102328

WATER RESOURCES RESEARCH, VOL. 43, W11403, doi:10.1029/2006WR005580, 2007

Click Here for Full Article

Conditioned backward probability modeling to identify sources of groundwater contaminants subject to sorption and decay

Roseanna M. Neupauer,[1] Ranhao Lin,[2,3] and Heather O'Shea[1]

Received 2 October 2006; revised 11 July 2007; accepted 31 July 2007; published 2 November 2007.

[1] If contamination is observed in an aquifer, a backward probability model can be used to obtain information about the former position of the observed contamination. A backward location probability density function (PDF) describes the possible former positions of the observed contaminant particle at a specified time in the past. If the source release time is known or can be estimated, the backward location PDF can be used to identify possible source locations. For sorbing solutes, the location PDF depends on the phase (aqueous or sorbed) of the observed contamination and on the phase of the contamination at the source. These PDFs are related to adjoint states of aqueous and sorbed phase concentrations. The adjoint states, however, do not take into account the measured concentrations. Neupauer and Lin (2006) presented an approach for conditioning backward location PDFs on measured concentrations of non-reactive solutes. In this paper, we present a related conditioning method to identify the location of an instantaneous point source of a solute that exhibits first-order decay and linear equilibrium or non-equilibrium sorption. We derive the conditioning equations and present an illustrative example to demonstrate important features of the technique. Finally, we illustrate the use of the conditioned location PDF to identify possible sources of contamination by using data from a trichloroethylene plume at the Massachusetts Military Reservation.

Citation: Neupauer, R. M., R. Lin, and H. O'Shea (2007), Conditioned backward probability modeling to identify sources of groundwater contaminants subject to sorption and decay, *Water Resour. Res.*, *43*, W11403, doi:10.1029/2006WR005580.

1. Introduction

[2] The movement of groundwater contaminants is uncertain because of both the spatial variability of flow and transport parameters and the scarcity of data. As a result, predicting the movement of a contaminant plume is imprecise. Likewise, identifying an unknown source of contamination is affected by these same limitations, and is further complicated by the sparse and imperfect measurements of the existing contaminant plume. This paper presents a method for identifying sources of reactive contaminants using probability density functions that are conditioned on the measured concentrations. We consider contaminants that undergo first-order decay and linear equilibrium or non-equilibrium sorption.

[3] In modeling the movement of a contaminant plume, the position of the plume is often treated as a random variable that accounts for the spatial variability and uncertainty of aquifer parameters. *Dagan* [1982] introduced the concept of a location probability density function (PDF) as a means of predicting the random concentration distribution, and used spatial moments of the PDF to determine the expected value and variance of the predicted concentration. The spread of the plume is a direct result of the aquifer heterogeneity, and therefore the second spatial moment (variance) of the plume depends on the covariance of the hydraulic conductivity. Moments of location PDFs have been used extensively in this context to predict solute concentrations in heterogeneous porous media [e.g., *Dagan*, 1982, 1984, 1987; *Kitanidis*, 1988] and to quantify macro-dispersivity [e.g., *Dagan*, 1990]. A similar approach has been followed for reactive solutes. *Dagan and Cvetkovic* [1993] and *Cvetkovic and Dagan* [1994] developed spatial and temporal moments of the aqueous-phase concentration distributions of a solute that undergoes linear sorption. The moments are derived from the moments of the concentration distribution of a non-reactive solute and a distribution function that characterizes the sorption kinetics. This approach has been extended to non-linear sorption [*Dagan and Cvetkovic*, 1996; *Cvetkovic and Dagan*, 1996] and to aquifers with spatially variable sorption properties [*Cvetkovic et al.*, 1998].

[4] Similar work has been done to analyze solute transport through the use of a traveltime probability density function that is related to the solute mass flux [e.g., *Selroos and Cvetkovic*, 1994], and through the use of temporal moments of mass flux of non-reactive [*Shapiro and Cvetkovic*, 1988]

[1]Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, Colorado, USA.
[2]Department of Civil Engineering, University of Virginia, Charlottesville, Virginia, USA.
[3]Now at Bank of Montreal, Toronto, Ontario, Canada.

Copyright 2007 by the American Geophysical Union.
0043-1397/07/2006WR005580$09.00

and reactive solutes [*Cvetkovic and Shapiro*, 1990; *Selroos and Cvetkovic*, 1994; *Andricevic and Cvetkovic*, 1996, 1998; *Rubin et al.*, 1997; *Lawrence et al.*, 2002; *Sanchez-Vila and Rubin*, 2003].

[5] The spatial and temporal moments of concentration distributions are useful in predicting the movement of a contaminant plume away from a known source. If contamination is observed in an aquifer, but the source is unknown, a backward location probability density function can be used to obtain information about the source location. A backward location probability density function (PDF) describes the possible former positions of the observed contaminant particle, and can be used to identify sources of groundwater contamination. *Neupauer and Wilson* [1999, 2001] have shown that backward PDFs are related to adjoint states of concentration and therefore can be obtained by solving the adjoint of a forward contaminant transport model. In the adjoint equation, the flow field is reversed, simulating upgradient movement in reversed time. The observation location is treated as a source of an adjoint state (state variable that is related to the PDF), and the movement is tracked upgradient and backward in time, leading to probability density functions of possible former locations or release times of the observed solute particle. The variance of the PDF is controlled by the dispersion coefficient, which accounts for the unmodeled velocity variations due to heterogeneity of hydraulic conductivity. Spatial variability of heterogeneity exists at multiple scales [e.g., *Dagan*, 1986]. The larger-scale heterogeneity can be explicitly modeled in a groundwater flow simulation. Heterogeneities at scales smaller than a grid block in the numerical model cannot be modeled explicitly, and therefore are accounted for through the dispersion coefficient.

[6] For sorbing solutes, both the former position and the former phase of the contaminant particle are unknown; therefore two different backward location PDFs can be obtained. The aqueous phase backward location PDF describes the possible positions of the observed particle if it had been in the aqueous phase at the former time of interest; and the sorbed phase backward location PDF describes its possible positions if it had been in the sorbed phase. For kinetically sorbing solutes, the possible former positions of the observed particle also depend on the phase of the particle at the observation location. As an example, consider a sorbing solute, with partitioning between the aqueous and sorbed phases defined by

$$\theta \frac{\partial C}{\partial t} = \alpha_s \rho_b (C_S - K_d C) \tag{1}$$

$$\rho_b \frac{\partial C_S}{\partial t} = \alpha_s \rho_b (K_d C - C_S), \tag{2}$$

where C and C_S are the aqueous and sorbed phase concentrations, θ is porosity, ρ_b is bulk density, K_d is the partition coefficient, and α_s is the rate constant. *Valocchi and Quinodoz* [1989] showed that for this first-order, linear, reversible reaction, the random amounts of time that particles spend in the aqueous and the sorbed phases are defined by exponential probability density functions. The PDF for the time a particle spends in the aqueous phase is given by

$$f_T(t) = \frac{\rho_b K_d \alpha_s}{\theta} e^{-\frac{\rho_b K_d \alpha_s t}{\theta}}, \tag{3}$$

where $\rho_b K_d \alpha_s / \theta$ is the exponential parameter; while the PDF for the time a particle spends in the sorbed phase has a parameter of α_s, and is given by

$$f_T(t) = \alpha_s e^{-\alpha_s t}. \tag{4}$$

From (3) and (4), the expected values of the amount of time a particle spends in the aqueous and sorbed phases are $\theta/(\rho_b K_d \alpha_s)$ and $1/\alpha_s$, respectively. For a slowly sorbing solute, α_s is small; therefore a particle that is presently in the sorbed phase is likely to remain in the sorbed phase for a relatively long time. Likewise, if a particle is observed in the sorbed phase, it is likely to have been in the sorbed phase in the recent past, and therefore is not likely to have traveled from distant points in the aquifer. Similarly, if a particle is observed in the aqueous phase, it is likely to have been in the aqueous phase in the recent past, and therefore could have traveled from more distant locations. For sorbing solutes, four different backward location PDFs can be obtained depending on the former phase (aqueous or sorbed) of the contaminant particle and the observed phase (aqueous or sorbed) [*Neupauer and Wilson*, 2004a].

