussian random variables having the mean and variance given by equations (4.5.7) and (4.5.8).

The SDE given in equation (4.5.4) is linear and strong solutions do exist. When $F_{x,i}$ are not functions of t, then the solution of equation (4.5.4) is given by

$$\underline{I}_x(t) = \underline{F}_x t + \sigma\underline{G}_x\underline{B}(t). \tag{4.5.11}$$

Equation (4.5.11) provides realizations which have the statistical properties of the realizations depicted in Figure 4.13. The vector $\underline{B}(t)$ consists of independent Wiener processes; Figure 4.14 shows some realizations based on equation (4.5.11).

The statistical properties of these realizations are essentially the same to those of the realizations given in Figure 4.13.

As we have mentioned earlier, $\underline{I}_x(t)$ are only dependent on the velocity patterns in the medium, and it is important ask the question how the correlation length, b, affects the realization of $\underline{I}_x(t)$. (In this discussion we focus on the kernel $\sigma^2 e^{\frac{-|x_1-x_2|^2}{b}}$ only). It is seen that σ (the square root of the variance of the kernel) acts as the multiplication factor to the diffusion form of the SDE given in equation (4.5.11). However, the correlation length b influence $\underline{I}_x(t)$ nonlinearly through $\underline{G}_x$, but this influence can always be captured by suitable changes in σ. Therefore, we can keep b at a constant value that is appropriate for the porous medium under study. We found that b =0.1 is suitable for our computational experiments in this chapter as well as in chapters 6, 7 and 8.

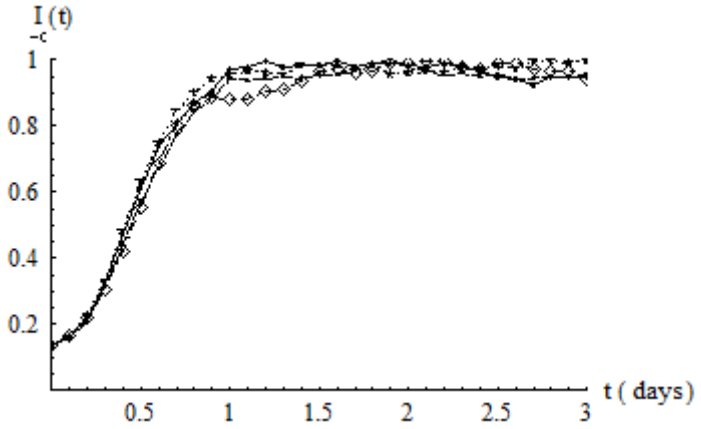


Figure 4.14. Some realization of $\underline{I}_x(t)$ for fixed $F_{x,i}s$ when $x = 0.5$ based on equation (4.5.11).

Equation (4.5.11) can be written in component forms,

$$I_{x,0}(t) = F_{x,0}(t) + \sigma \sum_{j=1}^{m} G_{x,0_j} B_j(t), \tag{4.5.12a}$$

$$I_{x,1}(t) = F_{x,1}(t) + \sigma \sum_{j=1}^{m} G_{x,1_j} B_j(t), \tag{4.5.12b}$$

$$\text{and} \quad I_{x,2}(t) = F_{x,2}(t) + \sigma \sum_{j=1}^{m} G_{x,2j} B_j(t). \tag{4.5.12c}$$

In the above equations, m is the number of significant eigen functions and is dependent on b. From equation (4.2.24), (4.2.25) and (4.2.26), we see that G_{ij} are related to $P_{ij}(x)$ which are given by equations (4.2.10), (4.2.11) and (4.2.12). For $b = 0.05$, $m = 8$, Figure 4.15 give $P_{ij}(x)s$ when $0.0 \le x \le 1.0$. The following observations can be noted from Figure 4.15: (a) P_{ij}s, therefore $G_{ij}s$ are sinesodial in nature; (b) amplitutes of $P_{ij}s$ increase with m (eigen function number) but as eigen values decrease with m, G_{ij}s diminish with m (not shown); (c) $P_{2j}s$ are insignificant in comparision to P_{0j}s and $P_{1j}s$ and therefore could be ignored; and (d) frequency of P_{ij} functions increases as m increase.

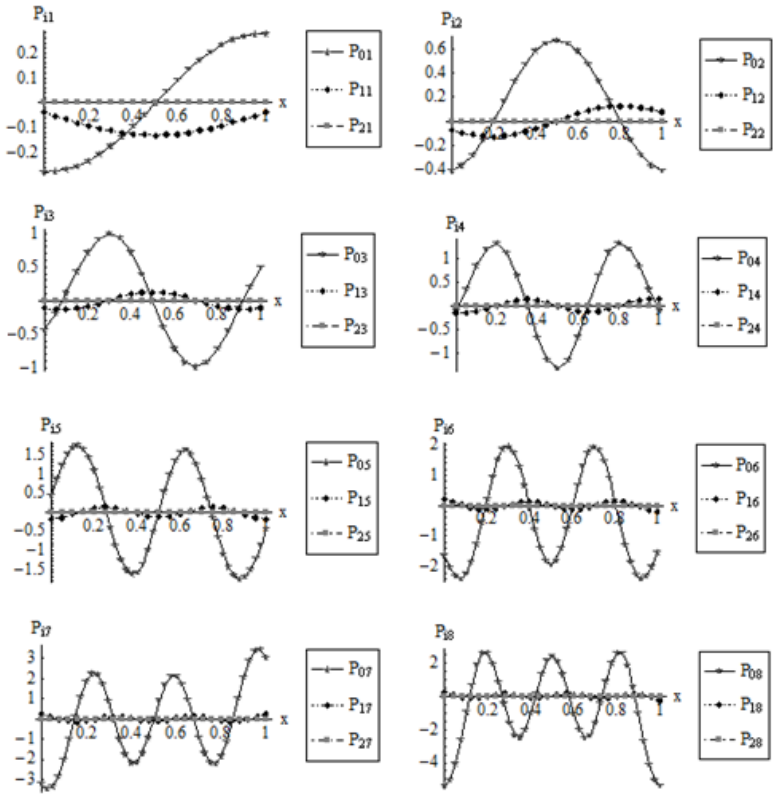


Figure 4.15. The approximated $P_{ij}(x)s$ given by equations (4.2.10), (4.2.11) and (4.2.12) when $0.0 \le x \le 1.0$ for $b = 0.05$.

