STOCHASTIC MODELLING OF SOLUTE TRANSPORT IN GROUNDWATER: APPLICATION TO A FIELD TRACER TEST

W.D. GRAHAM
Dept. of Agricultural Engineering, University of Florida, Gainesville, Florida 32611-0361, USA
D. McLAUGHLIN
Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

ABSTRACT Field data from the Borden, Ontario natural gradient tracer test are used to demonstrate the performance of the stochastic transport model developed in Graham & McLaughlin (1989a,b). This stochastic model provides 1) a prediction of the extent of the solute plume, 2) an estimate of the reliability of this prediction, and 3) a means of combining physically-based models with field data using the Bayesian concepts of measurement conditioning. Results show that the unconditional mean plumes are quite successful in predicting the peak concentrations and average spreading of the actual tracer plumes. There is still significant uncertainty associated with the mean prediction, however, which is estimated accurately by the unconditional concentration standard deviation. Some improvement in model prediction is achieved by conditioning the concentration moments using a small subset of the available concentration observations.

INTRODUCTION

In this paper results from a large-scale natural gradient tracer test conducted at a site in Borden Ontario (Canada) are used to evaluate the feasibility of applying the stochastic groundwater transport model developed in Graham & McLaughlin (1989a,b) to a realistic field problem. In the Borden experiment the trajectory of a pulse injection of known masses of inorganic and organic compounds was monitored extensively in three dimensions over a three year period (Mackay et al., 1986). The spatial variability of hydraulic conductivity at the site was also examined in detail by measuring the permeability of a series of soil cores taken along two cross-sections near the tracer plume (Sudicky, 1986).

Several researchers have used the Borden data to evaluate stochastic theories of subsurface transport. Freyberg (1986) evaluated the second spatial moments of the depth averaged tracer plume as a function of time, using point observations of solute concentration. He found that the theoretical two-dimensional ensemble spatial moments derived by Dagan (1984) calibrated well to the estimated values for the first 647 days of transport. Sudicky (1986) estimated the variance and correlation scale of the hydraulic conductivity field at the Borden site based on permeabilities measured from approximately 1280 soil cores. He then evaluated the Gelhar & Axness (1983) asymptotic ensemble longitudinal macrodispersivity and the Dagan (1984) time varying two-dimensional ensemble spatial moments using the estimated hydraulic conductivity statistics and found them both to be consistent with the spread of the injected tracer.

Barry et al. (1988) used the two dimensional Dagan model for the ensemble mean concentration to predict the extent of the chloride and bromide plumes at the Borden site.
They found that uncertainty in the specification of the initial condition, incomplete sampling of the plume during early sampling sessions, and assumptions required to convert the three dimensional discrete sampling data into a continuous vertically averaged form hindered a thorough evaluation of the Dagan model. Model predictions of the ensemble mean concentration computed using surfaces generated for some of the early sampling sessions as initial conditions were shown to be reasonably consistent with field concentration data for some of the later sampling sessions. However with increasing time the model tended to "over-disperse" the plume and the features of the measured data became less well reproduced. Barry et al. (1988) conclude that predictions based on the Dagan model are likely to be highly uncertain. However a formal analysis of prediction uncertainty was not performed.

In this paper the two-dimensional stochastic transport model developed in Graham & McLaughlin (1989a,b) is used to predict the trajectory of the chloride tracer plume at the Borden site. This model provides a convenient method for predicting the movement of solute plumes in heterogeneous aquifers where hydrogeological properties are unknown or highly uncertain, and differs from traditional methods in several fundamental ways:

First, the model's primary variables are the ensemble mean and the covariance of solute concentration. These probabilistic moments are derived from a stochastic version of the solute transport equation which explicitly recognizes that small-scale velocities at real sites are uncertain. The model's description of velocity uncertainty is related to soil heterogeneity and to the large-scale hydrologic features which control regional groundwater flow.

Second, the model does not assume that field-scale dispersion is Fickian. It derives the macrodispersive flux (velocity-concentration covariance) directly from the statistics of the velocity field and the stochastic transport equation. There is no need to specify or derive a macrodispersivity coefficient.