[7] In the adjoint model, the observation location is treated as an instantaneous point source of the adjoint state, representing a probability of unity that the contaminant particle was at the observation location at the time of sampling. The adjoint model does not, however, make use of the measured concentration, a quantity that can provide additional information about the source of contamination. If contamination is measured at two locations, the adjoint model weights each observation equally when calculating the backward location PDF [*Neupauer and Wilson*, 2005]; however the observation with the higher concentration is likely to be more informative and therefore should receive more weight. Recently, *Neupauer and Lin* [2006] presented a method for conditioning backward probability density functions on measured concentrations for non-reactive solutes. They showed that conditioning improved the accuracy and decreased the variance of the PDFs unless measurement errors were large. In this paper, we extend the conditioned backward probability model to handle reactive solutes. We consider solutes that undergo first-order decay and linear equilibrium or non-equilibrium sorption.

[8] Source locations can also be identified with a forward modeling approach. For each suspected source location, a forward simulation can be run with a simulated source of contamination at the suspected source location. The simulated concentration distributions can be compared to the measured concentrations to determine which of the suspected source locations are feasible. This approach requires one simulation for each suspected source location, so if the number of possible source locations is large, this approach is computationally inefficient. The adjoint model, on the other hand, obtains information about all possible source locations with one simulation, and therefore is more effi-

cient if the number of observations is small relative to the number of possible sources.

[9] Another method that can be used to identify source locations is backward particle tracking. In this method, particles are placed at the current location of the contaminant plume, and the particles are tracked backward in time following the advective flow paths. The final locations of the particles identify possible source locations. An advantage of the adjoint method over backward particle tracking is that the adjoint method accounts for unmodeled spatial variability of velocities through the dispersion coefficient. Because of these uncertainties, the adjoint method produces a PDF of the possible source locations, which can be used to differentiate among possible source locations. A related approach is the backward random walk, which incorporates a random particle movement into backward particle tracking to account for dispersion and uncertainty in particle velocities [*Uffink*, 1989; *Fogg et al.*, 1999]. This is equivalent to the adjoint method.

[10] In the next section, we present the adjoint equations for obtaining the unconditioned backward location PDFs for solutes that are subject to first-order decay and linear sorption. This combines the results of *Neupauer and Wilson* [2003, 2004a] for decaying and sorbing solutes, respectively. In subsequent sections, we present the theory for conditioning the backward location PDFs on the measured concentrations. This conditioning follows the approach of *Neupauer and Lin* [2006], with modifications to account for sorption and first-order decay. Finally, we present an example of the approach to identify possible sources of a trichloroethylene (TCE) plume at the Massachusetts Military Reservation. This application was first presented by *Neupauer and Wilson* [2005] using unconditioned PDFs. We use this application here to demonstrate the benefit of conditioning on measured concentrations.

2. Development of Conditioned Location Probability Density Functions

[11] If one or more observations of contamination are made, a backward probability model can be used to identify the location and phase (aqueous or sorbed) of the contamination source. We assume that N_A observations of contamination were made in the aqueous phase, and N_S observations were made in the sorbed phase (e.g., soil samples were analyzed and found to be contaminated). The observed aqueous concentrations are $\hat{\mathbf{c}} = [\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_{NA}]$ where $\hat{c}_i = \hat{c}(\mathbf{x}_{wi}, \tau_{wi})$ is the observed aqueous concentration at location $\mathbf{x}_{wi}$ at backward time τ_{wi}. Backward time is defined as the amount of time prior to the most recent sampling event. Similarly, the observed sorbed phase concentrations are $\hat{\mathbf{c}}_\mathbf{S} = [\hat{c}_{S_{N_A+1}}, \hat{c}_{S_{N_A+2}}, \ldots \hat{c}_{S_{N_A+N_S}}]$ where $\hat{c}_{Sj} = \hat{c}_S(\mathbf{x}_{wj}, \tau_{wj})$ is the observed sorbed phase concentration at location $\mathbf{x}_{wj}$ at backward time τ_{wj}, $j = N_A + 1, N_A + 2, \ldots, N_A + N_S$. We define $\{\mathbf{x}_w\} = [\mathbf{x}_{w1}, \mathbf{x}_{w2}, \ldots, \mathbf{x}_{N_A + N_s}]$ and we define $\{\tau_w\} = [\tau_{w1}, \tau_{w2}, \ldots, \tau_{N_A + N_s}]$ as the vectors of sample locations and sample times, respectively.

[12] In this section, we develop equations for conditioning backward location probability density functions on these measured aqueous and sorbed phase concentrations. We begin with background information on the unconditioned location PDFs for sorbing and decaying solutes, followed by the PDFs for measured concentrations. Finally, we use these two quantities, along with Bayes' theorem, to obtain conditioned backward location PDFs for decaying and sorbing solutes. This derivation follows a similar approach by *Neupauer and Lin* [2006] for a non-reactive solute. The extension to reactive solutes is complicated by having two possible observation phases (aqueous and sorbed), two possible source phases, and the possibility of decay.

2.1. Unconditioned Backward Location Probability Density Functions

[13] Unconditioned backward PDFs are related to adjoint states of concentration and can be obtained by solving the adjoint of a forward contaminant transport equation. *Neupauer and Wilson* [2003, 2004a] developed a model to produce unconditioned backward location probability density functions (PDFs) for reactive solutes that exhibit either first-order decay or linear equilibrium or non-equilibrium sorption. We present their main points in this section.

[14] As an example of a forward transport equation, we use the advection-dispersion-reaction equation, with linear non-equilibrium sorption and first-order decay, given by

$$\begin{aligned}
\theta\frac{\partial C}{\partial t} + \rho_b\frac{\partial C_S}{\partial t} &= \frac{\partial}{\partial x_i}\left(\theta D_{ij}\frac{\partial C}{\partial x_j}\right) - \frac{\partial}{\partial x_i}(\theta v_i C) - \theta\lambda_A C \\
&\quad - \rho_b\lambda_S C_S + q_I C_I - q_O C \\
\rho_b\frac{\partial C_S}{\partial t} &= \alpha_s\rho_b(K_d C - C_S) - \rho_b\lambda_S C_S \\
C(\mathbf{x}, t_o) &= C_{Ai}(\mathbf{x}) \\
C_S(\mathbf{x}, t_o) &= C_{Si}(\mathbf{x}) \\
C(\mathbf{x}, t) &= g_1(t) \text{ on } \Gamma_1 \\
\left[D_{ij}\frac{\partial C}{\partial x_j}\right]\cdot \mathrm{n}_i &= g_2(t) \text{ on } \Gamma_2 \\
\left[v_i C - D_{ij}\frac{\partial C}{\partial x_j}\right]\cdot \mathrm{n}_i &= g_3(t) \text{ on } \Gamma_3,
\end{aligned} \tag{5}$$

where $C(\mathbf{x}, t)$ (in units of mass of solute per volume of water) and $C_S(\mathbf{x}, t)$ (in units of mass of solute per mass of solid) are the aqueous and sorbed phase concentrations, respectively, t is time, $\mathbf{x}$ is the position vector ($\mathbf{x} = [x_1, x_2, x_3]$), θ is porosity, ρ_b is bulk density, D_{ij} is the i, jth entry of the dispersion tensor ($i, j = 1, 2, 3$), v_i is the groundwater velocity in the x_i direction, λ_A and λ_S are first-order decay rates in the aqueous and sorbed phases, respectively, q_I is fluid inflow rate per unit volume, C_I is inflow concentration, q_O is fluid outflow rate per unit volume, α_s is a rate constant, K_d is the partition coefficient, C_{Ai} and C_{Si} are the initial aqueous and sorbed phase concentrations, respectively, $g_1(t)$, $g_2(t)$, and $g_3(t)$ are known boundary functions, Γ_1, Γ_2, and Γ_3 are the domain boundaries, and n_i is the outward unit normal vector in the x_i direction. We assume that aquifer properties can vary in space but are constant in time. Note that the boundary conditions shown in (5) are consistent with the boundary conditions used in many transport codes, including MT3DMS [*Zheng and Wang*, 1999]. With these boundary conditions, an outflow boundary would be represented by Γ_2, with $g_2(t) = 0$. This homogeneous Neuman boundary condition allows advective mass flux out of the domain, but does not allow dispersive mass flux. Other codes, such as WATFLOW [*Molson et al.*, 2002], use an implicit Neuman boundary conditions [*Frind*, 1988; *Cornaton et al.*, 2004]

that allows for both advective and dispersive mass flux at an outflow boundary. *Cornaton et al.* [2004] have shown that the implicit Neuman boundary condition is more representative of the physical conditions at an outflow boundary. In this work, we consider only the three boundary types shown in (5).