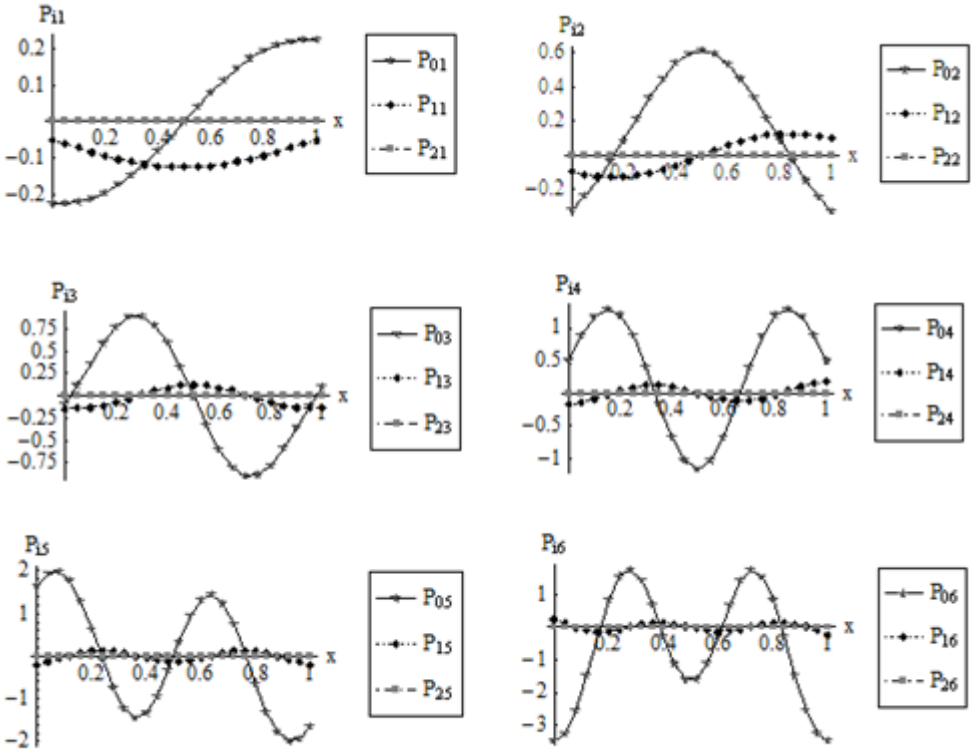


Figure 4.16. The approximated $P_{ij}(x)s$ given by equations (4.2.10), (4.2.11) and (4.2.12) when $0.0 \le x \le 1.0$ for $b = 0.1$.

When $b = 0.1$, the required number of significant eigen values is reduced to 6 and P_{ij} functions are depicted in Figure 4.16. Similar observation as before, when $b = 0.05$, can be made. Figure 4.17 shows P_{ij} functions when $b = 0.2$, and now the number of significant eigen values is 5 (i.e, $m = 5$). The same observations can be made for P_{ij} s when $b = 0.2$.

We produce the 3-dimensional graphs of P_{ij} when $i = 0,1$ and $j = 1,2,3,4,5$ in Figure 4.18 and Figure 4.19; b is plotted as the y-axis. As one could expect, it is reasonable to assume that function surface of P_{ij} is a smooth, continuous function of b. We can define continuous functions of x and b to define $P_{2j}s$.

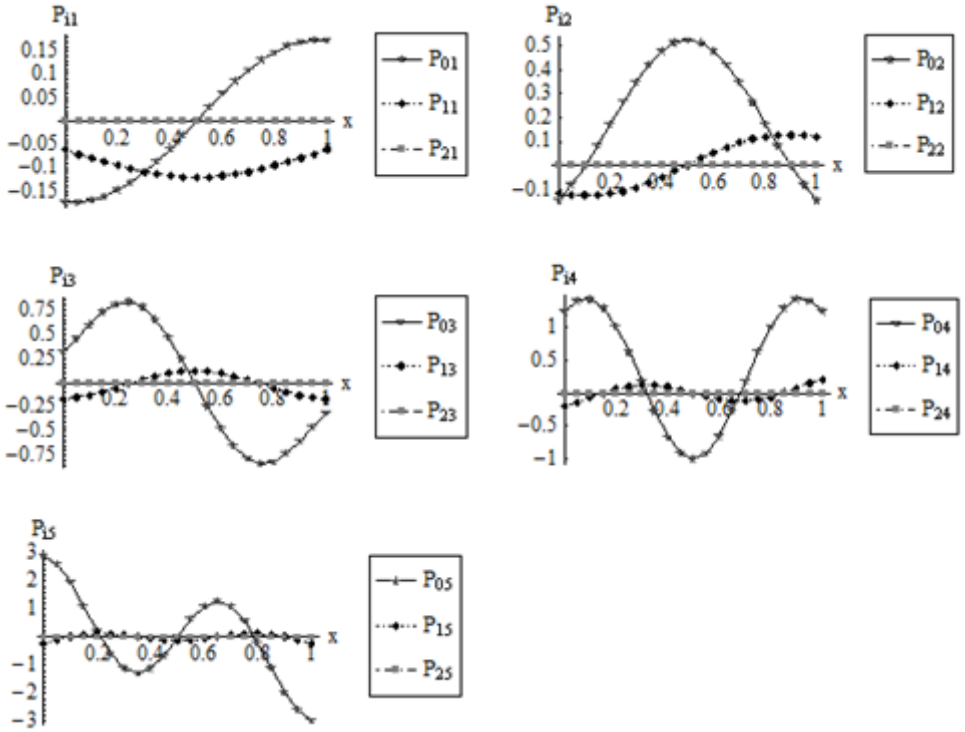
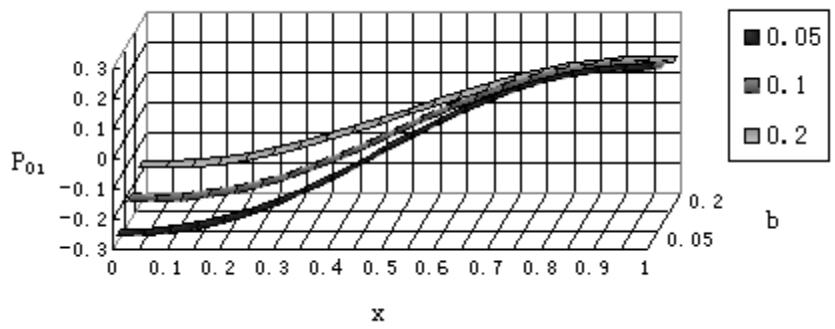


