

Semianalytical methods in stochastic groundwater transport

Sergio E. Serrano

Department of Civil Engineering, University of Kentucky, Lexington, KY, USA

In this article, semianalytical solutions of groundwater stochastic partial differential equations are used to build mathematical models to investigate the effect of erratic physical parameters on the dependent variables. Semianalytical solutions of random equations appear to exhibit many of the same advantages as semianalytical solutions of deterministic equations. They offer a higher degree of stability and accuracy and easier computational algorithms with respect to fully numerical solutions, while allowing the inclusion of irregular geometrical domains and less complex mathematics than analytical solutions. It is also observed that the Neumann expansion used to approximate the stochastic equation allows the consideration of higher parameter variances than the ones restricted to conventional small perturbation techniques.

Keywords: groundwater models, stochastic differential equations, semianalytical methods

1. Introduction

An important difficulty in the development of accurately forecasting mathematical models of groundwater flow and groundwater pollution is the uncertain (random) variability of the physical parameters describing porous media properties in natural aquifers. In recent years the modelling efforts have been concentrated on the specification of the stochastic properties of the aquifer parameters and the solution of the governing stochastic partial differential equation. Several authors have analyzed the effect of parameter variability, using various stochastic descriptions of spatial variability.¹⁻¹¹ Among the methods used are Monte Carlo simulations, spectral decomposition, small perturbation expansions, Neumann expansions, and semigroup solutions.

In the present article an attempt is made to explore the use of semianalytical solutions of groundwater stochastic equations. Semianalytical (or seminumerical) solutions of groundwater equations may be more stable than corresponding numerical solutions; they may open the way to include domains with irregular boundaries, which are difficult in analytical solutions; they offer simpler solutions than corresponding analytical solutions; they may offer a higher degree of accuracy with more efficient computational algorithms than corresponding numerical algorithms; and they have been

shown to have important applications in the solution of groundwater problems.¹² The procedure employed consists of discretizing one of the independent variables in the partial differential equation (either the time or the space coordinate, depending on whether the parameter uncertainty is represented as a space or a time random process) and obtaining a solution of the resulting equation. The solution is approximated by truncating the corresponding Neumann series expansion. Each term in the series contains the semigroup operator associated with the differential equation, that is, a convolution integral of the system impulse response function. The solution method, which we now wish to extend to partially discretized equations, has been tested with success in several full partial differential equations in groundwater transport.¹⁰ In particular, the method allows the consideration of larger variances than the ones restricted by conventional small perturbation techniques (see section 2, the appendix, and Ref. 13).

Section 2 contains a general description of the solution method of groundwater equations subject to random parameter variability. Section 3 introduces models based on semianalytical solutions of the stochastic differential equation, which can be employed when the input uncertainty occurs with respect to only one of the independent variables and therefore use the convenient features of numerical methods. The handling of the resulting stochastic matrices is discussed in section 3.1 with a groundwater flow case subject to time variability in the transmissivity. The case in which spatial erratic variability in the transmissivity is present is discussed in section 3.2, and in section 3.3 a groundwater pollution case in which spatial erratic variability in the velocity field is the most important uncertainty

Address reprint requests to Dr. Serrano at the Department of Civil Engineering, University of Kentucky, Lexington, KY 40506-0046, USA.

Received 13 March 1991; revised 24 September 1991; accepted 10 October 1991

term is described. Finally, the appendix presents a theorem with a proof showing the convergence criterion for the Neumann expansion and a description of the use of the method for relatively high variances in the random parameters.

2. Models to investigate the effect of parameter variability

2.1 Description of the methodology with a groundwater pollution case

Before studying the effect of random variability of a system parameter on the dependent variable of a groundwater differential equation, let us visualize the accuracy of the Neumann expansion solution method with a simple deterministic example. Consider the problem of groundwater pollution by a conservative contaminant in an infinite homogeneous aquifer, as governed by the advective-dispersive equation,¹⁴

$$\frac{\partial C}{\partial t} - D \frac{\partial^2 C}{\partial x^2} + u \frac{\partial C}{\partial x} = 0 \tag{1}$$

$$C(-\infty, t) = C(\infty, t) = 0 \quad C(x, 0) = C_0(x)$$

where C is the contaminant concentration in milligrams per liter, D is the average aquifer dispersion coefficient in square meters per day, u is the average aquifer pore velocity in meters per day, x is the horizontal distance from an origin in meters, t is time in days, and the initial condition, C_0 , is a smooth known function of x .

The semigroup solution to this differential equation is¹⁰

$$C(x, t) = J(x, t)C_0$$

$$= \frac{1}{(4\pi Dt)^{1/2}} \int_{-\infty}^{\infty} \exp \left[-\frac{(x - ut - s)^2}{4Dt} \right] C_0(s) ds \tag{2}$$

where J is the strongly continuous semigroup associated with the spatial differential operator in (1). Now assume that the initial condition may be represented as a pulse function of magnitude equal to c' mg/L for $x_1 \leq x \leq x_2$ and equal to zero everywhere else, that is

$$C_0(x) = c' H(x - x_1)H(x_2 - x) \tag{3}$$

where $H(\)$ denotes the unit step function. This initial condition would represent an accidental spill penetrating the saturated zone at an average concentration of c' mg/L over an area $(x_2 - x_1)$ m in length. Substituting (3) into (2) yields

$$C(x, t) = \Phi(x, t)$$

$$= \frac{c'}{2} \left\{ \operatorname{erf} \left[\frac{x_1 - x + ut}{(4Dt)^{1/2}} \right] - \operatorname{erf} \left[\frac{x_2 - x + ut}{(4Dt)^{1/2}} \right] \right\} \tag{4}$$

where $\operatorname{erf}(\)$ denotes the error function.

Let us now assume that we wish to investigate the effect of an increase of 100% in the magnitude of the velocity field u . In that case, (1) becomes

$$\frac{\partial C}{\partial t} - D \frac{\partial^2 C}{\partial x^2} + u \frac{\partial C}{\partial x} = -u \frac{\partial C}{\partial x} \tag{5}$$

The solution to this equation is

$$C(x, t) = \Phi(x, t) - u \sum_{i=1}^N \int_0^t J(x, t - \tau) \frac{\partial C_i}{\partial x} d\tau \tag{6}$$

where the C_i constitute a successive approximation series. To show the nature of (6), we note that the general solution of (5) is

$$C(x, t) = J(x, t)C_0 - u \int_0^t J(x, t - \tau) \frac{\partial C}{\partial x} d\tau \tag{7}$$

which can be expanded as

$$C(x, t) = \Phi(x, t) - u \int_0^t \frac{1}{(4\pi D(t - \tau))^{1/2}} \int_{-\infty}^{\infty} \exp \left[-\frac{(x - u(t - \tau) - s)^2}{4D(t - \tau)} \right] \frac{\partial C(s, \tau)}{\partial s} ds d\tau \tag{8}$$

after using (2) and (4)

Now, when C is expanded in the right side of (8) as the Neumann series¹⁰ $C = C_1 + C_2 + \dots$, (6) follows. $C_1 = \Phi$, which is the first and most important partial solution, and in general,

$$C_i = -u \int_0^t J(x, t - \tau) \frac{\partial C_{i-1}}{\partial x} d\tau \quad i \geq 2 \tag{9}$$

N in (6) is a suitable number of iterations such that C reaches a degree of accuracy on the order of the representative scale of resolution of the concentration measurement device used in the field.