Third, the model's predictions (concentration mean and covariance) can be conditioned on field observations of hydraulic conductivity, head, and/or concentration whenever such measurements become available. In the application discussed here, predictions of the Borden solute plume are conditioned on vertically-averaged concentration at three times. The conditioning process transforms the regularly shaped plumes obtained from unconditional predictions into more realistic irregular plumes.

Fourth, the conditional covariances computed by the model provide a convenient basis for designing field sampling programs and monitoring networks. Conditioning integrates modelling with data collection and provides a systematic conceptual framework for carrying out a site investigation. We believe that the concept of using measurement conditioning to combine physically-based models with field data will prove to be useful in many water resource applications. The Borden case study provides some useful insight into the strengths and limitations of this concept.

In the following sections the foundation and structure of the stochastic model are discussed. References, in which further details of the development and solution of the model equations may be found, are given. The unconditional ensemble moments for the chloride tracer plume at the Borden site are computed based on the initial solute mass and configuration described by Mackay et al. (1986), and the statistics of the hydraulic conductivity field determined by Sudicky (1986). Conditional moments are also evaluated using a small subset of the available concentration observations at three sampling times. At each time the updated conditional means are compared to the unconditional mean plume and the actual plume to evaluate the performance of the stochastic model.
FORMULATION OF THE MODELLING APPROACH

The basic objective of the model applied in this paper is to provide an accurate estimate of the two-dimensional vertically averaged solute concentration at any specified time and location. Accuracy is assumed to be limited by the inherent variability of the subsurface environment which can, at best, only be observed at a limited number of locations. If a probabilistic approach is used to describe this variability, the concepts of Bayesian estimation theory suggest that the best estimate of solute concentration is the conditional ensemble mean (Graham & McLaughlin, 1989a,b). The conditional ensemble covariance is a useful measure of the accuracy and correlation (over time and space) of this estimate. In both cases, the term "conditional" indicates that the quantity in question takes into account the information contributed by field observations of concentration and related hydrologic variables. In the special case when no measurements are available, conditioning is not possible and the appropriate estimates are "unconditional" moments derived solely from generic prior information.

If the measurements used for conditioning are available only at discrete times but the solute plume of interest is changing continuously over time, the conditional mean and covariance can be approximated by a two-stage recursive algorithm commonly known as the extended Kalman Filter (Jazwinski, 1978; Bencala & Seinfeld, 1979; Ljung 1983; Graham & McLaughlin, 1989b). The derivations and equations for the Kalman filtering algorithm we use to estimate solute concentration are discussed in Graham & McLaughlin (1989b). The model divides naturally into two parts: 1) the moment propagation equations, which describe the evolution of the conditional concentration moments between measurement times \( t_n \) and \( t_{n+1} \), and 2) the moment update equations, which account for the new information provided by measurements obtained at time \( t_{n+1} \). The propagation and moment equations are solved recursively, over a specified time period. At each discrete time within this period, estimates of the ensemble mean and covariance are provided throughout the region of interest.

The stochastic solute transport model requires the solution of a coupled set of partial differential equations between update times and solution of a coupled set of algebraic equations at update times. The resulting computational demands can be very great unless the numerical procedure used to solve the model's differential equations is selected carefully. The results generated in this study were computed with the dual-grid finite-element approach described in Graham & McLaughlin (in press). The use of a dual-grid approach makes it feasible to carry out a two-dimensional simulation of the Borden tracer test on a typical scientific workstation. More efficient numerical methods will probably be required before large three-dimensional problems can be solved on readily available computers. Such methods should be designed to take advantage of the special structure of the moment equations and to minimize the computation of intermediate variables which are not of direct interest.

RESULTS

To compare results from the two-dimensional stochastic transport model to the Borden chloride plume, the three-dimensional chloride concentration measurements must be averaged over the vertical dimension. For this purpose a uniform 6 m averaging interval extending from 1.5 to 7.5 m below ground reference was adopted and a trapezoidal integration scheme was used in accordance with the depth averaging procedure described by Freyberg (1986). For locations where the uppermost or lowermost vertical sampling point measured a concentration greater than background, the concentration was assumed
to reach background within a distance less than or equal to twice the vertical sampling interval. Similarly, in the horizontal dimension if the outermost sampling wells measured greater than background concentration, the concentration was assumed to reach background at the next (unsampled) row of monitoring wells.