Figure 4.17. The approximated $P_{ij}(x)s$ given by equations (4.2.10), (4.2.11) and (4.2.12) when $0.0 \le x \le 1.0$ for $b = 0.2$.



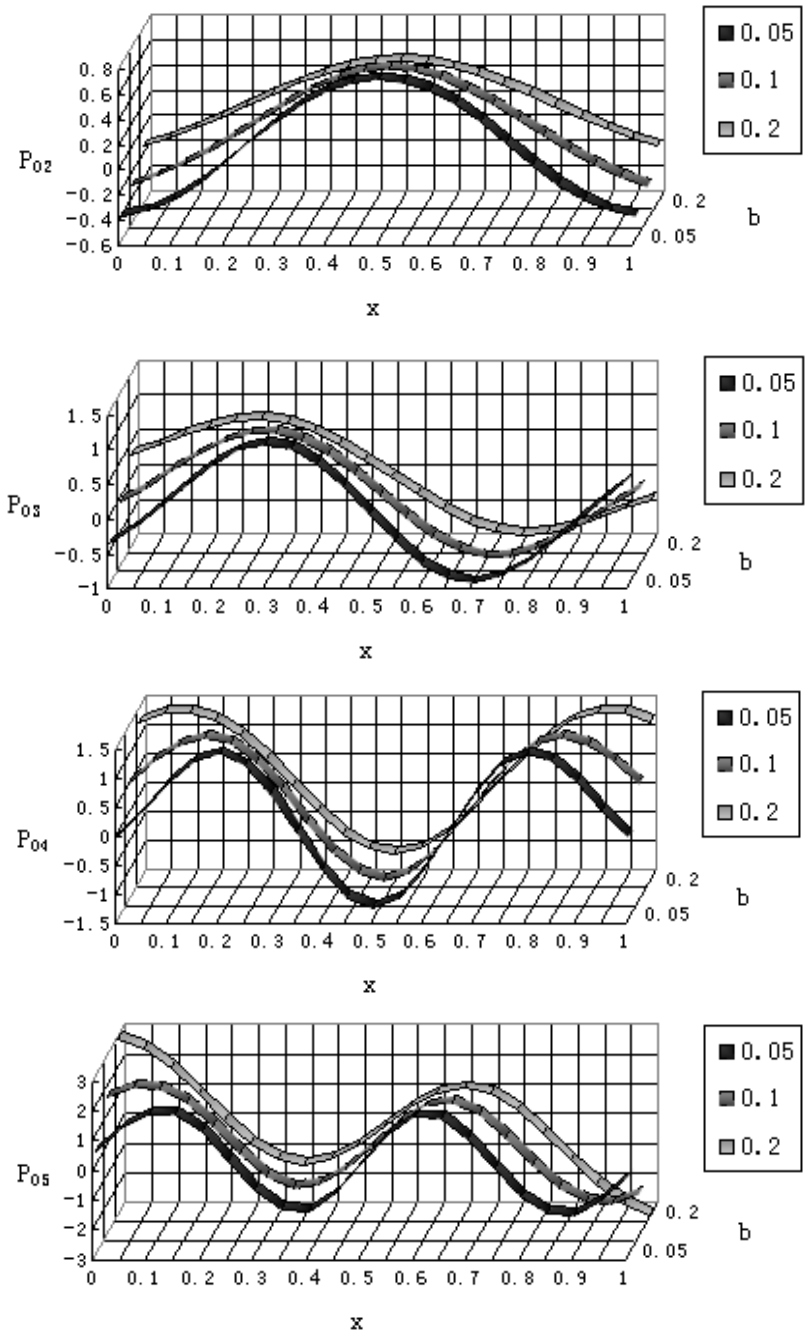
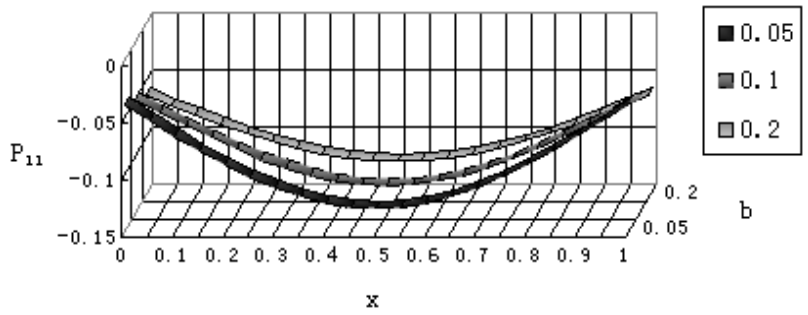
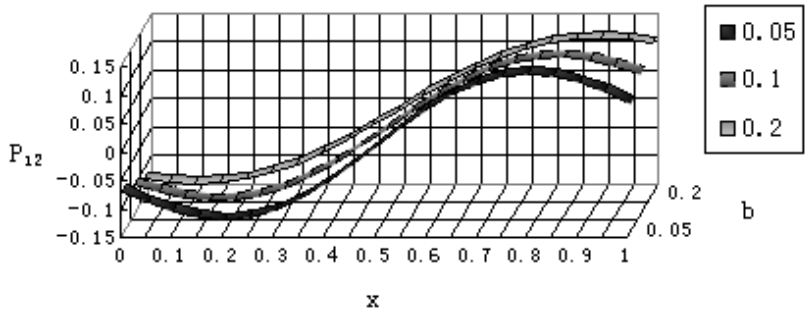
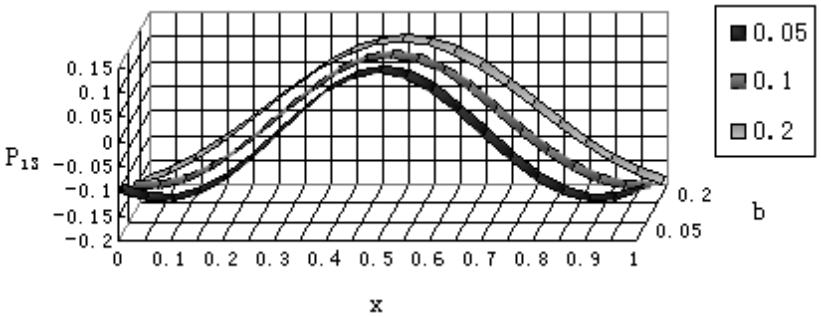
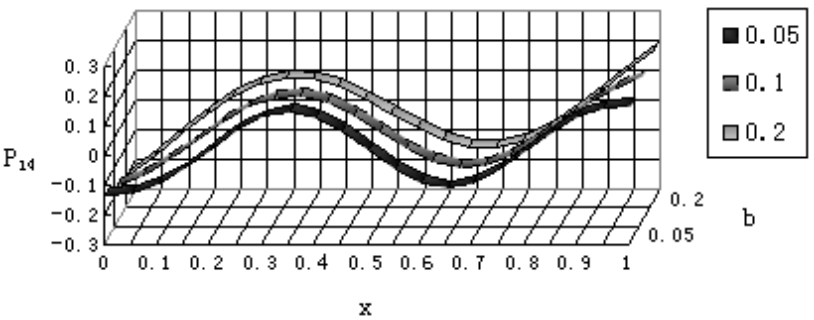