As an illustrative example, assume $c' = 10.0$ mg/L, $x_1 = 30.0$ m, $x_2 = 50.0$ m, an initial velocity $u = 0.2$ m/day, and $D = 0.1$ m²/day, and compute the breakthrough curves at $t = 30$ days after the spill. *Figure 1* illustrates the exact concentration magnitude versus distance, the concentration as obtained by the first-order approximation, and the concentration as obtained by a second-order approximation. The exact concentration versus distance was calculated from (4) for $u = 0.4$ m/day after the 100% increase in the original velocity. The concentration by a first-order approximation was computed from (6) with $u = 0.2$ m/day and corresponds to the first term in the equation. The concentration by a second-order approximation was obtained by adding the first two terms in (6) with $N = 1$. Thus *Figure 1* illustrates that the concentration profile after an increase in the velocity field of as much as 100% of the original magnitude is accurately estimated by a second-order approximation in (6). *Figure 2* illustrates the first three approximations individually. Note the uniform convergence of the approximation series to zero, indicating that considering only the first few terms may be sufficiently accurate.

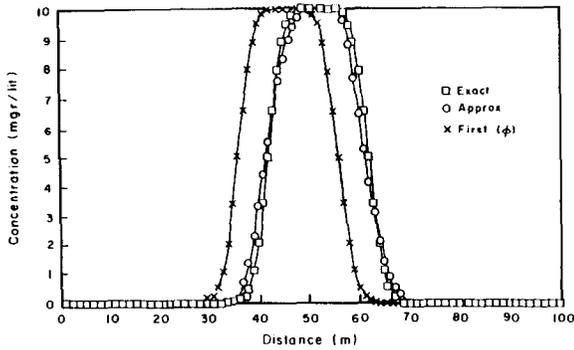


Figure 1. Concentration breakthrough curves after a 100% increase in the velocity

The programming effort or computer time required for the implementation of (6) is not high. A simple 24-point Gauss-Legendre quadrature was employed for the integral term with a first-order central finite difference approximation for the derivative within the integral.

2.2 Groundwater pollution subject to a spatially erratic velocity field

Consider again the model problem of section 2.1, and assume that the hydrogeology of the area is such that the measured groundwater velocities at discrete

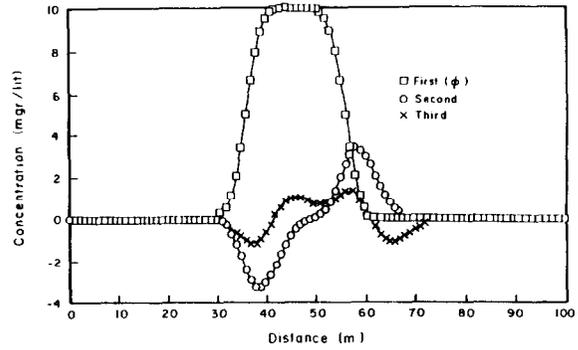


Figure 2. First three individual approximations versus distance

points in the aquifer over a long period of time suggest that the velocity field is a highly variable (uncertain) function that could be described as $u(x, \omega) = \bar{u} + u'(x, \omega)$, where \bar{u} is a long-term mean, $u'(x, \omega)$ represents the zero-mean spatially random component in the velocity field, and ω is the probability variable. In this case, (1) becomes

$$\frac{\partial C}{\partial t} - D \frac{\partial^2 C}{\partial x^2} + \bar{u} \frac{\partial C}{\partial x} = -u' \frac{\partial C}{\partial x} \tag{10}$$

subject to the same set of boundary conditions. Proceeding in a manner similar to the one before, we find the solution to this equation,

$$C(x, t) = \Phi(x, t) - \sum_{i=1}^N \int_0^t \frac{1}{(4\pi D(t-\tau))^{1/2}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\bar{u}(t-\tau)-s)^2}{4D(t-\tau)}\right] u'(s, \omega) \frac{\partial C_i(s, \tau)}{\partial s} ds d\tau \tag{11}$$

subject to $C_i(x, t) = \Phi(x, t)$, $\Phi(x, t)$ as given by (4) if we use the same initial condition (3), and

$$C_i(s, \tau) = - \int_0^\tau \frac{1}{(4\pi D(\tau-\gamma))^{1/2}} \int_{-\infty}^{\infty} \exp\left[-\frac{(s-\bar{u}(\tau-\gamma)-\rho)^2}{4D(\tau-\gamma)}\right] u'(\rho, \omega) \frac{\partial C_{i-1}}{\partial \rho} d\rho d\gamma \quad i \geq 2 \tag{12}$$

Interesting features of the Neumann expansion series for dissipative systems are the fast uniform convergence, which permits an accurate solution within a few iterations, and the possibility of including high variability in the random parameters. In the appendix a theorem establishing a criterion for the convergence of the series is given, and in Ref. 13 a comparison between exact solutions to the random equation and the Neumann series solution with $N = 1$ is given with favorable results. A sufficient condition for the convergence of the series is $\frac{1}{2}Mt < 1$, where M is the maximum sample value in the random velocity. Clearly, the simulation time step, t , can be adjusted to meet this requirement.

