

Application of geostatistical inverse modeling to contaminant source identification at Dover AFB, Delaware

Anna M. Michalak and Peter K. Kitanidis, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020

1. Introduction

Interest in techniques aimed at identifying sources of environmental contaminants has been growing over the past several years. The ability to conclusively identify the source of observed contamination can not only help in the remediation process, but can be critical to the identification of responsible parties, and therefore to the apportionment of any liability associated with a given site.

Methods that are currently available for contaminant source identification differ in their ranges of applicability in terms of contaminants and media, the level of confidence associated with their results and a variety of other factors. These methods can, however, be subdivided into three categories. The first category is compositional analysis, which determines the source by analyzing differences in the molecular or isotopic compositions of contaminants among potential sources. The second category is the use of either naturally occurring or introduced tracers discharged with a contaminant at a given potential source. The third category encompasses methods based on conclusions made from the contamination distribution itself, once the transport behavior of the contaminant and the medium in which it is being transported have been quantified. A review of these methods is available in Michalak (2001).

One set of methods based on analyzing the contamination distribution is inverse methods. These methods use modeling and statistical tools to determine either the prior location of observed contamination or the release history from a known source. Reviews of such methods are presented in Snodgrass and Kitanidis (1997) and Liu and Ball (1999), among others.

One subset of work uses a function estimate

to characterize the source location or release history. In this case, the source characteristics are not limited to a set number of parameters, but are instead free to vary in space and in time. This category includes methods that use a deterministic approach and others that offer a stochastic approach to the problem.

The assumption that the model parameters are deterministic but unknown differs from stochastic approaches, where parameters are viewed as jointly distributed random fields. Because there will always be uncertainty in contaminant concentration estimates, release history and release location, it makes sense to treat these quantities as random functions that can be described by their statistical properties. In this framework, estimation uncertainty is recognized and its importance can sometimes be determined.

One of the stochastic methods proposed in the past for the estimation of the release history of a contaminant is the use of the geostatistical approach to inverse modeling. Snodgrass and Kitanidis (1997) estimated the release history of a conservative solute being transported in a 1-dimensional homogeneous domain, given point concentration measurements at some time after the release.

In this paper, we present the first application of the geostatistical approach to contaminant source identification to the interpretation of field data. This work is also the first demonstration of the applicability of this approach to a physically non-uniform domain. Finally, we develop a Markov Chain Monte Carlo method for enforcing concentration non-negativity while maintaining the statistical rigor of the geostatistical methodology.

Aquitard cores taken from the Dover Air Force Base (DAFB) in Delaware are analyzed to infer the contamination history in the overlying aquifer. These data sets have

previously been examined by Ball et al. (1997) and Liu and Ball (1999). Ball et al. (1997) assumed that the history was made up of one-step and two-step constant concentrations at the aquifer/aquitard interface and the times of step concentration changes were estimated from the data. Liu and Ball (1999) applied Tikhonov regularization to obtain a function estimate of the concentration history. However, whereas Tikhonov regularization is a deterministic technique that identifies a single estimate of the source function, a geostatistical approach allows for more in-depth analysis. This method results in a best estimate that is the median of all possible contamination histories, as well as confidence intervals about that best estimate. Furthermore, conditional realizations can be generated which allow for better visualization of the unknown process. Finally, structural parameters used in the analysis that describe the continuity of the contamination-history function are optimized using the data itself.

2. Site and Data Description

The research site is located at DAFB. At the site, an unconfined sand aquifer is underlain by an aquitard, which consists of two layers of distinctly different characteristics: an upper layer of orange silty clay loam (OSCL) and a bottom layer of dark gray silt loam (DGSL). Tetrachloroethene (PCE) and trichloroethene (TCE) are two principal chemical contaminants of the overlying aquifer contaminant plume, and concentration profiles for both these chemicals have been obtained in the underlying aquitard at several locations. A detailed description of the site geology and hydrogeology can be found in Mackay et al. (1997) and Ball et al. (1997). A description of the sampling at the site is available in Liu and Ball (1999). The data sets used for the analysis presented in this work are at locations referred to as PPC11 and PPC13.

The soil core samples were also used to independently determine the sorption properties and porosity of the two aquitard layers (Ball et al., 1997). The physical parameters as used by

Ball et al. (1997) are presented in Table 1. Identical values were used in the current work, in order to facilitate a direct comparison between the two methods.