Figure 4.18. The 3-dimensional graphs of P_{ij} when $i=0$ and $j=1,2,3,4,5$


P11
0
-0.05
-0.1
-0.15
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
x
0.2
0.05
b
0.05
0.1
0.2


P12
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
x
0.2
0.05
b
0.05
0.1
0.2


P13
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
x
0.2
0.05
b
0.05
0.1
0.2


P14
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
x
0.2
0.05
b
0.05
0.1
0.2

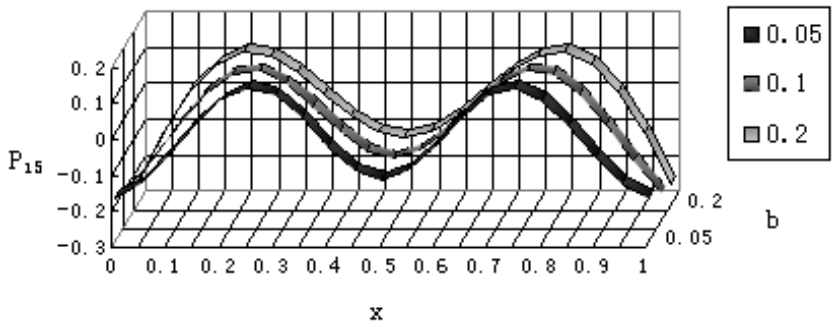


Figure 4.19. The 3-dimensional graphs of P_{ij} when $i=1$ and $j=1,2,3,4,5$

As we recall that the SSTM can be expressed as a diffusion process with martingale properties in time dimensions when another set of diffusion processes, $\underline{I}_x(t)$ are used:

$$dC_x(t) = -C_x(t)dI_{x,0} - \left(\frac{\partial C_x}{\partial x}\right)_x dI_{x,1} - \left(\frac{\partial^2 C_x}{\partial x^2}\right)_x dI_{x,2}. \tag{4.5.13}$$

We need to interpret this equation as follows: the solute concentration at a given point x consists of a combination of three diffusion processes which are solely based on velocity.

The spatial influence on the concentration is mediated through the prevailing concentration, and its spatial gradients. In keeping with the Ito definition of stochastic integration (Klebaner, 1998) we must use the concentration and its spatial gradients at a previous time. As a difference equation, we can write equation (4.5.13) as,

$$C_x(t+1) - C_x(t_n) = -C_x(t_n)dI_{x,0}(t_n) - \left(\frac{\partial C}{\partial x}\right)_{x,t_n} dI_{x,1}(t_n) - \left(\frac{\partial^2 C}{\partial x^2}\right)_{x,t_n} dI_{x,2}(t_n), \tag{4.5.14}$$

where t_n denotes the discretized time and the spatial gradients act as the coefficients of $d\underline{I}_x(t)$, and therefore can be written as,

$$dC_x(t) = \begin{bmatrix} -C_x & -\dfrac{\partial C_x}{\partial x} & -\dfrac{\partial^2 C_x}{\partial x^2} \end{bmatrix} \begin{bmatrix} dI_{x,0} \\ dI_{x,1} \\ dI_{x,2} \end{bmatrix} = \underline{F} d\underline{I}_x, \tag{4.5.15}$$

when, $\underline{F} = \begin{bmatrix} -C_x & -\dfrac{\partial C_x}{\partial x} & -\dfrac{\partial^2 C_x}{\partial x^2} \end{bmatrix}$.

Therefore, we can write,

$$C_x(t) = \int_0^t \underline{F} d\underline{I}_x(t).$$

As P_{2j} functions are insignificant compared to the other two functions, we can approximate equation (4.5.15) as

$$dC_x(t) = \begin{bmatrix} -C_x & -\dfrac{\partial C_x}{\partial x} \end{bmatrix} \begin{bmatrix} dI_{x,0} \\ dI_{x,1} \end{bmatrix}. \tag{4.5.16}$$

This approximation enables us to reduce the computational time further in obtaining solutions.