Note that in (11) the semigroup operator is the same as that given by (2) with a different source, \bar{u} instead of u , and $(t - \tau)$ instead of t . Equation (11) can be used to generate sample functions to observe the qualitative nature of the concentration evolution if sample

functions of u' are available. In most engineering applications, however, only the first lower order moments are available after statistical inference on the limited historical measurements of the velocity field. Thus it seems natural to develop expressions that would describe the corresponding lower order moments of the concentration field. According to the earlier results in this section, a value of $N = 1$ in (11) produced an accurate estimation of the concentration. It is shown¹³ that $N = 1$ is also sufficiently accurate for the case of variable parameters. After taking expectations on both sides of (11) we obtain an approximation to the mean concentration,

$$E\{C(x, t)\} = \Phi(x, t) \tag{13}$$

where $E\{ \}$ denotes the expectation operator. For $N = 1$ the mean concentration coincides with the deter-

ministic component. It is a modelling decision to include additional terms if it is judged that the variances of u' are so high that a first-order approximation is crude. The computational effort required by additional terms is small. Figure 3 shows the mean concentration versus distance 30 days after the spill according to (13), along with one sample realization of the concentration using (11) with the same deterministic parameters as before. In this experiment the random component of the velocity field, u' , was assumed to follow a Gaussian colored noise process in space of the form

$$E\{u'(x_1, t)u'(x_2, t)\} = R_u(x_1, x_2) = qe^{-\rho|x_1-x_2|} \quad (14)$$

where variance parameter $q = 0.1$ and decay parameter $\rho = 0.01$. This is a common assumption adopted by researchers, but of course in practical applications the functional form of u' must be supported by long-term field measurements. A sample function of a colored noise groundwater velocity with the above properties is necessary in (11) in order to obtain a sample function of the concentration. For the practical details about the numerical generation of such a sequence the reader is referred to Ref. 15. Figure 3 indicates that the lowest variability in the concentration field appears where the concentration gradients tend to zero.

The variance and the covariance of the concentration can be derived from the two-point (time or space) correlation function, which in turn depends on the correlation structure of the groundwater velocity as derived from field measurements.

Using (11) and (14), we deduce an estimate of the variance of the concentration at a particular point in time

$$\sigma_{C(x,t)}^2 = q \int_0^t \int_0^t J(x, s, t - \tau) J(x, \xi, t - \gamma) \times e^{-\rho|s-\xi|} \frac{\partial \Phi(s, \tau)}{\partial s} \frac{\partial \Phi(\xi, \gamma)}{\partial \xi} d\gamma d\tau \quad (15)$$

where the approximation $N = 1$ has been adopted again. Figure 4 shows the standard deviation of the concentration versus distance 30 days after the spill

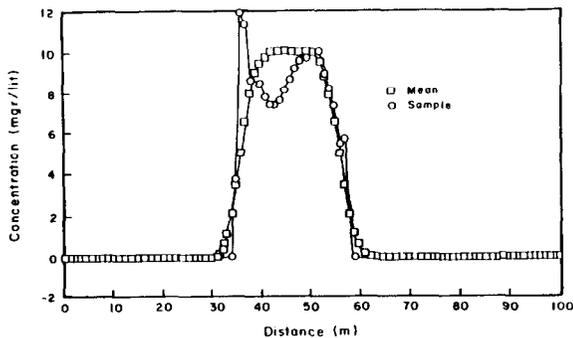


Figure 3. Mean and one sample concentration versus distance

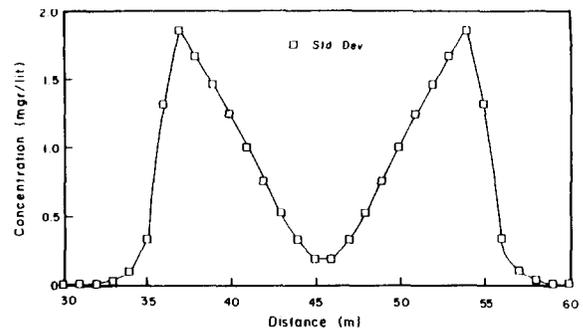


Figure 4. Standard deviation of concentration versus distance

according to (15). Once again, the lowest values of the standard deviation occur at locations where the concentration gradient tends to zero, that is, at $x \leq 33$, $x \geq 57$, and particularly at the peak concentration, $43 \leq x \leq 48$. The highest values of the standard deviation (the expected highest variability in concentration values) occur at locations where the concentration gradients are maximum, that is, at $36 \leq x \leq 40$ and at $51 \leq x \leq 55$. This phenomenon can be interpreted from (15) and an observation of Figure 3.

3. Semianalytical solutions of the random equations

Some modelling problems allow a discretization in certain dimension, and therefore the advantages of numerical methods can be beneficial. In certain circumstances the modeller observes that the variability associated with one or more of the parameters is more important with respect to one of the independent variables. For example, field measurements may demonstrate that the aquifer transmissivity is relatively constant with respect to the spatial coordinates while exhibiting an overall temporal erratic variability. This would suggest a convenient discretization of the deterministic (spatial) variables while keeping continuous the (time) stochastic variables. In these cases, where the stochasticity is important with respect to one of the independent variables and negligible with respect to the other independent variables, certain simplifications that reduce the complexity of the semigroup operator and that of the solution are possible.

3.1 Regional groundwater flow subject to time-variable transmissivity

Consider the case of a two-point boundary value problem describing regional groundwater flow in a horizontal phreatic aquifer bounded by rivers. Assume that the fluctuation of the aquifer transmissivity is a highly variable time function while at the same time its bulk magnitude, as estimated by a series of pumping tests, does not change drastically with respect to distance. This situation arises when the bulk hydraulic conductivity is relatively constant while the groundwater head oscillates daily and seasonally as a result

of random variability in the recharge and other environmental fluctuations (i.e., $T(t) = Kh(t)$). Thus the linearization of the groundwater flow equation with Dupuit assumptions results in the equation¹⁴

$$\frac{\partial h}{\partial t} - \frac{T(t, \omega)}{S} \frac{\partial^2 h}{\partial x^2} = \frac{i(t)}{S} \quad (16)$$

$$h(0, t) = h_1 \quad h(L, t) = h_2 \quad h(x, 0) = h_0(x)$$

where $h(x, t, \omega)$ is the elevation of the water table (Dupuit assumptions) with respect to the bottom of the aquifer, in meters; x is the horizontal distance from the left boundary in meters; h_1 and h_2 are the constant left and right heads at the rivers, respectively, in meters; h_0 is a smooth initial head across the aquifer, as interpolated from a few field piezometers; L is the length of the aquifer in meters; t is the time coordinate in days; S is the aquifer specific yield; $T(t, \omega)$ is the time random process representing the aquifer transmissivity, in square meters per day; and $i(t)$ is the aquifer recharge, which in this case is assumed spatially independent, in meters.

