Chapter 10

A Stochastic Inverse Method to Estimate Parameters in Groundwater Models

10.1 Introduction

Natural systems are heterogeneous and they contain noise due to random inputs, irregular varying coefficients and fluctuations in boundary conditions. In this chapter, we model the behavior of natural systems using stochastic differential equations, present a parameter estimation procedure for such models in a general setting, and extend it to simple groundwater models. The applications to groundwater models are within the context of one dimensional solute transport problem to estimate parameters for two governing equations, one consisting of a single parameter and other of two parameters. The results of this inverse methodology are reliable in the presence of noise. However, the investigation of solute transport parameter estimates shows an inverse relationship to the noise level. The main advantage of the estimation methodology presented here is its direct dependence on field observations of state variables of natural systems in the presence of uncertainty.

As we have seen before, we can model the behavior of natural systems such as groundwater flow and solute transport in porous media through differential equations based on conservation laws. In the process of developing the differential equations, we introduce the parameters, which we consider attributes or properties of the system. In the case of groundwater flow, for example, the parameters such as hydraulic conductivity, transmissivity and porosity are constant within the differential equation, and it is often necessary to assign numerical values to these parameters. These values of the parameters are obtained from laboratory experiments and/or field scale experiments. However, these values may not represent the often complex patterns across a large geographic area, hence limiting the effectiveness of the model. In addition, such field scale experiments can be expensive. Often we are interested in modeling for quantities such as the depth of water table and solute concentration. This is because they are directly relevant to environmental decision making, and we measure these variables regularly and the measuring techniques tend to be cheaper. Further we can continuously monitor these decision (output) variables in many situations. Therefore it is reasonable to assume that these observations of the output variables represent current status of the system. If the dynamics of the system can reliably be modeled by a relevant differential equation, we can expect the parameters estimated based on the observations may give us more reliable representative values than those obtained from laboratory tests and literature. However, such observations often contain noise from two different sources: experimental errors and noisy system dynamics. Noise in the system dynamics may be due to heterogeneity of the media, random nature of inputs such as rainfall and variable boundary conditions to name a few factors. The question of estimating the parameters from the observations naturally involves the models that represent the system noise as well. In this Chapter, we aim to illustrate a parameter estimation procedure for such models containing noise. We present models in a general setting so that the procedure can be used for groundwater models. Then we apply parameter estimation theory and procedures to the solute transport in saturated porous media in the presence of noise.

10.2 System Dynamics with Noise

Let us consider a differential equation of the form,

$$\frac{dy}{dt} = f(\theta, y, t) \tag{10.1}$$

where y is the dependent variable (output) that is observed, t is time and θ is a parameter upon which the characteristics of the model depends. Suppose we can include the noise ($\xi(t)$) contained in dy/dt as an additive component to equation (10.1). For simplicity we will assume that ξ depends only on time. In many engineering and natural systems, this noise is irregular, continuous and independent of each other, and white noise has been considered as a valid approximation (Øksendal, 1998). Therefore, we can express equation (10.1) as

$$\frac{dy}{dt} = f(\theta, y, t) + \xi(t) , \tag{10.2}$$

and multiplying by dt we get

$$dy = f(\theta, y, t)dt + \xi(t)dt. \tag{10.3}$$

Once we consider white noise as a model for the noise term, y becomes a stochastic process having many realizations or paths. A set of observations of y can be considered as a realization of y. Considering that the noise represents deviations away from the deterministic rate, the expected value of $\xi(t)$ over all realizations is zero; $\xi(t)$ is an independent stochastic process; the joint probability distribution of $\xi(t)$ is time-invariant; and $\xi(t)$ has to be continuous though irregular. The only stochastic process that can meet all these requirements is the Wiener process ($B(t)$) based on observations of the Brownian motion (Øksendal, 1998). It can be shown that,

$$dB(t) = \xi(t)dt, \tag{10.4}$$

taking the convergence in the probability.

In equation (10.4), $dB(t)$ are increments of the standard Wiener process which are normally distributed with a unit variance (for a detailed discussion refer to the chapters 2, 3 and 4).