Equation (4.2.3) gives the general form of eigen functions, $f_j(x)$ s. By inspecting equation (4.2.10), (4.2.11) and (4.2.12), we can deduce the following relationships between P_{ij} s and f_j s:

$$P_{0j}(x) = \frac{df_j(x)}{dx} + h_x \frac{d^2 f_j(x)}{dx^2}, \tag{4.5.17}$$

$$P_{1j}(x) = f_i(x) + h_x \frac{df_j(x)}{dx}, \tag{4.5.18}$$

$$\text{and} \quad P_{2j}(x) = \frac{h_x}{2} f_j(x). \tag{4.5.19}$$

Therefore G_{ij} are given by,

$$G_{0j} = \sqrt{\lambda_j} \frac{df_j}{dx} + h_x \sqrt{\lambda_j} \frac{d^2 f_j}{dx^2}, \tag{4.5.20}$$

$$G_{1j} = \sqrt{\lambda_j} f_j + h_x \sqrt{\lambda_j} \frac{df_j}{dx}, \tag{4.5.21}$$

$$\text{and} \quad G_{2j} = \frac{h_x}{2} \sqrt{\lambda_j} f_j. \tag{4.5.22}$$

The components of diffusion matrix $\underline{a}$ (see equation (4.5.5)) can be written as,

$$a_{ii} = \sigma^2 \sum_{k=1}^{m} G_{(i-1)k}^2, \qquad (i,1,2,3), \tag{4.5.23}$$

$$\text{and} \quad a_{ij} = \sigma^2 \sum_{k=1}^{m} G_{i-1k} G_{j-1k}, \qquad (j,1,2,3) \quad . \tag{4.5.24}$$

For example,

$$\begin{aligned} a_{11} &= \sigma^2 \sum_{k=1}^{m} G_{0k}^2 \\ &= \sigma^2 \sum_{k=1}^{m} \left(\sqrt{\lambda_k} \frac{df_k}{dx} + h_x \sqrt{\lambda_k} \frac{d^2 f_k}{dx^2} \right)^2, \end{aligned}$$

$$a_{11} = \sigma^2 \sum_{k=1}^{m} \left(\lambda_k \left(\frac{\partial f_k}{\partial x} \right)^2 + 2h_x \lambda_k \left(\frac{df_k}{dx} \right) \frac{d^2 f_k}{dx^2} + h_x^2 \lambda_k \left(\frac{\partial^2 f_k}{\partial x^2} \right)^2 \right). \tag{4.5.25}$$

Similarly,

$$a_{22} = \sigma^2 \sum_{k=1}^{m} G_{1k}^2 = \sigma^2 \sum_{k=1}^{m} \left(\lambda_k f_k^2 + 2h_x \lambda_k f_k \frac{df_k}{dx} + h_x^2 \lambda_k \left(\frac{df_k}{dx} \right)^2 \right). \tag{4.5.26}$$

$$a_{33} = \sigma^2 \sum_{k=1}^{m} \frac{h_x}{4} \lambda_j f_j^2. \tag{4.5.27}$$

a_{ij}s can also be evaluated using equation (4.5.24).

This menas that once we have eigen functions in the form given by equation (4.2.3), the diffusion matrix $\underline{a}$ can be evaluated and equation (4.5.7) and (4.5.9) can be used to calculate the strong solutions for $\underline{I}_x(t)$.

4.6 Propagator of $\underline{I}_x(t)$

The propagator of $\underline{I}_x(t)$ can be defined as

$$\varepsilon_i(\Delta t; I_{x,i}(t), t) \equiv \left(I_{x,i}(t + \Delta t) - I_{x,i} \right) \Big| I_{x,i}(t). \tag{4.6.1}$$

ε_i denotes the displacement of $I_{x,i}(t)$ during the infinitesimally small time interval Δt given the position (value) of $I_{x,i}(t)$ at time t. As $I_{x,i}(t)$ is an Ito diffusion, it is also a Markov process; the propagator $\varepsilon_i(\Delta t; I_{x,i}(t), t)$ is also a Markov process having a Gaussian probability density function, which completely specifics the propagator variable. To abbreviate the notation, we denote the propagator as $\varepsilon_i(\Delta t)$. It can be shown that the moments of $\varepsilon_i(\Delta t)$ has the analytical structure (Gillespie,1992),

$$E\left(\varepsilon_i^n(\Delta t)\right) \equiv \int_{-\infty}^{\infty} d\varepsilon \varepsilon^n P\left(\varepsilon \middle| \Delta t\right) = B_u(t)\Delta t + \mathrm{O}(\Delta t), \tag{4.6.2}$$

when B_n s are well behaved functions of t for a given $I_{x,i}(t)$ ($B_n(t)$ are also called the *n* th propagator moment function.) ; and $P\left(\varepsilon \middle| \Delta t\right)$ is the probability density function of $\varepsilon_i(\Delta t)$.

It can be shown that, for a given value of $I_{x,i}(t)$ at time t, the mean and variance of the propagator function can be expressed as follows (Gillespie, 1992):

$$E\left(\varepsilon_i(\Delta t)\right) = A_i(t)\Delta t + \mathrm{O}(\Delta t), \text{ and} \tag{4.6.3}$$

$$Var\left(\varepsilon_i(\Delta t)\right) = D_i(t)\Delta t + \mathrm{O}(\Delta t), \tag{4.6.4}$$

when $A_i(t)$ and $D_i(t)$ are independent of Δt. As the variance can not be negative, $D_i(t)$ should be a non-negative function. As mentioned before, $\varepsilon_i(\Delta t)$ is a Gaussian variable, i.e., it is normally distributed; therefore, we can write,

$$\varepsilon_i(\Delta t) \simeq N\left(A_i(t)\Delta t, D_i(t)\Delta t\right), \tag{4.6.5}$$

where " $\simeq$ " indicates the random variable is generated from the normal density function having the mean, $A_i(t)\Delta t$ and the variance $D_i(t)\Delta t$.

If we recall, equation (4.5.9) gives,

$$E\left\{\left(I_{x,i}(t+\Delta) - I_{x,i}(t)\right)^2 \middle| I_{x,i}(t)\right\} = a_{ii}\Delta t + \mathrm{O}(\Delta). \tag{4.6.6}$$

This can be written as,

$$E\left\{\varepsilon_i^2(\Delta t)\middle| I_{x,i}(t)\right\} = a_{ii}\Delta t. \tag{4.6.7}$$

when $t = t$, i.e., $\Delta t = 0$, therefore $\varepsilon_i(\Delta t) = 0$.