Equation (16) can be further reduced after assuming $h(x, t) = V(x) + u(x, t, \omega)$ where $V(x) = h_1 + [(h_2 - h_1)/L]x$ is a smooth steady-state function and u is the stochastic transient function satisfying

$$\frac{\partial u}{\partial t} - \frac{1}{S} \frac{\partial}{\partial x} \left(T(t, \omega) \frac{\partial u}{\partial x} \right) = \frac{i(t)}{S}$$

$$u(0, t) = u(L, t) = 0 \quad u(x, 0) = h_0(x) - V(x) \quad (17)$$

Now assume that the transmissivity function can be defined as $T(t, \omega) = \bar{T} + T'(t, \omega)$, where \bar{T} represents the average transmissivity and T' represents the zero-mean random component. Thus (17) becomes

$$\frac{\partial u}{\partial t} - \frac{\bar{T}}{S} \frac{\partial^2 u}{\partial x^2} = \frac{i(t)}{S} + \frac{T'}{S} \frac{\partial^2 u}{\partial x^2} \quad (18)$$

The similarity of (18) to (10) may suggest a straightforward solution at this stage. However, because (18) is an evolution equation forced by a time stochastic process, a discretization with respect to distance is plausible.

Let us formally subdivide the spatial domain into $n + 2$ nodes equally spaced by a distance Δx , so that $k = 0$ at $x = 0$ and $k = n + 1$ at $x = L$. Discretizing the spatial derivatives with a simple central finite difference approximation, we write (18) for the node k at time t as

$$\frac{\partial u_k}{\partial t} - \frac{\bar{T}}{S} \left(\frac{u_{k-1} - 2u_k + u_{k+1}}{\Delta x^2} \right) = \frac{i}{S} + \frac{T'}{S} \left(\frac{u_{k-1} - 2u_k + u_{k+1}}{\Delta x^2} \right) \quad (19)$$

After applying this equation to every node, including the boundary nodes, we can write the resulting system of differential equations in matrix form:

$$\frac{dU(t, \omega)}{dt} - \bar{A}U(t, \omega) = F(t) + A'(t, \omega)U(t, \omega) \quad (20)$$

$$U(0) = U_0$$

where U is a $n \times 1$ unknown column vector containing the values of u_k at the different nodes; \bar{A} is a $n \times n$ conductivity matrix containing values of $\bar{a} = \bar{T}/(S \cdot \Delta x^2)$; F is the known column vector containing recharge values, $b = i(t)/S$; A' is a square random matrix identical in form to A , except that instead of constants, a , it contains the random processes $a' = T'/(S \Delta x^2)$; and U_0 is the known column vector containing the initial condition, u_{0k} . The arrays are, respectively,

$$U = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \bar{A} = \begin{bmatrix} -2\bar{a} & \bar{a} & 0 & 0 & \cdots \\ \bar{a} & -2\bar{a} & \bar{a} & 0 & \cdots \\ 0 & \bar{a} & -2\bar{a} & \bar{a} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$F = \begin{bmatrix} b \\ b \\ \vdots \\ b \end{bmatrix} \quad U_0 = \begin{bmatrix} u_{01} \\ u_{02} \\ \vdots \\ u_{0n} \end{bmatrix}$$

Thus we have reduced the stochastic partial differential equation (18) into an ordinary vector stochastic differential equation (20) whose semigroup operator, and therefore its solution, is simpler. However, the price to be paid is measured in terms of loss of stability and increased computer space and time requirements.

The solution to (20) is

$$U(t, \omega) = J(t)U_0 + \int_0^t J(t-s)F(s) ds + \int_0^t J(t-s)A'(s, \omega)U(s, \omega) ds \quad (21)$$

where

$$J(t) = e^{\bar{A}t} = I + \bar{A}t + \frac{\bar{A}^2 t^2}{2!} + \frac{\bar{A}^3 t^3}{3!} + \cdots$$

is the semigroup and I is the identity matrix. The computation of the semigroup, or the exponential of a matrix in this case, should not be done by the approximation of its formal Taylor series expansion, since the results may prove inaccurate. There are many different methods of approaching the problem.^{16,17} For the present application we note that if the eigenvectors associated with the matrix \bar{A} are linearly independent, then an $n \times n$ modal matrix M may be formed (that is, a square matrix each of whose columns contains an eigenvector of \bar{A}) from the eigenvectors, and the similarity transformation¹⁸

$$S = M^{-1} \bar{A} M$$

where M^{-1} the inverse of M , can be used to diagonalize \bar{A} into the spectral matrix S composed of the eigenvalues, λ , of \bar{A} :

$$S = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

The matrix \bar{A} can be reclaimed from its spectral matrix via $\bar{A} = MSM^{-1}$. Any functional of \bar{A} can be expressed as $f(\bar{A}) = f(MSM^{-1}) = Mf(S)M^{-1} = Mf(\lambda)M^{-1}$. This expression may well be the simplest method for the evaluation of the exponential of a matrix, that is, $J(t) = e^{\bar{A}t} = Me^{St}M^{-1}$.

Returning to the solution equation (21) and following the same procedure as that in section 2 for the approximation of the third term in the equation, we obtain

$$U(t, \omega) = J(t)U_0 + \int_0^t J(t-s)F(s) ds + \sum_{i=1}^N \int_0^t J(t-s)A'(s, \omega)U'_i(s, \omega) ds \quad (22)$$

For practical applications, sample functions of a' in A' and measured values of b in F are approximated as staircase sequences A'_i and F_i , respectively. Setting $t = 1$ day (an assumed representative time scale for the recharge), then we can simulate realizations of U recursively at intervals of time one day apart:

$$U_t = J(1)U_{t-1} + \left[\int_0^1 J(1-s) ds \right] F_t + \sum_{i=1}^N \int_0^1 J(1-s)A'_i U'_i(s) ds \quad (23)$$

where F_i and A'_i are constants within the integration limits. A further simplification is obtained with the trapezoidal rule, valid for small Δt ,

$$U_t = J(1)U_{t-1} + \frac{1}{2}(I + J(1))F_t + \frac{1}{2}(I + J(1))A'_i \sum_{i=1}^N U'_i \quad (24)$$

where the semigroup properties of J have been used¹⁰; $U'_1 = J(1)U_{t-1}$, the first term in (24); and in general,

$$U'_i = \frac{1}{2}(I + J(1))A'_i U'_{i-1} \quad i \geq 2$$

Equation (24) indicates that realizations of the groundwater table elevation U_t can be forecasted on the basis of the previous realization U_{t-1} . This can be easily done if sample functions of the elements a' of A' are available. In the present experiment we assumed a' to be a white Gaussian noise process, that is, $E\{a'(t)\} =$

0 and $E\{a'(t_1)a'(t_2)\} = q\delta(t_1 - t_2)$, where $\delta(\cdot)$ is the Dirac's delta function and q is the variance parameter. Thus sample functions of the vector U in (24) can be computed by using generated white Gaussian noise sequences.¹⁹