[15] The form of the adjoint of (5) that is used in the backward probability model is [*Neupauer and Wilson*, 2003, 2004a]

$$\begin{aligned}
\theta\frac{\partial\phi_A}{\partial\tau}+\rho_b\frac{\partial\phi_S}{\partial\tau} &= \frac{\partial}{\partial x_i}\left(\theta D_{ij}\frac{\partial\phi_A}{\partial x_j}\right)+\frac{\partial}{\partial x_i}(\theta v_i\phi_A)-\theta\lambda_A\phi_A \\
&\quad -\rho_b\lambda_S\phi_S-q_I\phi_A+\delta(\mathbf{x}-\mathbf{x}_w)\delta(\tau) \\
\rho_b\frac{\partial\phi_S}{\partial\tau} &= \alpha_s\rho_b(K_d\phi_A-\phi_S)-\rho_b\lambda_S\phi_S+Q_S(\mathbf{x},\tau) \\
&\phi_A(\mathbf{x},\tau=\tau_w)=0 \\
&\phi_S(\mathbf{x},\tau=\tau_w)=0 \\
&\phi_A=0 \text{ on } \Gamma_1 \\
&\left[D_{ij}\frac{\partial\phi_A}{\partial x_j}+v_i\phi_A\right]\cdot \mathrm{n}_i=0 \text{ on } \Gamma_2 \\
&\left[D_{ij}\frac{\partial\phi_A}{\partial x_j}\right]\cdot \mathrm{n}_i=0 \text{ on } \Gamma_3,
\end{aligned} \tag{6}$$

where $\phi_A(\mathbf{x}, \tau)$ and $\phi_S(\mathbf{x}, \tau)$ are adjoint states of C and C_S, respectively, τ is backward time (time before sampling), τ_w is the observation time (in backward time), and $Q_S(\mathbf{x}, \tau)$ is a load term that depends on the phase (aqueous or sorbed) of the observation: $Q_S(\mathbf{x}, \tau) = 0$ for an aqueous phase observation, and $Q_S(\mathbf{x}, \tau) = \delta(\mathbf{x} - \mathbf{x}_w)\delta(\tau)$ for a sorbed phase observation. The adjoint equation has the same form as (5), except that the flow field is reversed, the time derivative is in terms of backward time τ, and the boundary conditions are slightly changed.

[16] The adjoint states are related to backward location PDFs through

$$f_{\mathbf{X},P_o}(\mathbf{x},p_o=A;p_w=l,\tau,\mathbf{x}_w,\tau_w) \;=\; \theta(\mathbf{x})\phi_A^l(\mathbf{x},\tau) \tag{7}$$

$$f_{\mathbf{X},P_o}(\mathbf{x},p_o=S;p_w=l,\tau,\mathbf{x}_w,\tau_w) \;=\; \rho_b(\mathbf{x})\phi_S^l(\mathbf{x},\tau), \tag{8}$$

where $f_{\mathbf{X},P_o}(\mathbf{x}, p_o = k; p_w = l, \tau, \mathbf{x}_w, \tau_w)$ describes the random former position $\mathbf{X}$ and random former phase P_o (A = aqueous, S = sorbed) at backward time τ of a contaminant particle that was observed in phase $p_w = l$ ($l = A, S$) at the observation location $\mathbf{x}_w$ and at backward time $\tau = \tau_w$, and the superscript l on the adjoint state denotes the observation phase [*Neupauer and Wilson*, 2004a]. In the probability notation, variables to the left of the semi-colon are random variables, and variables to the right of the semi-colon are deterministic parameters. Also, capital letters denote random variables and lower case letters denote a particular value. Finally, the subscript o denotes the source or former position, and the subscript w denotes the observation well or the current location. Note that the backward location PDFs are mixed continuous and discrete PDFs because the source location is a continuous random variable, while the source phase is a discrete random variable with two possible values: $p_o = A$ (aqueous) or $p_o = S$ (sorbed).

[17] For multiple observations, a multiple-observation location PDF describes the possible former position of all observed particles at time τ assuming that they all were at the same location and in the same phase at time τ. Suppose N_A observations of the contaminant were made in the aqueous phase, and N_S observations of the contaminant were made in the sorbed phase. The multiple-observation unconditioned backward location PDF is given by [*Neupauer and Wilson*, 2005]

$$f_{\mathbf{X}_o,P_o}(\mathbf{x}_o,p_o;\{\mathbf{x}_w\},\{\tau_w\})=\frac{\prod_{i=1}^{N_A}f_{\mathbf{X},P_o}(\mathbf{x},p_o;p_w=A,\tau,\mathbf{x}_{wi},\tau_{wi})\prod_{j=N_A+1}^{N_A+N_S}f_{\mathbf{X},P_o}\left(\mathbf{x},p_o;p_w=S,\tau,\mathbf{x}_{wN_A+j},\tau_{wN_A+j}\right)}{\int\prod_{i=1}^{N_A}f_{\mathbf{X},P_o}(\mathbf{x},p_o;p_w=A,\tau,\mathbf{x}_{wi},\tau_{wi})\prod_{j=N_A+1}^{N_A+N_S}f_{\mathbf{X},P_o}\left(\mathbf{x},p_o;p_w=S,\tau,\mathbf{x}_{wN_A+j},\tau_{wN_A+j}\right)d\mathbf{x}}, \tag{9}$$

where $f_{\mathbf{X},P_o}(\mathbf{x}, p_o = k; p_w = l, \tau, \mathbf{x}_w, \tau_w)$ are the unconditioned location PDFs shown in (7) and (8). At this time, the multiple-observation unconditioned backward location PDF has only been developed for a single instantaneous point source of contamination. For preliminary results on the extension to multiple instantaneous point sources of contamination, see *Lin* [2003].

[18] Notice that the equation for the adjoint states (6) has decay terms, indicating that the adjoint states (and therefore the PDFs) decay over time. This implies that more distant sources and earlier release times (later backward times) are less likely than sources near the observation location or with later release times (earlier backward times) because a contaminant particle that was released farther upgradient or at an earlier time would be more likely to have decayed before reaching the observation location [*Neupauer and Wilson*, 2003]. Since a contaminant particle that is observed obviously did not decay, a modeler can choose to disregard this possibility of decay by conditioning the PDF on the fact that the observed contaminant particle did not yet decay by the time it was observed. This is accomplished by normalizing the PDF by the total probability at backward time τ [*Neupauer and Wilson*, 2002]. For the location PDF, when the first-order decay rate is uniform, this normalization is identical to eliminating the decay term from the adjoint equation.

2.2. PDFs of Measured Concentrations

[19] The conditioning procedure described in the next section relies on the comparison of measured and simulated

contaminant concentrations. The conditioned backward modeling procedure identifies the most likely source locations to be the locations from which a release of contamination produces simulated concentrations at all observations that are similar to the measured concentrations. The simulated concentrations will not be identical to the measured concentrations. The model is a simplified version of reality, and therefore it does not capture all of the heterogeneities that exist in the true aquifer. In the model, some of the heterogeneity is accounted for in the dispersion coefficient; however, *Fitts* [1996] has shown that this leads to simulated concentration distributions that are smoother than actual distributions. Also, in a finite difference model, the simulated concentrations are averages over the grid block, while the measured concentrations are essentially taken at points in the horizontal plane; thus, the simulated concentrations do not represent the same sample volume as the measurements.

[20] To account for the differences between measured and simulated concentrations, we define a PDF on the measured concentration that is based on the simulated concentration. We assume that the measured concentrations, $\hat{\mathbf{c}}$ and $\hat{\mathbf{c}}_{\mathbf{S}}$, contain random measurement error [*Neupauer and Lin*, 2006], such that

$$\hat{c}_i = \hat{c}(\mathbf{x}_{wi}, \tau_{wi}) = \langle C(\mathbf{x}_{wi}, \tau_{wi}|m_o, \mathbf{x}_o, p_o, \tau)\rangle + \epsilon_i \quad i = 1, 2, \ldots, N_A \tag{10}$$

$$\hat{c}_{Sj} = \hat{c}_S(\mathbf{x}_{wj}, \tau_{wj}) = \langle C_S(\mathbf{x}_{wj}, \tau_{wj}|m_o, \mathbf{x}_o, p_o, \tau)\rangle + \epsilon_j \quad j = N_A + 1, N_A + 2, \ldots, N_A + N_S, \tag{11}$$

where $\langle C(\mathbf{x}_{wi}, \tau_{wi}|m_o, \mathbf{x}_o, p_o; \tau)\rangle$ is the expected value of the aqueous concentration at location $\mathbf{x}_{wi}$ at backward time τ_{wi}, given a source of mass m_o at location $\mathbf{x}_o$ at time τ in phase p_o ($p_o = A$ for an aqueous phase source, and $p_o = S$ for a sorbed phase source), $\langle C_S(\mathbf{x}_{wj}, \tau_{wj}|m_o, \mathbf{x}_o, p_o; \tau)\rangle$ is the expected value of the sorbed phase concentration at location $\mathbf{x}_{wj}$ at backward time τ_{wj}, given a source of mass m_o at location $\mathbf{x}_o$ at time τ in phase p_o, and ϵ_i and ϵ_j are the errors. In the notation in (10) and (11), the PDF is conditioned on variables to the right of the vertical bar.