Substituting in equation (10.3) we have,

$$dy = f(\theta, y, t)dt + dB(t); \quad 0 \le t \le T. \tag{10.5}$$

This is a stochastic differential equation giving the drift term ($f(\theta, y, t)$) and diffusive term ($dB(t)$). Kutoyants (1984), for example, gives the likelihood function $L(\theta)$ to estimate θ given the observation for y under certain conditions:

$$L(\theta) = \exp\left\{\int_0^T f(\theta, y, t)dy(t) - \frac{1}{2}\int_0^T f^2(\theta, y, t)dt\right\}. \tag{10.6}$$

Taking natural logarithm of both sides, the log likelihood is given by

$$l(\theta) = \int_0^T f(\theta, y, t)dy(t) - \frac{1}{2}\int_0^T f^2(\theta, y, t)dt. \tag{10.7}$$

By maximizing $l(\theta)$ with respect to θ we obtain,

$$\frac{\partial l(\theta)}{\partial \theta} = \int_0^T \frac{\partial}{\partial \theta}(f(\theta, y,t))dy(t) - \int_0^T f(\theta, y,t) \frac{\partial}{\partial \theta}(f(\theta, y,t))dt = 0 . \tag{10.8}$$

Equation (10.8) will give the maximum likelihood parameter ($\hat{\theta}$) for θ, given the values of y. We will illustrate the use of equation (10.8) by taking an example.

10.2.1 An Example

Suppose that the dynamics of a system could be expressed by,

$$\frac{dX(t)}{dt} = \theta \ X(t) + \xi(t); \quad X(0) = 1, \quad 0 \le t \le 1 \tag{10.9}$$

where $X(t)$ is the process under observation, θ is a parameter to be determined from the observations, and $\xi(t)$ is the noise component assumed to be white. Following the arguments mentioned in the previous section, we can express the process $X(t)$ in terms of a stochastic differential:

$$dX(t) = \theta \ X(t)dt + dB(t). \tag{10.10}$$

Following the definition of Ito integral (Øksendal, 1998) we can explicitly solve equation (10.10) and the solution can be expressed in terms of Wiener process, $B(t)$:

$$X(t) = X(0) \ \exp(B(t))\exp((\theta - 1/2)t) \tag{10.11}$$

where $B(t)$ is the standard Wiener process. The solution of equation (10.11) consists of a set of realizations of $X(t)$; and as an example, a realization of $X(t)$ is given in Figure 10. 1 for $\theta = 1.5$. Let us assume that we observe the realization of $X(t)$ depicted in Figure 10.1 and we seek to estimate θ given $X(t)$ and corresponding time values.

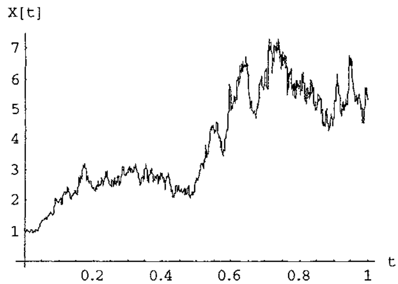

Figure 10.1 A realization of equation (10.11) for θ =1.5.

In this case, comparing to equation (10.5)

$$f(\theta,\ y,\ t) = \theta\ \ X(t)$$

and equation (10.8) can be expressed as

$$\frac{\partial l(\theta)}{\partial \theta} = \int_0^1 X(t)dX(t) - \int_0^1 \theta X(t)X(t)dt,$$
$$= \int_0^1 X(t)dX(t) - \int_0^1 \theta X^2(t)dt.$$

To maximize the log likelihood,

$$\frac{\partial l(\theta)}{\partial \theta} = 0.$$

Therefore, $\hat{\theta} = \dfrac{\int_0^T X(t)dX(t)}{\int_0^T X^2(t)dt}.$ (10.12)

From the given observations (Figure 10.1) we express $\hat{\theta}$ as

$$\hat{\theta} \approx \frac{\sum_{i=1}^{M} X(t_i) \ \ \Delta X(t_i)}{\sum_{i=1}^{M} X^2(t_i) \ \ \Delta t}. \tag{10.13}$$

The estimated value of θ $(\hat{\theta})$ for the given realization is 1.47849. Similarly, to investigate the robustness of the procedure, we compute $\hat{\theta}$ from 30 different realizations of $X(t)$. Calculated from equation (10.11) with $\theta = 1.5$, the mean of $\hat{\theta}$ is 1.48135 with a standard deviation of 0.586255. This shows that if we sample the $X(t)$ process from equation (10.11) a reasonable number of times, the mean value of $\hat{\theta}$ is very close to 1.5.