From (22), assuming $N = 1$ is sufficiently accurate, we may derive an estimate of the mean of U as

$$E\{U_t\} = J(t)U_{t-1} + \int_0^t J(t-s)F(s) ds \quad (25)$$

which coincides with the deterministic solution. $E\{U_t\}$ is a $n \times 1$ vector whose elements are $E\{u_k\}$, $k = 1, \dots, n$. From (22) the two-point correlation function can be derived. Combining the correlation function and the mean, we can easily show that an estimate of the variance of U is

$$\sigma^2_{U_t} = q \int_0^t J(2t-2s)\hat{A}^2 J(2s)[U_{t-1} \cdot U_{t-1}]_{n \times n} ds \quad (26)$$

where the product of the U_{t-1} vectors is the "outer product," making a square matrix, and the matrix \hat{A}^2 contains as elements the coefficients (either 0, 1, or -2) of A' . We can approximate a variance expression consistent with the sample functions model (equation (24)) by integrating at $\Delta t = 1$ day intervals:

$$\sigma^2_{U_t} = \frac{q}{2} \{ [J(2)\hat{A}^2 [U_{t-1}^2]_{n \times n}] + [J(1)\hat{A}^2 J(2)[U_{t-1}^2]_{n \times n}] \} \quad (27)$$

where $\sigma^2_{U_t}$ is a $n \times n$ covariance matrix. This equation indicates that the head variance at any point at time t in the aquifer may be calculated on the basis of the aquifer heads at time $t - 1$.

As an illustration, simulations were conducted by assuming $S = 0.14$, $L = 10.0$ m, $\Delta x = 1.0$ m, $h_1 = h_2 = 1.0$ m, $K = 0.06$ m/day, $T = 0.06$ m²/day, and $q = 0.25^2$. Since the calculations involve the estimation of the exponential of a matrix, as well as the recursive evaluation of head vectors and variance matrices, the problem is best suited for a spreadsheet software application. This of course would not be appropriate if the number of unknowns is considerable.

Figure 5 is a graphical output implementing (24), which illustrates an arbitrary initial head versus distance and two sample functions of the total head at $t = 5$ days and at $t = 10$ days after an arbitrary recharge sequence. Figure 6 is a breakthrough curve of the groundwater head with time at $x = 5$ m. Figure 7 shows the head standard deviation with distance at three different days, as computed from (27). As expected, the maximum standard deviation occurs in the middle of the aquifer, where the heads are maximum, and zero at the boundaries, where the heads are deterministic. Finally, Figure 8 shows the mean groundwater head and the mean head plus one standard deviation with respect to time at $x = 5.0$ m.

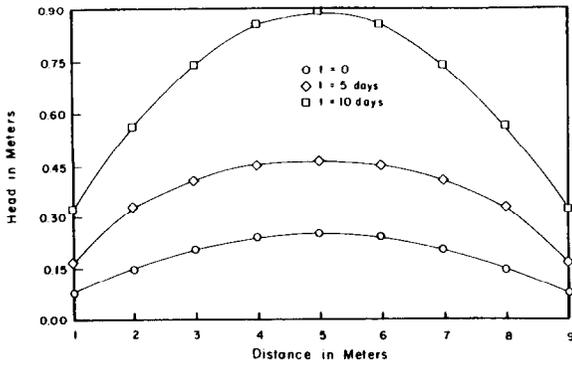


Figure 5. Sample groundwater head with distance

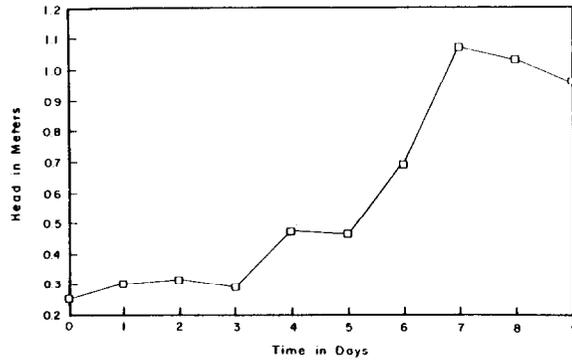


Figure 6. Sample groundwater head with time at $x = 5$ m

The above procedure is well suited for the investigation of the variability and the forecasting of groundwater flow subject to the time uncertainty associated with the transmissivity. In general, a similar methodology could be developed for the investigation of groundwater pollution when there is a time uncertainty associated with any of the parameters.

3.2 Regional groundwater flow subject to spatially variable transmissivity

Consider a situation similar to the one depicted in section 3.1, but in this case the dominant uncertainty is associated with spatial variability, rather than time variability, in the transmissivity. For instance, the series of field scale values of transmissivity, as obtained from long-term pumping tests, exhibit an erratic behavior with respect to distance, whereas they tend to be constant with respect to time at a particular location. This situation is realistic when the evolution of groundwater heads at the chosen representative time scale (i.e., one day) is very slow compared with its variability with respect to the spatial coordinates (i.e., the head changes considerably along the aquifer length). Thus the transmissivity may be represented as the random function $T(x, \omega) = \bar{T} + T'(x, \omega)$, where again \bar{T} represents the mean value and $T'(x, \omega)$ is a zero-mean spatially random process representing the variability

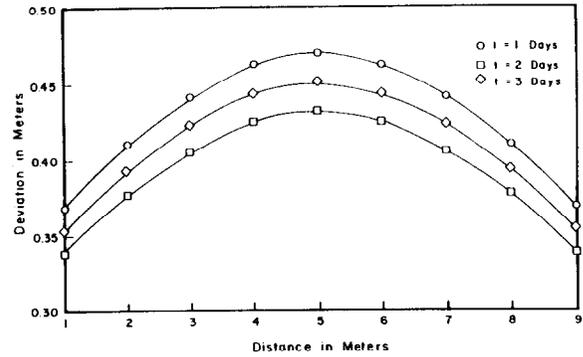


Figure 7. Head standard deviation with distance

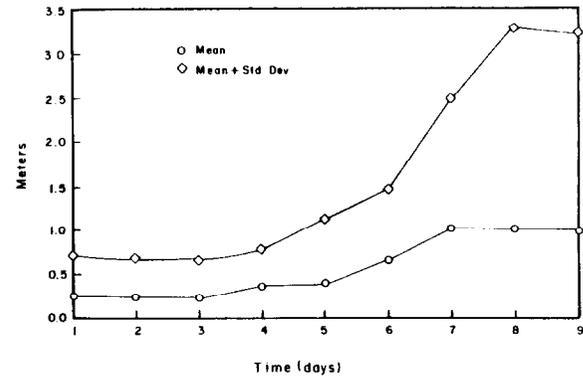


Figure 8. Mean head and mean plus one standard deviation with time at $x = 5$ m

with respect to the mean in the values of the transmissivity. Equation (17) reduces to

$$\frac{\partial u}{\partial t} - \frac{\bar{T}}{S} \frac{\partial^2 u}{\partial x^2} = \frac{i(t)}{S} + R(x, \omega)u \quad (28)$$

where the random operator R is given by

$$R(x, \omega)u = \frac{1}{S} \left(T'(x, \omega) \frac{\partial^2}{\partial x^2} + \frac{\partial T'(x, \omega)}{\partial x} \frac{\partial}{\partial x} \right) u \quad (29)$$