The experimental data gathered at the Borden site show significant vertical variability and extremely small vertical spreading. The total vertical plume displacement was small, however the vertical trajectory was found to be slightly concave upwards. These three-dimensional features will obviously be lost in a two-dimensional analysis. Nevertheless the depth averaged moments will still contain valuable information regarding more gross features such as the overall trajectory, areas of maximum concentration, and horizontal extent of the solute plume.

Table 1 summarizes the input parameters used to simulate the Borden plume. The vertically averaged log hydraulic conductivity variance and the horizontal log hydraulic conductivity correlation scale are taken from the Sudicky analysis (Sudicky, 1986). The values for the mean hydraulic gradient, geometric mean hydraulic conductivity and porosity are those used by Frind et al. (1987) and together give the observed mean plume displacement. The injected solute mass and initial configuration are taken from Mackay et al. (1986); however since the initial spatial configuration of the solute pulse was not well defined (Barry et al., 1988) the initial concentration was determined by spreading the total injected mass evenly over a rectangular prism which preserved the spatial moments observed one day after injection, as presented in Freyberg (1986).

**TABLE 1 Inputs for the Field Problem.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln K Covariance</td>
<td>( P_{l}(\xi) = \sigma_{l}^{2} \left( \frac{\pi^{2} \xi}{4\lambda} \right) K_{1}\left( \frac{\pi^{2} \xi}{4\lambda} \right) - \frac{1}{2} \left( \frac{\pi^{2} \xi}{4\lambda} \right)^{2} K_{0}\left( \frac{\pi^{2} \xi}{4\lambda} \right) )</td>
</tr>
<tr>
<td>Ln K Variance</td>
<td>( \sigma_{l}^{2} = 0.29 )</td>
</tr>
<tr>
<td>Ln K Integral Scale</td>
<td>( \lambda = 2.8 \text{ m} )</td>
</tr>
<tr>
<td>Mean Hydraulic Conductivity</td>
<td>( K_{G} = 6.18 \text{ m day}^{-1} )</td>
</tr>
<tr>
<td>Porosity</td>
<td>( n = 0.33 )</td>
</tr>
<tr>
<td>Longitudinal Dispersivity</td>
<td>( \alpha_{l} = 0.05 \text{ m} )</td>
</tr>
<tr>
<td>Transverse Dispersivity</td>
<td>( \alpha_{t} = 0.05 \text{ m} )</td>
</tr>
<tr>
<td>Initial Concentration</td>
<td>337 mg/l over 4 mx4 mx6 m prism</td>
</tr>
<tr>
<td>Domain Size</td>
<td>60 m by 20 m</td>
</tr>
<tr>
<td>Computational Grid</td>
<td>( \Delta x = \Delta y = 1 \text{ m} )</td>
</tr>
<tr>
<td>Covariance Grid</td>
<td>( \Delta x = \Delta y = 2 \text{ m} )</td>
</tr>
<tr>
<td>Time Step (0 to 45 days)</td>
<td>( \Delta t = 2.5 \text{ days} )</td>
</tr>
<tr>
<td>Time Step (45 to 260 days)</td>
<td>( \Delta t = 5 \text{ days} )</td>
</tr>
</tbody>
</table>

(Note that \( K_{n} \) represents the modified Bessel function of order \( n \). The Ln K covariance is used to derive the velocity covariance (Graham & McLaughlin, 1989b))

The values selected for local dispersivity are on the high end of the generally accepted range, primarily for numerical convenience. The numerical limitation of concern in our stochastic solute transport model is not numerical dispersion or oscillation of the mean concentration plume, however, since this equation explicitly contains a macrodispersive-flux term. Rather it is the grid resolution required to capture the spatial
variability of the higher order moment equations which are driven by forcing terms with pronounced spatial gradients.