$E(\varepsilon_i / \Delta t) = 0$ at $t = t,$.

This leads to

$$Var\left(\varepsilon_i(\Delta t)\right) = a_{ii}\Delta + \mathrm{O}(\Delta). \tag{4.6.8}$$

By comparing equation (4.5.7) with equation (4.6.3), and equation (4.6.4) with equation (4.6.8), we deduce that,

$$A_i(t) = F_{x,i}, \text{ and} \tag{4.6.9}$$

$$D_i(t) = a_{ii}. \tag{4.6.10}$$

We can now write the Langevin equation for the Ito diffusion, $I_{x,i}(t)$.

4.7 Langevin Equation for $I_{x,i}(t)$

From equation (4.6.1), the propagator can be written as,

$$dI_{x,i}(\Delta t) \equiv I_{x,i}(t+\Delta t) - I_{x,i}(t), \text{ for a given } I_{x,i}(t). \tag{4.7.1}$$

At the same time, we can write equation (4.6.5) as,

$$dI_{x,i}(t) = N\left[a_{ii}(t)\Delta t\right]^{1/2} + F_{x,i}\Delta t, \tag{4.7.2}$$

where N is a unit normal random variable generated from $N(0,1)$ density function. In the limit, $\Delta t \to 0$,

$$dI_{x,i}(t) = F_{x,i}dt + \left[a_{ii}(t)dt\right]^{1/2} N. \tag{4.7.3}$$

This is the first form of the Langevin equation for $I_{x,i}(t)$.

It can be shown that, if $N \simeq Normal(0,1)$,

$$\beta N + \lambda = Normal(\lambda, \beta^2), \tag{4.7.4}$$

Therefore, $\sqrt{dt}N$ is a normally distributed random variable having the density function $Normal(0,dt)$. This is the density function of the standard Wiener process increments, $dw(t)$. Therefore, we can rewrite equation (4.7.3) as,

$$dI_{x,i}(t) = F_{x,i}dt + \sqrt{a_{ii}}dw_i(t). \tag{4.7.5}$$

Equation (4.7.5) is the second form of the Langevin equation.

The advantage of this Langevin approximation for $I_{x,i}(t)$ over equation (4.5.2) is that it is not a multidimensional SDE but a one-dimensional one. This would allow us to compute $I_{x,i}(t)$ s more efficiently.

The moment evolution equations for $I_{x,i}(t)$ can be given as follows: (See Gillespie (1992) for derivations.)

$$\frac{d\left[E\left(I_{x,i}(t)\right)\right]}{dt} = E\left[F_{x,i}\right], \tag{4.7.6}$$

$$\frac{d\left[Var\left(I_{x,i}(t)\right)\right]}{dt} = 2\left(E\left[I_{x,i}(t)F_{x,i}\right] - E\left[I_{x,i}(t)\right]E\left[F_{x,i}\right]\right) + E\left[a_{ii}\right], \tag{4.7.7}$$

with initial conditions,

$$E\left[I_{x,i}(0)\right] = 0, \text{ and} \tag{4.7.8}$$

$$Var\left[I_{x,i}(0)\right] = 0. \tag{4.7.9}$$

If $F_{x,i}$ and a_{ii} are deterministic functions of x then the moment evolution equations can be simplified to equations (4.5.7) and (4.5.8).

4.8 The Evaluation of $C_x(t)$ Diffusions

The SSTM can be written as, for given x, (see equation (4.5.13)),

$$dC_x(t) = -C_x(t)dI_{x,0} - \left(\frac{\partial C_x}{\partial x}\right)_x dI_{x,1} - \left(\frac{\partial^2 C_x}{\partial x^2}\right)_x dI_{x,2} \quad , \tag{4.8.1}$$

where subscript x refers to the first and second derivatives with respect to x. Note that the coefficients are functions of t for given x, and $I_{x,i}$ s are Ito diffusions based on the velocity structure. Equation (4.8.1) is a stochastic diffusion and a SDE which displays the interplay between the concentration profile and velocity structures in the medium. By substituting the equations of the form of equation (4.7.5) for $dI_{x,i}$ s we obtain,

$$\begin{aligned} dC_x(t) = &-C_x(t)\left(F_{x,0}dt + \sqrt{a_{00}}dw_0(t)\right) \\ &-\left(\frac{\partial C}{\partial x}\right)_x \left(F_{x,1}dt + \sqrt{a_{11}}dw_1(t)\right) \\ &-\left(\frac{\partial^2 C}{\partial x^2}\right)_x \left(F_{x,2}dt + \sqrt{a_{22}}dw_2(t)\right). \end{aligned} \tag{4.8.2}$$

In equation (4.8.2), dw_0, dw_1 and dw_2 are independent standard Wiener increments.

Equation (4.8.2) can be written as,

$$dC_x(t) = -\alpha\left(C_x(t),t\right)dt - \sum_{k=0}^{2}\beta_k\left(C_x(t),t\right)dw_k, \tag{4.8.3}$$

where,

$$\alpha\left(C_x(t),t\right) = \alpha = \left(C_x(t)F_{x,0} + \left(\frac{\partial C}{\partial x}\right)F_{x,1} + \left(\frac{\partial^2 C}{\partial x^2}\right)F_{x,2}\right), \tag{4.8.4}$$

$$\beta_0 = C_x(t)\sqrt{a_{00}}\,, \tag{4.8.5a}$$

$$\beta_1 = \left(\frac{\partial C}{\partial x}\right)_x \sqrt{a_{11}}\,, \text{ and} \tag{4.8.5b}$$

$$\beta_2 = \left(\frac{\partial^2 C}{\partial x^2}\right)_x \sqrt{a_{22}}\,. \tag{4.8.5c}$$