Since the transmissivity varies randomly with respect to distance, we may discretize the time domain in the differential equation at equal intervals separated by Δt . Replacing the time derivative by a simple backward finite difference approximation, we can write (28) as an ordinary stochastic differential equation whose independent variable is x only and whose dependent variable is the head at a fixed time t :

$$-\frac{d^2 u}{dx^2} + a^2 u = bi + a^2 u_0 + Ru \quad (30)$$

where $u_0(x)$ is the head across the aquifer at the previous time step ($t - \Delta t$), $a^2 = S/(\bar{T}\Delta t)$, $b = 1/\bar{T}$, i again is the recharge at time t in meters, and the operator R is given by (29) with the partial derivatives replaced by total derivatives. A similar differential equation would be obtained with an implicit type of discretization of the time domain.

The solution to (30) is given by

$$u(x, \omega) = \int_0^L G(x, \xi)[bi(\xi) + a^2u_0(\xi)] d\xi + \int_0^L G(x, \xi)R(\xi, \omega)u(\xi, \omega) d\xi \quad (31)$$

where $G(x, \xi)$ is the Green's function or the impulse response function of the system given by (30). It is easy to show that it is given by¹¹

$$G(x, \xi) = H(\xi - x)G_1(x, \xi) + H(x - \xi)G_2(x, \xi) \quad (32)$$

where $H(\cdot)$ is the unit step function,

$$G_1(x, \xi) = C_2(\xi) \sinh(ax) \quad (33)$$

$$G_2(x, \xi) = C_3(\xi) \cosh(ax) + C_4(\xi) \sinh(ax) \quad (34)$$

$$C_2(\xi) = C_4(\xi) \left[1 - \frac{\tanh(aL)}{\tanh(a\xi)} \right] \quad (35)$$

$$C_3(\xi) = -C_4(\xi) \tanh(aL) \quad (36)$$

$$C_4(\xi) = \frac{1}{a \tanh(aL) \left[\sinh(a\xi) - \frac{\cosh(a\xi)}{\tanh(a\xi)} \right]} \quad (37)$$

and $\sinh(\cdot)$ denotes the hyperbolic sine function. We now follow a procedure similar to that applied in the previous problems for the treatment of u in the third term of (31). Thus the general solution of (30) reduces to

$$u(x, \omega) = a \int_0^L G(x, \xi)u_0(\xi) d\xi + b \int_0^L G(x, \xi)i(\xi) d\xi + \sum_{i=1}^N \int_0^L G(x, \xi)R(\xi, \omega)u_i(\xi, \omega) d\xi \quad (38)$$

subject to

$$u_i(\xi) = \Phi(\xi) = a \int_0^L G(x, \xi)u_0(\xi) d\xi$$

which is the first term in the right side of (38), and in general,

$$u_i(\xi) = \int_0^L G(x, \xi)R(\xi, \omega)u_{i-1}(\xi) d\xi \quad i \geq 2$$

Equation (38) indicates that we can generate sample realizations of the groundwater head recursively, step by step in time. For instance, the heads across the aquifer at time t can be forecast if the observed or generated values of the head at time $t - \Delta t$ and sample functions of the processes T' and dT'/dx are available. Generally (38) is evaluated numerically using an appropriate quadrature.

For very large variances, N should be greater than or equal to 2, since the convergence rate of the series is somewhat slower. In these cases, smoothing of each iteration may be required because of the small instability generated by the numerical approximations to the derivatives of the operator R in (38). An objective way to smooth each iteration is to fit a polynomial of the form $u_{i-1} = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ and to use its algebraic expression to compute the derivatives of u_{i-1} . By choosing n points from the raw values of the previous iteration, u'_i for $i = 0, \dots, n$, forming the Vandermonde matrix²⁰ and the vector equation

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} u'_0 \\ u'_1 \\ u'_2 \\ \vdots \\ u'_n \end{bmatrix}$$

one can calculate the coefficients a_i . This can be easily done by interfacing a subroutine to solve the above vector equation with the main program doing the actual iteration for u .

For the present experiment we assumed the random component of the transmissivity to follow a colored noise process with the properties $E\{T'(x)\} = 0$ and $E\{T'(x_1)T'(x_2)\} = \rho e^{-\rho|x_1-x_2|}$. The first spatial derivative of this process has the properties¹⁹

$$E\left\{\frac{dT'(x)}{dx}\right\} = 0 \quad E\left\{\frac{dT'(x_1)}{dx_1} \frac{dT'(x_2)}{dx_2}\right\} = \frac{\partial^2}{\partial x_1 \partial x_2} E\{T'(x_1)T'(x_2)\} = \rho^2 q e^{-\rho|x_1-x_2|}$$

Sample functions of these processes can be generated by a linear ordinary differential equation forced by white Gaussian noise.¹⁵ The mean groundwater head can be computed from (38). Assuming that $N = 1$ gives an acceptable degree of accuracy (although this is not always the case), we reduce the estimate of the mean to

$$E\{u(x)\} = a \int_0^L G(x, \xi)u_0(\xi) d\xi + b \int_0^L G(x, \xi)i(\xi) d\xi \quad (39)$$

Under the same assumption the estimated head variance can be deduced from (38) as

$$\sigma_{u(x)}^2 = \int_0^L \int_0^L G(x, \xi)G(x, \gamma)E\{R(\xi)R(\gamma)\}\Phi(\xi)\Phi(\gamma) d\xi d\gamma \tag{40}$$

where the correlation of the operator R , after some manipulation, is given by

$$E\{R(\xi)R(\gamma)\} = \frac{q}{S^2} e^{-\rho|\xi-\gamma|} [\nabla_\xi^2 \nabla_\gamma^2 - \rho \nabla_\xi \nabla_\gamma - \rho \nabla_\xi \nabla_\gamma^2 + \rho^2 \nabla_\xi \nabla_\gamma] \tag{41}$$

and the operator $\nabla_\xi^2 \Phi(\xi)$ denotes $[\partial^2 \Phi(\xi)]/\partial \xi^2$. The evaluation of (40) and (41) has to be done numerically. Figure 9 illustrates the mean and the mean plus or minus one standard deviation as computed from (39) and (40), respectively, after one time step, with $q = 1.0$, $\rho = 0.1$, $\Delta t = 1.0$ day, an arbitrary smooth initial condition u_0 , and the rest of the parameters as before. A program in C was written for the calculations. The results obtained with a computer program proved more efficient with fewer memory requirements than a similar problem run in spreadsheets. However, the main advantage of using spreadsheets lies in the possibility of visualizing and controlling the partial results step by step and in the graphics capabilities. A model such as this could be used to forecast the groundwater heads across the aquifer subject to strong and erratic spatial variability in the transmissivity.