It would be interesting to see what happens when the standard Wiener increment term ($dB(t)$) in equation (10.5) is modified by an amplitude (σ^2):

$$dy = f(\theta, \ y, \ t)dt + \sigma^2 \ dB(t). \tag{10.14}$$

Equation (10.12) is used to calculate θ, and Table 10.1 shows the mean and the standard deviations of $\hat{\theta}$ based on 30 distinct realizations of $X(t)$ from equation (11) with $\theta = 1.5$. It is evident that an amplitude of 1.0 or less, and slightly above 1.0 would produce reliable estimates from this procedure.

Table 10.1 Mean and standard deviations of parameter estimates.

	$\hat{\theta}$	
σ^2	Mean	Std. Deviation
0.01	1.47319	0.117916
0.10	1.55374	0.234585
0.25	1.39394	0.385624
0.50	1.52886	0.426577
1.00	1.48135	0.586255
1.50	1.76876	1.424700

10.3 Applications in Groundwater Models

In this section, the above described general parameter estimation procedure is applied in the context of solute transport in saturated porous media in the presence of noise. Unny (1989) presented the basis for this application by describing groundwater system in the form of stochastic partial differential equations and then estimating parameters.

10.3.1 Estimation Related to One Parameter Case

The stochastic one-dimensional advective transport equation can be expressed as,

$$\frac{\partial C}{\partial t} = -v_x \left(\frac{\partial C}{\partial x} \right) + \xi(x,t) \tag{10.15}$$

where

v_x = average linear velocity, m/day,
C = solute concentration, mg/l,
and $\xi(x, t)$ is described by a zero-mean stochastic process.

We multiply equation (10.15) by dt throughout and as in equation (10.14), formally replace $\xi(x, t)dt$ by $\sigma^2 dB(t)$ (see Unny (1989) for the derivation). Now, we can obtain the stochastic partial differential equation as follows,

$$\partial C = -v_x \left(\frac{\partial C}{\partial x}\right)dt + \sigma^2 dB(t). \tag{10.16}$$

Suppose we have observations of solute concentration, C_i at M independent space coordinates along x-axis, where $1 \le i \le M$, at different time intervals, t (where $0 \le t \le T$). In other words we have M number of C_i observations for each time step. Hence, altogether, there are $((T+1)*M)$ number of C_i observations. We use these observations to estimate the parameter θ, which, in this case v_x (or hydraulic conductivity, if hydraulic gradient and porosity are known), of all possible parameter values using maximum likelihood approach. As we explained above, we can write equation (10.16) in the form of equation (10.14),where

$$f(\theta, C, t) = -v_x \left(\frac{\partial C}{\partial x}\right).$$

The likelihood expression for the estimation of parameter θ can be given by

$$L(\theta) = \exp\left\{\int_0^T f(\theta, C, t)\, dC(t) - \frac{1}{2}\int_0^T f^2(\theta, C, t)dt\right\}. \tag{10.17}$$

The estimate $\hat{\theta}$ can be obtained by maximizing $L(\theta)$; therefore,

$$\frac{\partial L(\theta)}{\partial \theta} = 0. \tag{10.18}$$

If, $l(\theta) = \ln L(\theta)$, (10.19)

taking the natural log on both sides of the equation (10.17)

$$l(\theta) = \int_0^T f(\theta, C, t)\, dC(t) - \frac{1}{2}\int_0^T f^2(\theta, C, t)dt. \tag{10.20}$$

The parameter is estimated as the solution to the equation

$$\int_0^T \frac{\partial}{\partial\theta} f(\theta,\ C,\ t)\ dC(t) - \int_0^T f(\theta,\ C,\ t)\frac{\partial}{\partial\theta} f(\theta,\ C,\ t)dt = 0\,. \tag{10.21}$$

If we observe M independent sample paths, the likelihood-function becomes the product of the likelihood functions for M individual sample paths,

$$L(\theta) = L(\theta,\ C_1)\, L(\theta,\ C_2) \ldots\ldots L(\theta,\ C_M). \tag{10.22}$$

Taking the natural log on both sides of the equation (10.22) we have the log-likelihood,

$$l(\theta) = l(\theta,\ C_1) + \ l(\theta,\ C_2) + \ldots\ldots + l(\theta,\ C_M). \tag{10.23}$$