Now the equation (4.8.3) can be written as

$$\begin{aligned} dC_x(t) &= -\alpha dt - \sum_{k=0}^{2}\beta_k dw_k\,, \\ &= -\alpha dt + \begin{bmatrix} -\beta_0 & -\beta_1 & -\beta_2 \end{bmatrix}\begin{bmatrix} dw_0 \\ dw_1 \\ dw_2 \end{bmatrix}. \end{aligned} \tag{4.8.6}$$

Equation (4.8.6) gives a diffusion matrix,

$$\begin{aligned} f &= \begin{bmatrix} -\beta_0 & -\beta_1 & -\beta_2 \end{bmatrix}\begin{bmatrix} -\beta_0 & -\beta_1 & -\beta_2 \end{bmatrix}^T, \\ &= \beta_0^2 + \beta_1^2 + \beta_2^2 \end{aligned} \tag{4.8.7}$$

Following Klebaner (1998), the expectations of infinitesimal differences of $C_x(t)$ can be written as (see also equation (4.5.7) and (4.5.9)),

$$E\left(C_x(t+\Delta)-C_x(t)\middle|C_x(t)\right)=-\alpha\Delta+\mathrm{O}(\Delta)\quad , \tag{4.8.8}$$

and

$$E\left(\left(C_x(t+\Delta)-C_x(t)\right)^2\middle|C_x(t)\right)=\left(\beta_0^2+\beta_1^2+\beta_2^2\right)\Delta+\mathrm{O}(\Delta). \tag{4.8.9}$$

Following the same arguments as in the case of deriving the Langevin equation for $I_{x,i}(t)$, we can obtain the Langevin equation for $C_x(t)$:

$$dC_x(t)=-\alpha dt+\sqrt{\beta_0^2+\beta_1^2+\beta_2^2}\,dw(t), \tag{4.8.10}$$

where $dw(t)$ is independent increments of the standard Wiener process, and if

$$\beta_x\equiv\sqrt{\beta_0^2+\beta_1^2+\beta_2^2}\,,\ \text{then}$$

$$dC_x(t)=-\alpha dt+\beta_x dw(t). \tag{4.8.11}$$

This shows that the concentration at given x can be characterised by a Langevin type stochastic differential equation. This equation can be used to develop numerical solutions of the concentration profiles.

The time evolution of the probability density function of $C_x(t)$, $P_x\left(C_x,t|C_x(t),t_0\right)$, is described by Fokker-Plank equation (Klebaner, 1998; Gillespie, 1992). The Fokker-Plank equation for $P_x\left(y,t|y_0,t_0\right)$ is,

$$\frac{\partial P_x\left(y,t|y_0,t_0\right)}{\partial t}=\frac{\partial\left(\alpha P_x\right)}{\partial y}+\frac{1}{2}\frac{\partial^2\left(\beta_x^2P_x\right)}{\partial y^2}, \tag{4.8.12}$$

where y denotes $C_x(t)$ and P_x stands for $P_x\left(y,t|y_0,t_0\right)$.

Equation (4.8.12) has the initial condition,

$$P\left(y,t=t_0|y_0,t_0\right)=\delta\left(y-y_0\right), \tag{4.8.13}$$

where δ is the Dirac-delta function.

Once we solve the Fokker-Plank equation (4.8.12) along with its initial condition, the time evolution of probability density function can be found. This is also a weak solution to equation (4.8.10). Equation (4.8.10) also has strong solutions which can be obtained by integrating the SDE (4.8.11) using Ito integration. The drift coefficient $(-\alpha)$ in equation (4.8.11) is a stochastic variable in x and t.

Similar to equations (4.7.6) and (4.7.7), we can write time evolution of the moments $C_x(t)$. The derivation of these equations are very similar to those given by Gillespie (1992).

The time evolution of the mean of $C_x(t)$ is given by,

$$\frac{d\left(E\left(C_x(t)\right)\right)}{dt} = E\left[-\alpha\right], \tag{4.8.14}$$

therefore,

$$E\left(C_x(t)\right) = \int_0^t E\left[-\alpha\right]dt. \tag{4.8.15}$$

The mean of $C_x(t)$ at a given x is expressed as an integral of the expectation of $(-\alpha)$. As can be seen from equation (4.8.4), α is not only dependent on $C_x(t)$ and its first and second derivatives with respect to x, but also dependent on the mean velocity and its first and second derivatives with respect to x, according to equations (4.2.20), (4.2.21) and (4.2.22). The initial condition for equation (4.8.15) is $C_x(0)$, the value of the concentration at time is zero for a given x.

The evolution of the variance of $C_x(t), \left(Var\left(C_x(t)\right)\right)$ is given by,

$$\frac{d\left[Var\left(C_x(t)\right)\right]}{dt} = 2\left(E\left(-\alpha C_x(t)\right) + E\left(C_x(t)\right)E(\alpha)\right) + E\left[\beta_x^2\right], \tag{4.8.16}$$

and the variance of $C_x(t)$ can be obtained by integrating equation (4.8.16) with respect to t,

$$Var\left(C_x(t)\right) = 2\int_0^t E\left(-\alpha C_x(t)\right)dt + 2\int_0^t E\left(C_x(t)\right)E(\alpha)\,dt + \int_0^t E\left[\beta_x^2\right]dt. \tag{4.8.17}$$

By Fubini's theorem (Klebaner,1998), for continuous stochastic variable such as α, $C_x(t)$ and β_x, we can rewrite equation (4.8.17) as,

$$Var\left(C_x(t)\right) = -2E\left[\int_0^t \alpha C_x(t)\,dt\right] + 2\int_0^t E\left(C_x(t)\right)E(\alpha)\,dt + E\left[\int_0^t \beta_x^2\right]dt. \tag{4.8.18}$$

In equation (4.8.18), Fubini's theorem is only applied to the first and third term of equation (4.8.17).