3.3 Groundwater pollution subject to spatially variable velocity field

Let us consider again the case of dispersion in an infinite aquifer subject to a space stochastic velocity field as described by (10) in section 2.1. Since $u'(x, \omega)$ is a random process in space, we may attempt a discretization of the time domain with the hope that the resulting ordinary stochastic differential equation is simpler to solve, while maintaining an accurate, stable estimation of the concentration field. A physical justification comes from the observation that the time evolution of the concentration at a point may be small

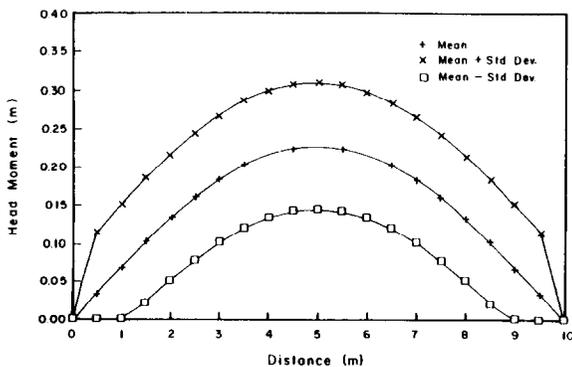


Figure 9. Mean head and mean plus and minus one standard deviation with distance

at time intervals of a few weeks, while its spatial variability is important. Proceeding as in section 3.2, we reduce (10) to

$$\frac{d^2C}{dx^2} - a \frac{dC}{dx} - bC = - \left(bC_0 - m \frac{dC}{dx} \right) \tag{42}$$

where $a = \bar{u}/D$, $b = 1/(D\Delta t)$, $m = u'/D$, $C_0(x)$ is the concentration at time $t - \Delta t$, $C(x, \omega)$ is the concentration profile at time t , and the rest of the terms are as before. The solution to this equation is given by

$$C(x, \omega) = b \int_{-\infty}^{\infty} G(x - \xi)C_0(\xi) d\xi - \sum_{i=1}^N \int_{-\infty}^{\infty} G(x - \xi)m(\xi, \omega) \frac{dC_i(\xi)}{d\xi} d\xi \tag{43}$$

subject to

$$C_1(x) = \Phi(x) = b \int_{-\infty}^{\infty} G(x - \xi)C_0(\xi) d\xi$$

and in general,

$$C_i(x) = \int_{-\infty}^{\infty} G(x - \xi)m(\xi, \omega) \frac{dC_{i-1}(\xi)}{d\xi} d\xi \quad i \geq 2$$

The Green's function, G , can be easily derived after applying the Fourier transformation with complex integration or any other suitable technique. This will give

$$G(x - \xi) = \frac{1}{2k} e^{-k|x-\xi| + a(x-\xi)/2}$$

where $k = (a^2/4) + b$.

Comparing the analytical solution of (10), that is, (11), to the corresponding semianalytical solution (43), we can observe that the semianalytical solution is simpler to evaluate.

An estimate of the mean concentration function is obtained after assuming an appropriate value of N (i.e., $N = 1$) and taking expectations on both sides of (43):

$$E\{C(x)\} = \Phi(x) \tag{44}$$

From (43) and (44) an estimate of the variance of the concentration at a point can be deduced as

$$\sigma_{C(x)}^2 = \frac{q}{D^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(x - \xi) G(x - \gamma) e^{-\rho|\xi - \gamma|} \nabla_{\xi} \Phi(\xi) \nabla_{\gamma} \Phi(\gamma) d\gamma d\xi \quad (45)$$

Comparing (15) with (45), we note again that the variance of the semianalytical solution is simpler to evaluate than that of the analytical solution. Thus we conclude that whenever a semianalytical solution is reasonably accurate in a problem, it should be attempted.

Summary and conclusions

Mathematical models that use semianalytical solutions of groundwater flow and pollution equations subject to uncertain physical parameters were discussed. The methodology combined the concept of fundamental solutions of Neumann expansions of the associated stochastic differential equation with the theory of stochastic processes. An interesting feature is that higher parameter variances not accepted by conventional small perturbation techniques can be considered. The procedure may be used to investigate the effect of erratic variability in the values of parameters of groundwater models on the behavior of the dependent variables.

Models that use the analytical solution are applicable to the general case when the uncertainty terms are random processes in time and in space. Models that use a semianalytical solution are simplified versions that take advantage of numerical methods. These models are applicable to practical cases when the random variability exists with respect to one of the independent variables while being smooth (deterministic) with respect to the other independent variables. Thus semianalytical methods offer simpler solutions (with the possibility of including irregular boundaries) than analytical methods, while producing more accurate (more stable and computationally efficient algorithms) than corresponding numerical solutions.

The methodologies presented were described with simple one-dimensional models in space. This was done for illustration purposes only, and the extension to higher dimensions should not pose much difficulty. For additional ideas in this respect the reader is referred to Ref. 21.

Acknowledgments

Special thanks to Mr. Jiyang Liang from the Chinese Academy of Sciences, Beijing, China, for his extremely valuable contributions and suggestions in the early stages of the research, particularly in section 3.1. The support of the Department of Civil Engineering of the University of Kentucky, the U.S. Geological Survey through Kentucky Water Resources Research Institute grant G1564 89-04, and National Science Foundation grant BCS-892-1004 is greatly appreciated.