Therefore, the log likelihood function can be expressed as

$$l(\theta) = \sum_{i=1}^{M} \int_0^T f(\theta,\ C_i,\ t)\ dC_i(t) - \frac{1}{2}\sum_{i=1}^{M} \int_0^T f^2(\theta,\ C_i,\ t)dt\,, \tag{10.24}$$

and the parameter estimate $\hat{\theta}$ is obtained as the solution to maximum likelihood

$$\sum_{i=1}^{M}\int_0^T \frac{\partial f(\theta,\ C_i,\ t)}{\partial\theta}\ dC_i(t) - \sum_{i=1}^{M}\int_0^T f(\theta,\ C_i,\ t)\frac{\partial f(\theta,\ C_i,\ t)}{\partial\theta} dt = 0\,. \tag{10.25}$$

Let us assume that the drift term in equation (10.14), $f(t, C, \theta)$, depends linearly on its parameters θ, then we can express it as

$$f(\theta,\ C,\ t) \ = \ a_0(C,t) \ + \theta \ \ a_1(C,\ t). \tag{10.26}$$

The log-likelihood function from equation (10.24) is

$$l(\theta_1) = \sum_{i=1}^{M} \int_0^T \{a_0(C_i,t) + \theta a_1(C_i,t)\} dC_i(t) - \frac{1}{2} \sum_{i=1}^{M} \int_0^T \{a_0(C_i,t) + \theta a_1(C_i,t)\}^2 \, dt. \qquad (10.27)$$

The estimate $\hat{\theta}$ is obtained as a solution to the equation

$$\sum_{i=1}^{M} \int_0^T \{a_1(C_i,t)\} dC_i(t) - \sum_{i=1}^{M} \int_0^T \left\{a_0(C_i,t) + \theta a_1(C_i,t)\right\} \{a_1(C_i,t)\} dt = 0. \qquad (10.28)$$

Hence the estimate of $\theta(\hat{\theta})$ is given by

$$\hat{\theta} = \frac{\sum_{i=1}^{M} \int_0^T \{a_1(C_i,t)\} dC_i(t) - \sum_{i=1}^{M} \int_0^T \{a_0(C_i,t)\}\{a_1(C_i,t)\} dt}{\sum_{i=1}^{M} \int_0^T \{a_1^2(C_i,t)\} dt}. \qquad (10.29)$$

When we compare equations (10.16) and (10.26), we have

$$a_0(C_i,\ t) = 0; \quad a_1(C_i,\ t) = -\left(\frac{\partial C}{\partial x}\right)_i; \ \theta = v_x.$$

Therefore, the estimate of v_x is given by,

$$\hat{v}_x = \frac{\sum_{i=1}^{M} \int_0^T -\left(\frac{\partial C}{\partial x}\right) dC_i(t)}{\sum_{i=1}^{M} \int_0^T \left\{\left(\frac{\partial C}{\partial x}\right)_i\right\}^2 dt}. \qquad (10.30)$$

10.3.2 Estimation Related to Two Parameter Case

We can use the same theoretical basis to estimate two parameter space problems. As an example, let us consider a one-dimensional stochastic advection-dispersion equation, which is given by

$$\frac{\partial C}{\partial t} = D_L \left(\frac{\partial^2 C}{\partial x^2}\right) - v_x \left(\frac{\partial C}{\partial x}\right) + \xi(x,t), \tag{10.31}$$

where D_L is the longitudinal hydrodynamic dispersion coefficient, m^2/day. Two parameters to be estimated are D_L and v_x. Equation (10.26) can be written in the following form:

$$f(t,C,\theta) = a_0(C,t) + \theta_1 a_1(C,t) + \theta_2 a_2(C, t). \tag{10.32}$$

In a similar way to the one parameter problem, we can compare equation (10.32) and the drift term of equation (10.31):

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

The log-likelihood function from equation (10.21) is

$$\begin{aligned} l(\theta_1,\theta_2) = &\sum_{i=1}^{M}\int_0^T \{a_0(C_i,t) + \theta_1 a_1(C_i,t) + \theta_2 a_2(C_i,t)\}\, dC_i(t) \\ &- \frac{1}{2}\sum_{i=1}^{M}\int_0^T \{a_0(C_i,t) + \theta_1 a_1(C_i,t) + \theta_2 a_2(C_i,t)\}^2\, dt. \end{aligned} \tag{10.33}$$

Differentiating (10.33) with respect to θ_1 and θ_2 respectively we get the following two simultaneous equations:

$$\sum_{i=1}^{M}\int_{0}^{T}\{a_1(C_i,t)\}dC_i(t)-\sum_{i=1}^{M}\int_{0}^{T}\{a_0(C_i,t)+\theta_1 a_1(C_i,t)+\theta_2 a_2(C_i,t)\}\{a_1(C_i,t)\}dt=0. \quad (10.34a)$$

$$\sum_{i=1}^{M}\int_{0}^{T}\{a_2(C_i,t)\}dC_i(t)-\sum_{i=1}^{M}\int_{0}^{T}\{a_0(C_i,t)+\theta_1 a_1(C_i,t)+\theta_2 a_2(C_i,t)\}\{a_2(C_i,t)\}dt=0. \quad (10.34b)$$

Now we obtain the values for $\hat{\theta}_1$ and $\hat{\theta}_2$ as the solutions to these two equations.