[21] We assume that the errors, ϵ_i and ϵ_j, are normally distributed with zero mean and variances of σ_i^2 and σ_j^2, respectively. If measurement errors do not follow a normal distribution, other error models can be used. *Lin* [2003] found that the variance of the measurement error has a substantially greater effect on the conditioned backward PDF than does the choice of distributions of measurement error. With the assumption of normally distributed random measurement error, the measured aqueous and sorbed phase concentrations, $\hat{c}_i$ and $\hat{c}_{S_j}$, are also normally distributed random variables whose probability density functions are given by

$$f_{\hat{C}_i|M_o,\mathbf{X}_o,P_o}(\hat{c}_i|m_o, \mathbf{x}_o, p_o, \tau) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left\{-\frac{[\hat{c}_i - \langle C(\mathbf{x}_{wi}, \tau_{wi}|m_o, \mathbf{x}_o, p_o, \tau)\rangle]^2}{2\sigma_i^2}\right\} \tag{12}$$

$$f_{\hat{C}_{Sj}|M_o,\mathbf{X}_o,P_o}(\hat{c}_{Sj}|m_o, \mathbf{x}_o, p_o, \tau) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \cdot \exp\left\{-\frac{[\hat{c}_{Sj} - \langle C_S(\mathbf{x}_{wj}, \tau_{wj}|m_o, \mathbf{x}_o, p_o, \tau)\rangle]^2}{2\sigma_j^2}\right\}. \tag{13}$$

[22] Assuming that the measurement errors are independent, the joint PDF of all measured concentrations is the product of the PDFs for each measured concentration, given by

$$f_{\hat{\mathbf{C}},\hat{\mathbf{C}}_S|M_o,\mathbf{X}_o,P_o}(\hat{\mathbf{c}}, \hat{\mathbf{c}}_S|m_o, \mathbf{x}_o, p_o; \tau) = \prod_{i=1}^{N_A} f_{\hat{C}_i|M_o,\mathbf{X}_o,P_o}(\hat{c}_i|m_o, \mathbf{x}_o, p_o, \tau) \cdot \prod_{j=N_A+1}^{N_A+N_S} f_{\hat{C}_{Sj}|M_o,\mathbf{X}_o,P_o}(\hat{c}_{Sj}|m_o, \mathbf{x}_o, p_o, \tau) \tag{14}$$

where the two PDFs on the right-hand side are the normal distributions defined in (12) and (13).

2.3. Conditioned Backward Location PDFs

[23] In this section, we develop the backward location PDF for a reactive solute (first-order decay and linear sorption) that is conditioned on measured concentrations. We assume that the source of contamination was an instantaneous point source in either the aqueous phase or the sorbed phase, and that the release time was known. The backward PDF identifies the possible former positions and the former phase of an observed contaminant particle at the time of release. Since the source mass M_o, location $\mathbf{X}_o$, and phase P_o are all unknown (random) variables, we first develop a joint PDF on these three random variables that is conditioned on the measured concentrations, and then we find the marginal distribution over all possible values of source mass. This is similar to the approach followed by *Neupauer and Lin* [2006]; however the source phase for their conservative solute was known to be the aqueous phase, so the joint distribution only contained the two random variables $\mathbf{X}_o$ and M_o.

[24] Let us define $f_{M,\mathbf{X}_o,P_o|\hat{\mathbf{C}},\hat{\mathbf{C}}_S}(m_o, \mathbf{x}_o, p_o|\hat{\mathbf{c}},\hat{\mathbf{c}}_S; \tau, \{\mathbf{x}_w\}, \{\tau_w\})$ as the joint PDF of source mass, source location, and source phase conditioned on the measured concentrations. Through Bayes' theorem, a joint PDF for the unknowns $(m_o, \mathbf{x}_o, p_o)$, conditioned on the known data, can be written as

$$f_{M,\mathbf{X}_o,P_o|\hat{\mathbf{C}},\hat{\mathbf{C}}_S}(m_o, \mathbf{x}_o, p_o|\hat{\mathbf{c}}, \hat{\mathbf{c}}_S; \tau, \{\mathbf{x}_w\}, \{\tau_w\}) = \frac{f_{\hat{\mathbf{C}},\hat{\mathbf{C}}_S|M,\mathbf{X}_o,P_o}(\hat{\mathbf{c}}, \hat{\mathbf{c}}_S|m_o, \mathbf{x}_o, p_o; \tau) f_{M,\mathbf{X}_o,P_o}(m_o, \mathbf{x}_o, p_o; \{\mathbf{x}_w\}, \{\tau_w\})}{\sum_{p_o}\left\{\int\int\left[f_{\hat{\mathbf{C}},\hat{\mathbf{C}}_S|M,\mathbf{X}_o,P_o}(\hat{\mathbf{c}}, \hat{\mathbf{c}}_S|m_o, \mathbf{x}_o, p_o; \tau) f_{M,\mathbf{X}_o,P_o}(m_o, \mathbf{x}_o, p_o; \{\mathbf{x}_w\}, \{\tau_w\})\right] dm_o d\mathbf{x}_o\right\}}. \tag{15}$$

The denominator is identical to the numerator, except that it is integrated over the source mass and spatial domains, and summed over the two possible (discrete) former phases, $p_o = A, S$, where A = aqueous and S = sorbed.

[25] The first probability density function in both the numerator and the denominator in (15) is given by (14). The second probability density function in the numerator and denominator represents the joint PDF of source mass, location, and phase in the absence of any concentration information. Since source mass is independent of source location and phase, this joint PDF can be separated into the product of two PDFs, given by

$$f_{M,\mathbf{X}_o,P_o}(m_o,\mathbf{x}_o,p_o;\{\mathbf{x}_w\},\{\tau_w\}) = f_M(m_o;\{\mathbf{x}_w\},\{\tau_w\}) \cdot f_{\mathbf{X}_o,P_o}(\mathbf{x}_o,p_o;\{\mathbf{x}_w\},\{\tau_w\}). \quad (16)$$

[26] The first PDF on the right-hand side is a PDF of source mass alone. In the absence of any other information, we assume a uniform distribution for source mass. The second distribution is the unconditioned location PDF for multiple observations given by (9).

[27] Substituting (9), (14), and (16) into (15), and integrating over the source mass domain, we obtain

$$f_{\mathbf{X}_o,P_o|\hat{\mathbf{C}},\hat{\mathbf{C}}_S}(\mathbf{x}_o,p_o|\hat{\mathbf{c}},\hat{\mathbf{c}}_S) = \beta_{\mathbf{x}} \int \left[\prod_{i=1}^{N_A} f_{\hat{C}_i|M,\mathbf{X}_o,P_o}(\hat{c}_i|m_o,\mathbf{x}_o,p_o) \cdot f_{\mathbf{X},P_o}(\mathbf{x},p_o;p_w=A,\tau,\mathbf{x}_{wi},\tau_{wi}) \cdot \prod_{j=N_A+1}^{N_A+N_S} f_{\hat{C}_{Sj}|M,\mathbf{X}_o,P_o}(\hat{c}_{Sj}|m_o,\mathbf{x}_o,p_o) \cdot f_{\mathbf{X},P_o}(\mathbf{x},p_o;p_w=S,\tau,\mathbf{x}_{wj},\tau_{wj}) \right] dm_o, \quad (17)$$

where $\beta_{\mathbf{x}}$ accounts for the uniform distribution of source mass and for the denominators in (15) and (9) by ensuring that the total probability is unity. The distributions on $\hat{C}_i$ and $\hat{C}_{Sj}$ given source mass, location, and phase are the normal distributions defined in (12) and (13), respectively, with mean values equal to the expected values of concentrations. The expected values are related to the unconditioned location probability density functions [*Dagan*, 1982], and therefore are related to the adjoint states (backward model results) through

$$\langle C(\mathbf{x}_w,t=T,P_o=l)\rangle = m_o\phi_l^A(\mathbf{x}_o,\tau=T;\mathbf{x}_{wi},\tau_{wi},P_w=A) \quad (18)$$

$$\langle C_S(\mathbf{x}_w,t=T,P_o=l)\rangle = m_o\phi_l^S(\mathbf{x}_o,\tau=T;\mathbf{x}_{wi},\tau_{wi},P_w=S), \quad (19)$$

Because the adjoint states are used as shown here to estimate the true concentration, it is important that appropriate adjoint states are chosen. As mentioned above, if first-order decay is modeled, the adjoint state (and PDF) also decays, indicating that the observed contaminant particle could have decayed if it had come from more distant sources. The modeler can instead choose to condition the PDF on the fact that the observed contaminant particle did not decay. The appropriate adjoint states to use in (18) and (19) are those obtained from (6) with the decay.