References

- 1 Warren, J. E. and Skiba, F. F. Macroscopic dispersion. *Trans. Am. Inst. Min. Metall. Pet. Eng.* 1964, **231**, 215-230

- 2 Heller, J. P. Observations of mixing and diffusion in porous media. *Proceedings of the Second Symposium on Fundamental Transport Phenomena in Porous Media*, 1972
- 3 Winter, C. L., Newman, C. M., and Neumann, S. P. A perturbation expansion for diffusion in a random velocity field. *SIAM J. Appl. Math.* 1984, **44**(2), 411-424
- 4 Matheron, G. and De Marsily, G. Is transport in porous media always diffusive? *Water Resour. Res.* 1980, **16**, 901-917
- 5 Smith, L. and Schwartz, F. W. Mass transport. 1. A stochastic analysis of macroscopic dispersion. *Water Resour. Res.* 1980, **16**, 303-313
- 6 Dagan, G. Solute transport in heterogeneous porous formations. *J. Fluid Mech.* 1984, **145**, 151-177
- 7 Tang, D. H., Schwartz, F. W., and Smith, L. Stochastic modelling of mass transport in a random velocity field. *Water Resour. Res.* 1982, **18**(2), 231-244
- 8 Sposito, G. and Jury, W. A. Fundamental problems in the stochastic convection-dispersion model of solute transport in aquifers and field soils. *Water Resour. Res.* 1986, **22**(1), 77-88
- 9 Cushman, J. H. Development of stochastic partial differential equations in subsurface hydrology. *J. Stochastic Hydrol. Hydraul.* 1987, **1**(4), 241-262
- 10 Serrano, S. E. General solution to random advective-dispersive transport equation in porous media. Part 2: Stochasticity in the parameters. *Stochastic Hydrol. Hydraul.* 1988, **2**(2), 99-112
- 11 Serrano, S. E. Stochastic differential equation models of erratic infiltration. *Water Resour. Res.* 1990a, **26**(4), 703-711
- 12 Sahuquillo, A. An eigenvalue numerical technique for solving unsteady linear groundwater flow models continuously in time. *Water Resour. Res.* 1983, **19**(1), 87-93
- 13 Serrano, S. E. Semigroup solutions to stochastic unsteady groundwater flow subject to random parameters. *Stochastic Hydrol. Hydraul.* 1987, **1**(4), 281-296
- 14 Bear, J. *Hydraulics of Groundwater*. McGraw-Hill, New York, 1979
- 15 Serrano, S. E. Modelling infiltration in hysteretic soils. *Adv. Water Res.* 1990b, **13**(1), 12-23
- 16 Moler, C. and Van Loan, C. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Rev.* 1978, **20**(4), 801-836
- 17 Umari, A. M. J. and Gorelick, S. M. The problem of complex eigensystems in the semianalytical solution for advancement of time in solute transport simulations: A new method using real arithmetic. *Water Resour. Res.* 1986, **22**(7), 1149-1154
- 18 Kraus, A. D. *Matrices for Engineers*. Hemisphere, Washington, DC, 1987
- 19 Jazwinski, A. H. *Stochastic Processes and Filtering Theory*. Academic, San Diego, CA, 1970
- 20 Chapra, S. C. and Canale, R. P. *Numerical Methods for Engineers*, 2nd ed. McGraw-Hill, New York, 1988
- 21 Liron, N. and Rubinstein, J. Calculating the fundamental solution to linear convection-diffusion problems. *SIAM J. Appl. Math.* 1984, **44**(3), 493-511

Appendix: convergence theorem for large variances

In this section we attempt to demonstrate that the Neumann expansion used to approximate solutions of stochastic partial differential equations in groundwater flow and transport, which is the main tool in the present article, converges uniformly even when relatively large variances in the random parameters are present. The following theorem concentrates on a one-dimensional

spatial domain. It is easy to show that similar theorems can be constructed for the general three-dimensional

domain and for the case of more than one random parameter.

Let us rewrite (11) as

$$C(x, t, \omega) = J(x, t)C_0 - \int_0^t J(x, t - \tau)u'(x, \omega) \sum_{n=1}^{\infty} \phi_n(x, \tau, \omega) d\tau \quad t \in [0, T] \quad (A1)$$

where the semigroup operator, $J(x, t)$, is given by (2) and the series $\phi_n(x, \tau, \omega) = \partial C_n / \partial x$. As described in section 2.2 (A1) is subject to

$$\phi_1(x, \tau, \omega) = \frac{\partial}{\partial x} [J(x, \tau)C_0] \quad (A2)$$

$$\phi_{n+1}(x, \tau, \omega) = - \int_0^{\tau} J(x, \tau - \gamma)u'(x, \omega)\phi_n(x, \gamma, \omega) d\gamma \quad (A3)$$

Theorem

If $|u'(x, \omega)| \leq M$ for almost every $\omega \in \Omega$ and $\|\cdot\|_2$ is the norm of the space¹⁰

$$L^2(0, T; V) = \{f : [0, T] \rightarrow V : \int_0^T \|f\|_V^2 dt < \infty\}$$

where $V = H^1$ is the first-order Sobolev space of $L^2(\Omega)$ -valued random functions, then a sufficient condition for the almost sure convergence of the series $\sum_0^{\infty} \phi_n(x, \tau, \omega)$ in the space $L^2(0, T; V)$ is that $t(M/2) < 1$.

Proof

It follows from the integral form of the Minkowski inequality that

$$\begin{aligned} \|\phi_{n+1}(x, \tau, \omega)\|_2 &= \left\| - \int_0^{\tau} J(x, \tau - \gamma)u'(x, \omega)\phi_n(x, \gamma, \omega) d\gamma \right\|_2 \\ &\leq M \left\| \int_0^{\tau} J(x, \tau - \gamma)\phi_n(x, \gamma, \omega) d\gamma \right\|_2 \\ &\leq M \|\phi_n(x, \tau, \omega)\|_2 \left\| \int_0^{\tau} J(x, \tau - \gamma) d\gamma \right\|_2 \\ &\leq M \|\phi_n(x, t, \omega)\|_2 \sup_{x \in V; \tau, \gamma \in [0, T]} \int_0^{\tau} |J(x, \tau - \gamma) d\gamma| \\ &\leq M \|\phi_n(x, t, \omega)\|_2 \int_0^t J(x, t - \gamma) d\gamma \\ &\leq M \|\phi_n(x, t, \omega)\|_2 \cdot \frac{1}{2} \int_0^t d\gamma \\ &\leq M \|\phi_n(x, t, \omega)\|_2 \cdot \frac{t}{2} \end{aligned}$$

Since $L^2(0, T; V)$ is a complete space, the convergence of $\sum_0^{\infty} \phi_n(x, \tau, \omega)$ follows from the convergence of $\sum_0^{\infty} \|\phi_n(x, t, \omega)\|_2$. If $Mt/2 < 1$, the convergence of the Neumann series can be used to construct approximate solutions to (10). Q.E.D.

The above result indicates that it is always possible to manipulate the simulation time step in order to consider high variability in the random parameter and assure convergence. For instance, if $\bar{u} = 0.1$ m/day with

a $\pm 80\%$ value of maximum measured variability, that is, $u' = \pm 0.08$ m/day maximum, then simulation time steps $t < 25.0$ days will generate an almost sure convergent Neumann series. Note that in the example presented in section 2.1 this condition is slightly violated and yet convergence is achieved in some weaker sense.