10.3.3 Investigation of the Methods

We use the above-mentioned method to estimate parameters in equations (10.30) and (10.34) by using a noisy dataset. The one dimensional solute transport dataset was generated by using equation (10.15) for one parameter case and equation (10.31) for the two parameter case. First, data was generated by using the deterministic solutions for each case and then noise was added randomly to each deterministic concentration value to generate a stochastic dataset. As an example, in the case of a maximum of ±5% introduced randomness, the noise component was generated by a random function which gives a maximum of 5% of deterministic concentration and another randomness function selects + or – operation. The spatial domain of the solution is 10m ($0 \le x \le 10$).

10.4 Results

The example in section 2.1 shows that the parameter estimation methodology described in this Chapter produces reliable estimates for a noisy dynamic system. The expected values of the estimates are closer to the actual parameters at low noise levels. As the percentage of noise is increased by changing the σ^2, the difference between actual and estimated parameters becomes larger. However, it is interesting to see that the mean value shows a close correlation to the actual value though the standard deviation increases with the noise (Table 10.1).

We present only a sample of results for the simulation study of solute transport. Figure 10.2 shows the estimated average linear velocity, v_x (0.3), that was used to generate the deterministic solution, against the actual parameter value for one parameter case. Figure 10.3 shows the comparison of the longitudinal dispersion coefficient, D_L in the two parameter estimation.

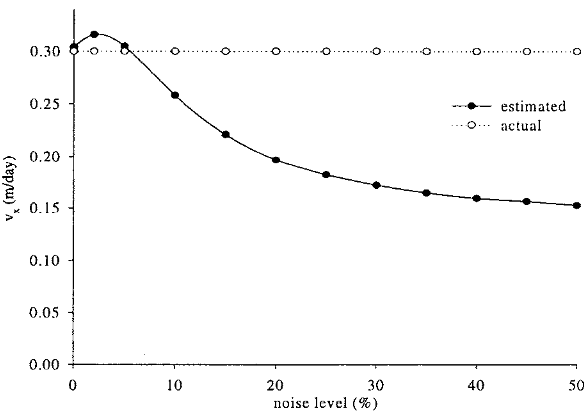

Figure 10.2 Actual and estimated velocity for different noise levels for one-parameter case.

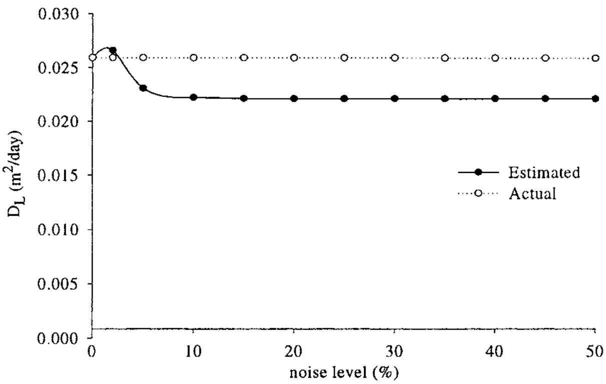

Figure 10.3 Actual and estimated longitudinal dispersion coefficient (D_L) for different noise levels in two-parameter case.

As seen in Figure 10.2 and Figure 10.3, the deviations of the estimated parameters from the corresponding actual values increase at first and then begin to flatten as the noise level increases. For example, the onset of flattening is 5% in Figure 10.2 whereas it is 2% in the Figure 10.3.

10.5 Concluding Remarks

In this Chapter, we have shown a straightforward procedure to estimate parameters of stochastic differential equations, which model the dynamics of systems containing noise. A sample of results has been discussed in two different cases to show that the likelihood functions give reasonable results even with significant levels of noise contained in the data. This procedure can be extended to the cases where the amplitude of noise is non-linear, but it is beyond the scope of this chapter.