[28] In summary, the conditioned backward location PDF in (17) is obtained by first solving the adjoint equation once for each observation. The resulting adjoint states are used in the expressions in (7) and (8) to obtain the unconditioned PDFs, and in (18) and (19) to obtain the true concentrations. The true concentrations from (18) and (19) are used in (12) and (13) to obtain the PDFs of the measured concentrations. The unconditioned PDFs and the PDFs of the measured concentrations are substituted into (17), to obtain

$$f_{\mathbf{X}_o,P_o|\hat{\mathbf{C}},\hat{\mathbf{C}}_S}(\mathbf{x}_o,p_o=l|\hat{\mathbf{c}},\hat{\mathbf{c}}_S) = \beta_{\mathbf{x}} \int \left[\prod_{i=1}^{N_A} \frac{\theta(\mathbf{x})\phi_l^A(\mathbf{x}_o,\tau;\mathbf{x}_{wi},\tau_{wi})}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left\{ \frac{-\left[\hat{c}_i - m_o\phi_l^A(\mathbf{x}_o,\tau;\mathbf{x}_{wi},\tau_{wi})\right]^2}{2\sigma_i^2} \right\} \cdot \prod_{j=N_A+1}^{N_A+N_S} \frac{\theta(\mathbf{x})\phi_l^S(\mathbf{x}_o,\tau;\mathbf{x}_{wj},\tau_{wj})}{\sqrt{2\pi\sigma_j^2}} \cdot \exp\left\{ \frac{-\left[\hat{c}_{Sj} - m_o\phi_l^S(\mathbf{x}_o,\tau;\mathbf{x}_{wj},\tau_{wj})\right]^2}{2\sigma_j^2} \right\} \right] dm_o, \quad (20)$$

where $l = A$ for the aqueous phase backward location PDF, $l = S$ for the sorbed phase backward location PDF, and $\beta_{\mathbf{x}}$ is given by

$$\beta_{\mathbf{x}} = \left[\int_{\mathbf{x}} \left[f_{\mathbf{X}_o,P_o|\hat{\mathbf{C}},\hat{\mathbf{C}}_S}(\mathbf{x}_o,p_o=A|\hat{\mathbf{c}},\hat{\mathbf{c}}_S) + f_{\mathbf{X}_o,P_o|\hat{\mathbf{C}},\hat{\mathbf{C}}_S}(\mathbf{x}_o,p_o=S|\hat{\mathbf{c}},\hat{\mathbf{c}}_S) \right] d\mathbf{x} \right]^{-1}. \quad (21)$$

3. Illustrative Example

[29] Consider a one-dimensional, semi-infinite domain ($x_1 \geq 0$), with a pumping well at $x_1 = 0$ and an instantaneous point source of contamination at x_{1o} = 200 m, with flow toward the pumping well (i.e., in the $-x_1$ direction). For the parameter values shown in Table 1, the resulting aqueous and sorbed phase concentration distributions at t = 100 days after release from an instantaneous point source of contamination in the aqueous phase ($C_{Ai} = (m_o/\theta)\,\delta(x_1 - x_{1o})$, $C_{Si} = 0$) are shown in Figure 1a. These plumes are obtained by solving (5) in Laplace space and numerically inverting the Laplace transforms using the Talbot algorithm [*Jury and Roth*, 1990]. The aqueous plume travels toward the pumping well, with the peak aqueous phase concentration occurring near $x_1 \approx 148$ m, and the peak sorbed phase concentration occurring near $x_1 \approx 178$ m. After t = 100 days, 30% of the original mass has decayed, while 70% ($\approx e^{-\lambda t}$, $\lambda = \lambda_A = \lambda_S$) remains in the aquifer.

[30] To illustrate the conditioned backward location probability density function, we use two observations – an aqueous phase observation of $\hat{C} = 5.5 \times 10^{-3}$ g/m^3 at x_{w1} = 100 m, and a sorbed phase observation of $\hat{C}_S = 2.7 \times 10^{-6}$ g/kg ($\rho_b\hat{C}_S/\theta$ = 0.013 g/m^3) at x_{w2} = 150 m. Both samples are taken at forward time t = 100 d, which is equivalent to backward time $\tau = \tau_1 = \tau_2 = 0$. The

Table 1. Parameters Used in the Examples

Parameter	Value
Source location, x_{1o}	200 m
Groundwater velocity, v_1	−1 m/d
Dispersion coefficient, D_{11}	3 m^2/d
Porosity, θ	0.3
Partition coefficient, K_d	0.3 cm^3/g
Bulk density, ρ_b	1.5 g/cm^3
Retardation coefficient, R	2.5
Rate constant, α_s	0.02 d^{-1}
Source mass, m_o	1 g/cm^2
Inflow rate, q_I	0
Outflow rate, q_O	0
Decay rate in aqueous phase, λ_A	0.0035 d^{-1}
Decay rate in sorbed phase, λ_S	0.0035 d^{-1}
Aqueous concentration at x_1 = 100 m	5.5×10^{-3} g/m^3
Sorbed phase concentration at x_1 = 150 m	2.7×10^{-6} g/kg^3

unconditioned backward location PDFs at τ = 100 days before observation are shown in Figures 1b and 1c, for the aqueous and sorbed phase observations, respectively (See *Neupauer and Wilson* [2004b] for a discussion of the numerical implementation of the backward probabilistic model).

[31] For the observation made in the aqueous phase at x_{1w} = 100 m, the solid line in Figure 1b represents the possible former positions of the contaminant particle at τ = 100 days before observation if it had been in the aqueous phase at that time, and the dashed line represents the possible former positions if it had been in the sorbed phase at that time. If the observed particle had been in the aqueous phase at τ = 100 days prior to observation, its most likely former position is $x_1 \approx 152$ m. If it had been in the sorbed phase, its most likely former position is $x_1 \approx 122$ m. The area under the PDFs represent the probability that the observed contaminant particle was in the given phase at any location. The results show a probability of 0.29 that the observed contaminant particle was in the aqueous phase at time τ = 100 days before observation, a probability of 0.42 that it was in the sorbed phase. The remaining probability (a probability of 0.29) accounts for the possibility that a particle released into the aquifer at τ = 100 days prior to sampling would have decayed prior to the sampling time, and therefore was not present in the aquifer at τ = 100 days.

[32] For the sorbed phase observation at x_{w2} = 150 m, Figure 1c shows the backward location probabilities at τ = 100 days before observation, for both the aqueous and the sorbed phases. For an aqueous former phase, the most likely prior location was at $x_1 \approx 172$ m; while for a sorbed former