Once we evaluate the mean and variance of $C_x(t)$, we can obtain the probability density function, $P_x\left(y,t|y_0,t_0\right)$ of $C_x(t)$ (y represents the value of $C_x(t)$). As Ito diffusions are Martingales with Markovian properties (Klebaner, 1998), and $C_x(t)$ is Gaussian,

$$P_x\left(y,t|y_0,0\right) \equiv P_x\left(y,t\right) = \frac{1}{\left[2Var\left(y\right)\right]^{1/2}} \exp\left(\frac{-\left(y-E\left(y\right)\right)^2}{2Var\left(y\right)}\right). \tag{4.8.19}$$

when $E(y)$ and $Var(y)$ are given by equations (4.8.15) and (4.8.18). When time is zero, $P_x\left(C_x\left(0\right)\right) = 1.0$.

Equation (4.8.19) should also be the solution to the Fokker-Plank equation (4.8.12). The Fokker-Plank equation (4.8.12) is a stochastic partial differential equation for which analytical equations can not easily be obtained. Therefore, we can make use of this fact to verify the numerical solutions for the Fokker-Planck equation.

4.9 Numerical Solutions

We have seen in the previous section 4.8, the following SDE gives the time course of concertration $C_x(t)$ for a given x in the vicinity of x:

$$dC_x\left(t\right) = -\alpha dt + \beta_x dw\left(t\right), \tag{4.9.1}$$

where $dw(t)$ is the standard Wiener increments with a zero mean and dt variance, if dt is sufficiently small; the drift coefficient α is given by (see equation (4.8.4)),

$$\alpha = C_x\left(t\right)F_{x,0} + \left(\frac{\partial C}{\partial x}\right)_x F_{x,1} + \left(\frac{\partial^2 C}{\partial x^2}\right)_x F_{x,2}. \tag{4.9.2}$$

In equation (4.9.2),

$$F_{x,0} = \frac{\partial \overline{V}\left(x,t\right)}{\partial x} + \frac{h_x}{2}\frac{\partial^2 \overline{V}\left(x,t\right)}{\partial x^2}, \tag{4.9.3}$$

$$F_{x,1} = \overline{V}\left(x,t\right) + h_x\frac{\partial \overline{V}\left(x,t\right)}{\partial x}, \tag{4.9.4}$$

and,

$$F_{x,3} = \frac{h_x}{2}\overline{V}\left(x,t\right); \tag{4.9.5}$$

where $\overline{V}(x,t)$ in the mean velocity which is assumed to be regular differentiable continuous function.

β_x in equation (4.9.1) is given by,

$$\beta_x = \left(\beta_0^2 + \beta_1^2 + \beta_2^2\right)^{1/2}, \tag{4.9.6}$$

where,

$$\beta_0 = C_x\left(t\right)\sqrt{a_{00}}, \tag{4.9.7}$$

$$\beta_1 = \left(\frac{\partial C}{\partial x}\right)_x \sqrt{a_{11}}, \text{ and}$$

$$\beta_2 = \left(\frac{\partial^2 C}{\partial x^2}\right)_x \sqrt{a_{22}}, \tag{4.9.8}$$

The following equation gives the expressions of a_{00}, a_{11}, and a_{22},

$$a_{ii} = \sigma^2 \sum_{j=1}^{m} G_{ij}^2, \qquad (i,0,1,2), \tag{4.9.9}$$

where m is the number of effective eigen functions, and

$$G_{ij} = \sqrt{\lambda_j} P_{ij}, \quad (i,0,1,2). \tag{4.9.10}$$

In the numerical solutions, we make use of the finite differences, for a given dependent variable, say U, based on the grid given in Figure 4.20

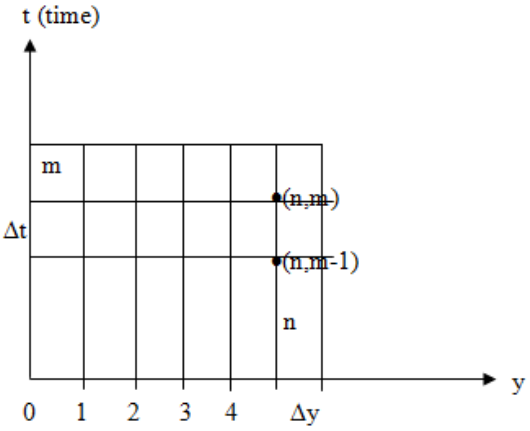


Figure 4.20. Space-time grid used in the numerical solutions with respect to y

$$\left(\frac{\partial U}{\partial y}\right)_n^m = \frac{\left(U_n^m - U_{n-1}^m\right)}{\Delta y}, \tag{4.9.11}$$

$$\left(\frac{\partial^2 U}{\partial y^2}\right)_n^m = \frac{\left(U_n^m - 2U_{n-1}^m + U_{n-2}^m\right)}{\Delta y^2}, \tag{4.9.12}$$

and

$$\left(\frac{\partial U}{\partial t}\right)_n^m = \frac{\left(U_n^{m+1} - U_n^m\right)}{\Delta t}. \tag{4.9.13}$$

By using a numerical scheme developed, we obtain realizations of $C_x(t)$ as strong solutions to equation (4.9.1). We can also obtain solutions to the Fokker-Plank equation (4.8.12) using the same finite differences.

4.10 Remarks for the Chapter

In this Chapter, we develop a generalized form of SSTM that can include any arbitrary velocity covariance kernel in principle. We have demonstrates that for a given kernel, a generalized analytical forms for eigen functions can be obtained by using the computational methods developed. We have also developed a Langevin form of the SSTM for a given x, and the time evolution of concentration, $C_x(t)$, follows a stochastic differential equation having the coefficients α and β_x which are again functions of $C_x(t)$ and eigen functions. In other words, if one monitors the concentration $C_x(t)$ at a given point in space, the data collected along with time would constitute a realization of the strong solution of the SDE. The solution is a function of the covariance kernel, i.e. a function of σ^2 and b for an exponentially decaying kernel, and also a function of $C_x(t)$ itself and its first- and second derivatives with respect to x. This focus of SDE provides a very convenient and computationally efficient way to solve the stochastic partial differential equation associated with the SSTM.

By deriving a Langevin form of the SSTM, we essentially prove that any time course of the concentration at a given point behaves according to the underlying SDE, which would characterize the nature of local porous medium and is a statement of mass concentration of the solute.