Figure 1 shows the actual vertically averaged Borden chloride plumes and Fig. 2 shows the unconditional mean chloride concentration plumes at 42.5, 85 and 260 days after initial solute release. The unconditional mean plumes represent the best estimates of the trajectory of this plume, if no site-specific concentration measurements are available. The mean concentration plumes are regular and symmetric, resembling those that would be obtained from the Dagan (1984) unconditional stochastic transport model. It is important to interpret the mean concentration plumes together with the concentration standard deviations shown in Fig. 3, since the standard deviation plots give an estimate of the likely range of deviation of any particular solute plume replicate around the ensemble mean.

Figure 4 is a plot of the 42.5, 85, and 260 day normalized unconditional residuals which were calculated by subtracting the unconditional mean concentration field from the actual depth averaged plume and dividing by the local unconditional standard deviation. If the concentration residuals are normal they are expected to fall within two standard deviations of the theoretical residual mean of zero 95% of the time. Figure 4 shows that the unconditional concentration residuals fall within the two standard deviation limit at all times and locations except near the center of mass of the mean plume at 42.5 days. (Note that the actual concentration observations were interpolated onto the computational grid for convenience in calculating the concentration residuals and taking concentration observations.)

Figures 1 through 4 indicate that the unconditional mean plumes are quite successful in predicting the peak concentrations and average spreading of the actual depth-averaged tracer plumes. However, there is still significant uncertainty associated with these predictions. The magnitude of this uncertainty is estimated accurately by the unconditional concentration standard deviation.

To condition the concentration moments, eight monitoring wells were located in areas of high uncertainty at 42.5 days, five additional wells were similarly located at 85 days, and 12 additional wells were similarly located at 260 days. Existing wells were resampled at subsequent times if they fell inside the area of significant uncertainty for the propagated conditional mean plume. Thus a total of eight wells were sampled at 42.5 days, 10 wells were sampled at 85 days, and 12 wells were sampled at 260 days.

Figure 5 shows conditional mean plumes at each sampling time. The locations of the wells sampled at each time are indicated with asterisks on the respective conditional mean plume plots. Note that the conditional mean plumes are more irregular than the unconditional mean plumes and they have begun to pick up the spatial variability exhibited by the actual solute plume.

Figure 6 shows the normalized conditional concentration residuals which were calculated by subtracting the conditional mean concentration field from the actual depth averaged plume and dividing by the local conditional standard deviation. At each sampling time the normalized residuals exceed the two standard deviation limit over a significant portion of the domain. This indicates that the conditional prediction uncertainty is under estimated by the stochastic model, i.e. the model is over-confident of its conditional mean predictions. It may be possible to alleviate this problem by including measurement errors in the conditioning process, and allowing for uncertainty in the specified initial and boundary conditions. However the problem may also stem from the fact that the solute transport problem is non-gaussian. In this case the conditional moments would not necessarily be predicted well by the Kalman filtering algorithm (which is a linear estimator), and the two standard deviation limit would not apply.
FIG. 1  Actual Concentration Plumes.

FIG. 2  Unconditional Mean Concentration Plumes.
FIG. 3  Unconditional Standard Deviation Plumes

FIG. 4  Normalized Unconditional Residuals.
FIG. 5  Conditional Mean Concentration Plumes.

FIG. 6  Normalized Conditional Residuals.
DISCUSSION

Analysis of the Borden field data indicates that the spatially distributed stochastic transport model we have developed holds promise for application to real-world problems. This model provides 1) a prediction of the extent of the solute plume, 2) an estimate of the reliability of this prediction, and 3) a means of combining physically based models with field data using the Bayesian concepts of measurement conditioning. To our knowledge this is the first time a rigorous, stochastic conditioning procedure has been applied to a solute plume at an actual field site.

Because hydrogeologic variability at the Borden site is quite small, the unconditional stochastic model provides adequate predictions in this case. Nevertheless some improvement is achieved by conditioning the concentration moments using a small subset of the available concentration observations. Since no head or hydraulic conductivity measurements are available to condition the velocity field in the vicinity of the solute plume the full benefits of site-specific conditioning cannot be realized for this example. Thus the somewhat limited improvement demonstrated by the conditional moments reinforces the results of Graham & McLaughlin (1989b) which indicate that the key to the conditioning process is to define the site-specific spatial variability of the groundwater velocity field.

REFERENCES