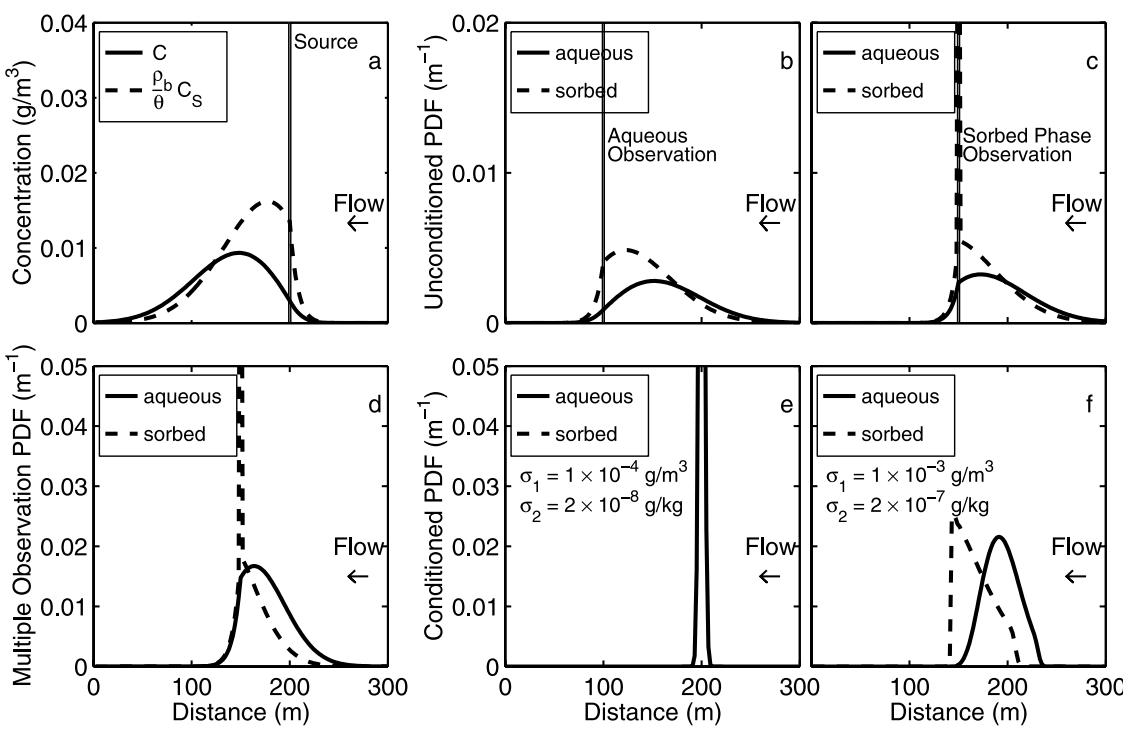

Figure 1. (a) Concentration at t = 100 days after release from an instantaneous point source in the aqueous phase at x_{1o} = 200 m. (b) Unconditioned backward location PDFs at τ = 100 days before observation in the aqueous phase at x_{w2} = 150 m. (c) Unconditioned backward location PDFs at τ = 100 days before observation in the sorbed phase at x_{w2} = 150 m (maximum value of the dashed line in subplot (c) is 0.056 m^{-1}), (d) Unconditioned multiple-observation backward location PDFs at τ = 100 days before observations at x_{w1} = 100 m in the aqueous phase and at x_{w2} = 150 m in the sorbed phase (maximum value of the dashed line in subplot (d) is 0.19 m^{-1}). (e) Conditioned backward location PDFs for τ = 100 days before observation at x_{w1} = 100 m in the aqueous phase and at x_{w2} = 150 m in the sorbed phase with small error levels. (f) Conditioned backward location PDFs for τ = 100 days before observation at x_{w1} = 100 m in the aqueous phase and at x_{w2} = 150 m in the sorbed phase with moderate error levels.

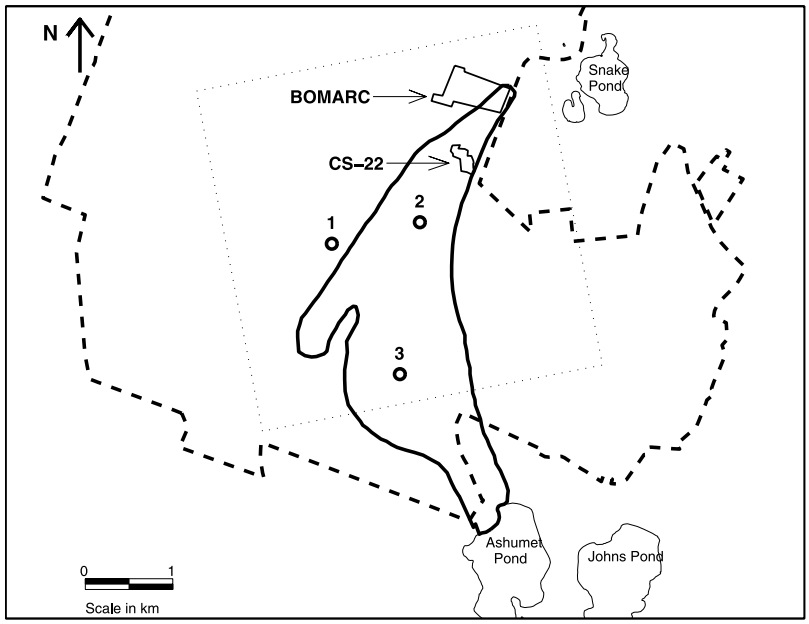

Figure 2. Inferred TCE plume (5 μg/L contour in 1997) at the Massachusetts Military Reservation (MMR). Circles denote sample locations used in the backward model. The dashed line represents MMR boundary. The dot-dash line represents the plot region used in Figure 3. Two suspected sources of contamination are BOMARC and CS-22.

phase, the most likely prior location is the observation location (x_1 = 150 m). The observed contaminant particle has a probability of 0.28 of having been in the aqueous phase, a probability of 0.43 of having been in the sorbed phase, and a probability of 0.29 of not being in the aquifer at that time (i.e, if it were in the aquifer, it would have decayed prior to observation).

[33] The unconditioned multiple observation location probability density functions, obtained from (9), are shown in Figure 1d. The solid line represents the possible source locations assuming that the two contaminant particles originated from the same source in the aqueous phase. The most likely source location is at $x_1 \approx 164$ m. The dashed line represents the possible source locations assuming that the two contaminant particles originated from the same source in the sorbed phase, and indicates that the location of the sorbed phase observation (x_1 = 150 m) is the most likely source location. The normalization in the denominator of (9) forces the total probability to be unity; therefore unlike the single observation probability density functions, it is not possible to determine the likelihood that the source was in a particular phase. In the absence of any concentration information, the unconditioned backward location probability density functions are controlled by the location of the observations.

[34] The unconditioned PDFs have a large variance, with the range of possible source locations extending from $x_1 \approx$ 90 m to $x_1 \approx 300$ m. This large range of possible source locations represents any source location that could produce an observable concentration at the observation location at t = 100 days after release from the source. Source locations that would result in high concentrations at the observation location have high values of the unconditioned PDF, and source locations that would result in low concentrations at the observation location have low values of the PDF. Since there are many possible source locations that could produce a finite (i.e., non-zero) concentration at the observation location, the unconditioned PDF has a large variance. The variance can be reduced by conditioning on the measured concentrations.

[35] The conditioned backward location PDFs, obtained from (20), are shown in Figure 1e and f for small (1-2% of the measured concentration) and moderate (10-20% of the measured concentration) levels of measurement error, respectively. For small error levels, (Figure 1e, $\sigma_1 = 1 \times 10^{-4}$ g/m^3 and $\sigma_2 = 2 \times 10^{-8}$ g/kg), the conditioned aqueous backward location PDF correctly identifies the true source location and phase (x_{1o} = 200 m, in the aqueous phase) as the most likely location of the source. The uncertainty is small, as indicated by the narrow spread of the PDF. This demonstrates that conditioning on measured concentrations can substantially improve our ability to identify the location of a contaminant source. The conditioned sorbed phase backward location PDF is essentially zero everywhere, indicating a probability of essentially zero that the contaminant was released from the source in the sorbed phase, which is consistent with the true conditions.

[36] For moderate error levels (Figure 1f, $\sigma_1 = 1 \times 10^{-3}$ g/m^3 and $\sigma_2 = 2 \times 10^{-7}$ g/kg), the conditioned aqueous phase backward location PDF indicates that the most likely source location is near the true source location (x_{1o} = 200 m); however, the uncertainty is increased relative to the PDF shown in Figure 1e for smaller error levels. Relative to the unconditioned PDF (Figure 1d), the uncertainty in this conditioned PDF is decreased, as shown by the narrower spread of the PDF. The conditioned sorbed phase backward location PDF shows that the most likely position of a sorbed phase source is near the location of the sorbed phase observation, at x_1 = 150 m. These results show that the uncertainty in the backward location PDFs increases as

Table 2. TCE Samples Used in the Backward Model

Sample	Sampling Date	Concentration (μg/L)	Standard Deviation (μg/L)
1	11 September 1996	58	5
2	21 June 2000	150	10
3	26 October 1998	203.1	5

measurement error increases. *Neupauer and Lin* [2006] have shown that as the error level increases, the concentration measurements provide less information and the results approach the unconditioned PDFs.

4. Field Application: Massachusetts Military Reservation

[37] In this section, we demonstrate the use of the conditioned backward PDF using an example of identifying the source of a trichloroethylene (TCE) plume at the Massachusetts Military Reservation (MMR). The outline of the plume in 1997 is shown in Figure 2. Two suspected sources of TCE are the former BOMARC Missile Site (see Figure 2), which was in operation between 1962 and 1973, and CS-22 (see Figure 2), where unspecified waste, possibly from the BOMARC facility, may have been disposed [*Air Force Center for Environmental Excellence* (*AFCEE*), 2000].

[38] *Zheng and Wang* [2002] developed a flow and transport model to optimize the design of the remediation system for the MMR TCE plume. *Neupauer and Wilson* [2005] used this model without the remediation system in operation to obtain a pre-remediation flow field that represents the flow field that controlled the movement of the plume. This pre-remediation flow field was used to model unconditioned backward location PDFs for the three observations of TCE shown in Figure 2 (See Table 2 for additional information). The unconditioned PDFs represent the possible positions of the three observed contaminant particles in 1962, the year that BOMARC operations began. *Neupauer and Wilson* [2005] concluded that the samples were likely to have originated near BOMARC and CS-22.

[39] The results of *Neupauer and Wilson* [2005] were obtained by ignoring natural recharge, and were based on the modified adjoint states that were conditioned on the fact that the observed contaminant particle did not decay (i.e., elimination of the decay term in (6)). As mentioned above, these modified adjoint states are not appropriate with the conditioning step, so we reproduced the unconditioned PDFs to allow for decay, and to include natural recharge. The unconditioned PDFs for 1962 are shown in Figures 3a–3c for Samples 1–3, respectively. The plot region for these plots is shown in Figure 2. Figure 3d shows the multiple-observation backward location PDF for these three samples, representing the possible source location of the three samples assuming they were released from the same location in 1962. In all cases, the unconditioned backward location

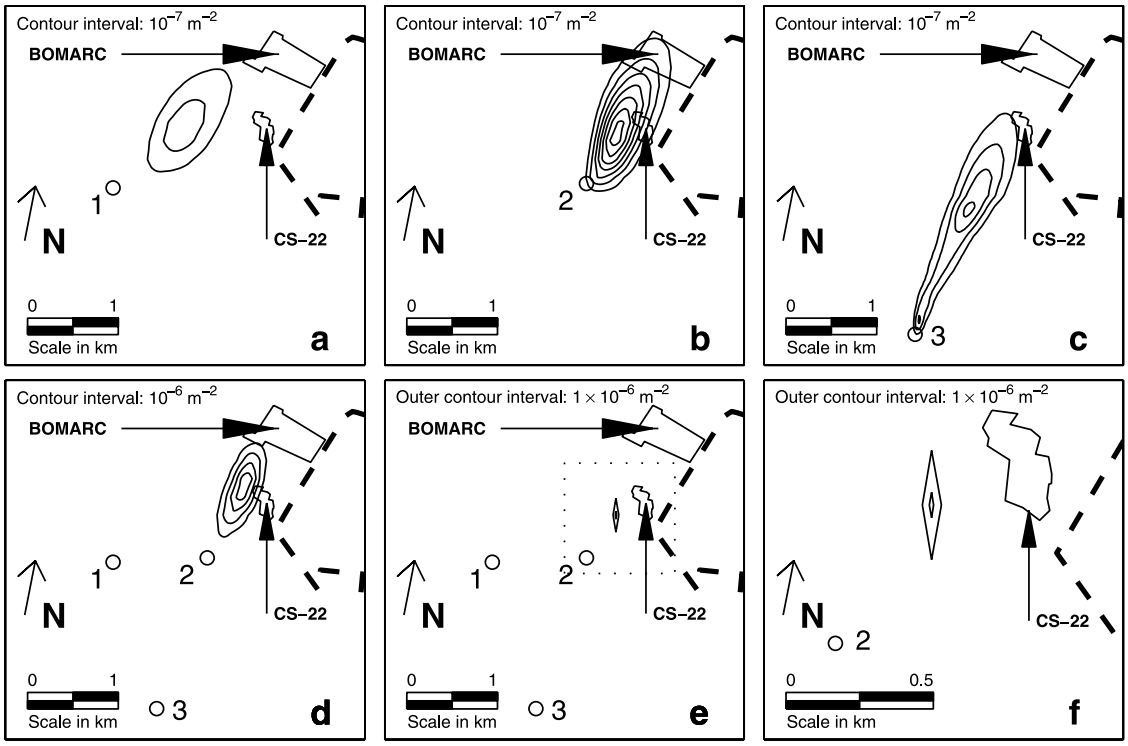

Figure 3. Aqueous phase backward location PDFs for the source of TCE at the Massachusetts Military Reservation. (a) Unconditioned PDF for Sample 1. (b) Unconditioned PDF for Sample 2. (c) Unconditioned PDF for Sample 3. (d) Unconditioned multiple observation PDF for Samples 1–3. (e) Conditioned PDF for Samples 1–3. The box outlined with a dotted line represents the plot region for subplot (f). (f) Enlargement of subplot (e). Contours in subplots (e) and (f) have values of 1×10^{-6} m^{-2} (outer) and 1×10^{-5} m^{-2} (inner).

PDFs are in the vicinity of BOMARC and CS-22, indicating that both are possible sources of contamination. The backward location PDFs are three-dimensional, but to simplify the presentation, we present vertically averaged PDFs.

[40] Next, we condition these PDFs on their measured concentrations. All three samples were observed in the aqueous phase, therefore $N_A = 3$ and $N_S = 0$ in (20). The measured concentrations and standard deviations of measurement error are shown in Table 2; these were used in (12) to obtain PDFs for the measured concentrations. The true concentrations in (12) are obtained from (18) using the adjoint states and a range of possible source masses. For this example, we used a range of source masses with 50 mass values uniformly spaced in log space between 100 g to 10,000 g, such that m_o = [100 g,. . ., (100 g)$100^{k/49}$,. . ., 10,000 g], for k = 1, 2,. . . 50. The selected range of source masses can affect the location of the conditioned backward location PDFs. To determine the appropriate source mass range, we evaluated (20) with several different source mass ranges that extended over one order of magnitude, e.g., 0.1 g – 1 g; 1 g – 10 g; 10 g – 100 g, etc, and the magnitude of the integral in (20) was analyzed. The ranges that had a relatively large value of the integral were included in the final calculation.

[41] The integration over the source mass domain in (20) was approximated as a summation, given by

$$f_{\mathbf{X}_o,P_o|\hat{\mathbf{C}},\hat{\mathbf{C}}_S}(\mathbf{x}_o,p_o=l|\hat{\mathbf{c}},\hat{\mathbf{c}}_S) \approx \beta_{\mathbf{x}} \sum_{k=1}^{50} w_k \left[\prod_{i=1}^{N_A} \frac{\theta(\mathbf{x})\phi_l^A(\mathbf{x}_o,\tau;\mathbf{x}_{wi},\tau_{wi})}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left\{ \frac{-\left[\hat{c}_i - m_{ok}\phi_l^A(\mathbf{x}_o,\tau;\mathbf{x}_{wi},\tau_{wi})\right]^2}{2\sigma_i^2} \right\} \cdot \prod_{j=N_A+1}^{N_A+N_S} \frac{\theta(\mathbf{x})\phi_l^S(\mathbf{x}_o,\tau;\mathbf{x}_{wj},\tau_{wj})}{\sqrt{2\pi\sigma_j^2}} \cdot \exp\left\{ \frac{-\left[\hat{c}_{Sj} - m_{ok}\phi_l^S(\mathbf{x}_o,\tau;\mathbf{x}_{wj},\tau_{wj})\right]^2}{2\sigma_j^2} \right\} \right], \tag{22}$$

where m_{ok} = (100 g)$100^{k/49}$, and w_k is a weighting function. Here we used $w_1 = w_2 = \cdots = w_k = \cdots = w_{50}$. The resulting conditioned backward location PDF was vertically integrated to simplify the presentation, and is shown in Figures 3e and 3f. The highest values of the conditioned backward location PDF occur to the southwest of CS-22, indicating that the region near CS-22 is the more likely source location. The results also show that conditioning substantially decreases the uncertainty.

5. Conclusion

[42] Backward location probability density functions can be used to identify former positions or source locations of contaminant particles that are observed in an aquifer. For a sorbing solute, the PDFs depend on the phase of the observation and on the former phase of the contaminant particle. Backward aqueous and sorbed phase location PDFs can be obtained by solving the adjoint of a forward contaminant transport model; however, these PDFs are based only on the observation location and time, and ignore the measured concentrations. We developed a new approach for conditioning backward location PDFs on measured concentrations. The approach consists of two steps. In the first step, unconditioned backward location PDFs are obtained for each observation by solving the adjoint of a forward contaminant transport equation. In the second step, this backward location PDF is conditioned on the measured concentrations using a formula developed through Bayes' theorem. We showed that the conditioned PDF has a smaller variance than the unconditioned PDF, and it is more accurate. As the measurement error increases, the conditioned PDF approaches the unconditioned PDF.

[43] **Acknowledgments.** The authors are grateful to Jacobs Engineering Group, Inc. at Otis Air National Guard Base for providing the TCE data, Spence Smith for providing information on the TCE contamination and remediation, and Chunmaio Zheng for providing his flow and transport model of the CS-10 remediation system. The authors thank Emil Frind, two anonymous reviewers, and the associate editor, Alberto Bellin, for their valuable comments on an earlier version of this manuscript.

References

Air Force Center for Environmental Excellence (AFCEE) (2000), Air Force will conduce site investigations (SI) for the Chemical Spill-22 (CS-22), Coast Guard Chemical Spill-8 (CS-8 [CG]) and Chemical Spill-18 (CS-18) Sites, *News Release 2000-51*, Otis Air Natl. Guard Base, Mass., 26 Sept. (Available at http://www.mmr.org/cleanup/index.htm)

Andricevic, R., and V. Cvetkovic (1996), Evaluation of risk from contaminants migrating in groundwater, *Water Resour. Res.*, *32*, 611–621.

Andricevic, R., and V. Cvetkovic (1998), Relative dispersion for solute flux in aquifers, *J. Fluid Mech.*, *361*, 145–174.

Cornaton, F., P. Perrochet, and H.-J. Diersch (2004), A finite element formulation of the outlet gradient boundary condition for convective-diffusive transport problems, *Int. J. Numer. Meth. Engng*, *61*, 2716–2732.

Cvetkovic, V., and G. Dagan (1994), Transport of kinetically sorbing solute by steady random velocity in heterogeneous porous formations, *J. Fluid Mech.*, *265*, 189–215.

Cvetkovic, V., and G. Dagan (1996), Reactive transport and immiscible flow in geologic media, II, Applications, *Proc. R. Soc. Lond. A*, *452*, 303–328.

Cvetkovic, V. D., and A. M. Shapiro (1990), Mass arrival of sorptive solute in heterogeneous porous media, *Water Resour. Res.*, *26*, 2057–2067.

Cvetkovic, V., G. Dagan, and H. Cheng (1998), Contaminant transport in aquifers with spatially variable hydraulic and sorption properties, *Proc. R. Soc. Lond. A*, *454*, 2173–2207.

Dagan, G. (1982), Stochastic modeling of groundwater flow by unconditional and conditional probabilities, 2. The solute transport, *Water Resour. Res.*, *18*, 835–848.

Dagan, G. (1984), Solute transport in heterogeneous porous formations, *J. Fluid Mech.*, *145*, 151–177.

Dagan, G. (1986), Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale, *Water Resour. Res.*, *22*, 120S–134S.

Dagan, G. (1987), Theory of solute transport in groundwater, *Annu. Rev. Fluid Mech.*, *19*, 183–215.

Dagan, G. (1990), Transport in heterogeneous porous formations: Spatial moments, ergodicity, and effective dispersion, *Water Resour. Res.*, *26*, 1281–1290.

Dagan, G., and V. Cvetkovic (1993), Spatial moments of a kinetically sorbing solute plume in a heterogeneous aquifer, *Water Resour. Res.*, *29*, 4053–4061.

Dagan, G., and V. Cvetkovic (1996), Reactive transport and immiscible flow in geologic media, I. General theory, *Proc. R. Soc. Lond. A*, *452*, 285–301.

Fitts, C. R. (1996), Uncertainty in deterministic groundwater transport models due to the assumption of macrodispersive mixing: Evidence from Cape Cod (Massachusetts, U.S.A.) and Borden (Ontario, Canada) tracer tests, *J. Contam. Hydrol.*, *21*, 69–84.

Fogg, G. E., E. M. LaBolle, and G. S. Weissmann (1999), Groundwater vulnerability assessment: Hydrogeologic perspective and example from Salinas Valley, California. Assessment of Non-Point Source Pollution in the Vadose Zone, Geophysical Monograph 108, *AGU*, 45–61.

Frind, E. O. (1988), Solution of the advection-dispersion equation with free exit boundary, *J. Numer. Meth.*, *4*, 301–314.

Jury, W. A., and K. Roth (1990), *Transfer Functions and Solute Movement through Soil: Theory and Applications*, Birkhauser, Verlag, Boston.

Kitanidis, P. K. (1988), Prediction by the method of moments or transport in a heterogeneous formation, *J. Hydrol.*, *102*, 453–473.

Lawrence, A. E., X. Sanchez-Vila, and Y. Rubin (2002), Conditional moments of the breakthrough curves of kinetically sorbing solute in heterogeneous porous media using multirate mass transfer models for sorption and desorption, *Water Resour. Res.*, *38*(11), 1238, doi:10.1029/2001WR001006.

Lin, R. (2003), Identification of groundwater contamination sources using probabilities conditioned on measured concentrations, M. S. thesis, Dept. of Civil Eng., Univ. of Virginia, Charlottesville, Virginia.

Molson, J. W., J. Beckers, E. O. Frind, and P. J. Martin (2002), WATFLOW/3D, *Version 4.0, A Three-Dimensional Numerical Model for Coupled Groundwater and Surface Water Flow*, Dept. of Earth Sciences, Univ. of Waterloo, Waterloo, Ontario, Canada.

Neupauer, R. M., and R. Lin (2006), Identifying sources of a conservative groundwater contaminant using backward probabilities conditioned on measured concentrations, *Water Resour. Res.*, *42*(3), W03424, doi:10.1029/2005WR004115.

Neupauer, R. M., and J. L. Wilson (1999), Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant, *Water Resour. Res.*, *35*(11), 3389–3398.

Neupauer, R. M., and J. L. Wilson (2001), Adjoint-derived location and travel time probabilities for a multi-dimensional groundwater system, *Water Resour. Res.*, *37*(6), 1657–1668.

Neupauer, R. M., and J. L. Wilson (2002), Backward probabilistic model of groundwater contamination in non-uniform and transient flow, *Adv. Water Resour.*, *25*(7), 733–746.

Neupauer, R. M., and J. L. Wilson (2003), Backward location and travel time probabilities for a decaying contaminant in an aquifer, *J. Contam. Hydrol.*, *66*, 39–58.

Neupauer, R. M., and J. L. Wilson (2004a), Forward and backward location probabilities for sorbing solutes in groundwater, *Adv. Water Resour.*, *27*, 689–785.

Neupauer, R. M., and J. L. Wilson (2004b), Numerical implementation of a backward probabilistic model of groundwater contamination, *Ground Water*, *42*(2), 175–189.

Neupauer, R. M., and J. L. Wilson (2005), Backward probability model using multiple observations of contamination to identify groundwater contamination sources at the Massachusetts Military Reservation, *Water Resour. Res.*, *41*, W02015, doi:10.1029/2003WR002974.

Rubin, Y., M. A. Cushey, and A. Wilson (1997), The moments of the breakthrough curves of instantaneously and kinetically sorbing solutes in heterogeneous geologic media: Prediction and parameter inference from field measurement, *Water Resour. Res.*, *33*(11), 2465–2481.

Sanchez-Vila, X., and Y. Rubin (2003), Travel time moments for sorbing solutes in heterogeneous domains under nonuniform flow conditions, *Water Resour. Res.*, *39*(4), 1086, doi:10.1029/2002WR001399.

Selroos, J.-O., and V. Cvetkovic (1994), Mass flux statistics of kinetically sorbing solute in heterogeneous aquifers: Analytical solution and comparison with simulations, *Water Resour. Res.*, *30*, 63–69.

Shapiro, A. M., and V. D. Cvetkovic (1988), Stochastic analysis of solute arrival time in heterogeneous porous media, *Water Resour. Res.*, *24*(10), 1711–1718.

Uffink, G. J. M. (1989), Application of Kolmogorov's backward equation in random walk simulations of groundwater contaminant transport, in *Contaminant Transport in Groundwater*, edited by H. E Kobus and W Kinzelbach, 283–289, Vt: A.A. Balkema, Brookfield.

Valocchi, A. J., and H. A. M. Quinodoz (1989), Application of the random walk method to simulate the transport of kinetically adsorbing solutes, Proceedings of the Third IAHS Scientific Assembly, Baltimore, MD, *IAHS Publ.*, *185*, 35–42.

Zheng, C., and P. P. Wang (1999), *MT3DMS: Documentation and User's Guide*, Release DoD_3.50.A, U.S. Army Corps of Eng., Contract Report, SERDP-99.

Zheng, C., and P. P. Wang (2002), A field demonstration of the simulation-optimization approach for remediation system design, *Ground Water*, *40*(3), 258–265.

R. Lin, Bank of Montreal, 100 King Street West, M5X 1A1, Toronto, Ontario, Canada.

R. M. Neupauer and H. O'Shea, Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, CO 80309, USA. (neupauer@colorado.edu)