Table 1. Summary of parameters in two-layer aquitard

Physical Definition	Parameter	Unit	Layer 1 (OSCL)	Layer 2 (DGSL)
Effective diffusivity	D (PCE)	m ² /s	4.2×10 ⁻¹⁰	4.2×10 ⁻¹⁰
	D (TCE)	m ² /s	4.9×10 ⁻¹⁰	4.9×10 ⁻¹⁰
Retardation factor	R (PCE)		2	45
	R (TCE)		1.4	20
Porosity	η		0.53	0.56
Bulk density	ρ _b	kg/L	1.22	1.15

3. Model

Physical Model

Solute transport in this two-layer aquitard is mainly controlled by a diffusive process which is assumed to be mathematically described by the following differential equation:

$$R_1 \frac{\partial c_1^{aq}}{\partial t} = D_1 \frac{\partial^2 c_1^{aq}}{\partial x^2} \quad 0 < x < L$$

$$R_2 \frac{\partial c_2^{aq}}{\partial t} = D_2 \frac{\partial^2 c_2^{aq}}{\partial x^2} \quad L < x < +\infty$$

where c_1^{aq} and c_2^{aq} are aqueous concentrations, R_1 and R_2 are retardation factors, D_1 and D_2 are effective diffusion coefficients in layer 1 (OSCL) and layer 2 (DGSL), respectively; x is the depth within the aquitard, L is the thickness of the first layer (OSCL) and is 0.74m for location PPC11 and 0.91m for location PPC13.

Inverse Model

The objective is to estimate an unknown function. The standard estimation problem may be expressed in the following form:

$$\mathbf{z} = \mathbf{h}(\mathbf{s}, \mathbf{r}) + \boldsymbol{\varepsilon}$$

where \mathbf{z} is an $n \times 1$ vector of observations and \mathbf{s} is an $m \times 1$ "state vector" obtained from the discretization of the unknown function that we wish to estimate. The vector \mathbf{r} contains other parameters needed by the model function $\mathbf{h}(\mathbf{s}, \mathbf{r})$ (e.g. diffusivity). The measurement error is represented by the vector $\boldsymbol{\varepsilon}$. Following geostatistical methodology, \mathbf{s} and $\boldsymbol{\varepsilon}$ are represented as random vectors. In this case, the function $\mathbf{h}(\mathbf{s}, \mathbf{r})$ is linear in \mathbf{s} such that

$$\mathbf{h}(\mathbf{s}, \mathbf{r}) = \mathbf{H}\mathbf{s}$$

where \mathbf{H} is a known matrix.

In the 1-D case we have an analytical solution for the forward problem (Liu and Ball 1999):

$$c(x, T) = \int_0^T s(t) f(x, T-t) dt$$

where c is the total concentration (aqueous and sorbed) and T is measurement time. The source is a function of time and is expressed by $s(t)$. The transfer function $f(x, T-t)$ applies the appropriate weight to the source function:

$$f(x, T-t) = \frac{\eta_1 R_1}{\rho_{b1}} \left(\frac{R_1}{D_1} \right)^{1/2} \sum_{i=0}^{\infty} \vartheta^i \left(\frac{2iL+x}{2[\pi(T-t)^3]^{1/2}} \right) \cdot \exp\left(-\frac{R_1(2iL+x)^2}{4D_1(T-t)} \right) - \vartheta \frac{(2i+2)t-x}{2[\pi(T-t)^3]^{1/2}} \cdot \exp\left(-\frac{R_1[(2i+2)Ltx]^2}{4D_1(T-t)} \right) \quad 0 < x < L$$

$$f(x, T-t) = \frac{\eta_2 R_2}{\rho_{b2}} \left(\frac{R_1}{D_1} \right)^{1/2} (1-\vartheta) \sum_{i=0}^{\infty} \vartheta^i \frac{\gamma_i}{2[\pi(T-t)^3]^{1/2}} \cdot \exp\left(-\frac{R_1 \gamma_i^2}{4D_1(T-t)} \right) \quad L < x < \infty$$

where

$$\vartheta = \frac{\eta_2 (D_2 R_2)^{1/2} - \eta_1 (D_1 R_1)^{1/2}}{\eta_2 (D_2 R_2)^{1/2} + \eta_1 (D_1 R_1)^{1/2}}$$

$$\gamma_i = (2i+1)L + \left(\frac{D_1 R_2}{D_2 R_1} \right)^{1/2} (x-L)$$

where the subscripts 1 and 2 refer to parameter values in the upper and lower aquitard layers, respectively.

Let $x_i, i=1, \dots, n$ be the n points at which the measurements are taken, and let us discretize the time domain into m temporal points $t_j, j=1, \dots, m$, with a time step $\Delta t = T/m$. All measurements are taken at time $t=T$. In this case, the sensitivity matrix is:

$$\mathbf{H} = \Delta t \begin{bmatrix} f(x_1, T-t_1) & \dots & f(x_1, T-t_m) \\ f(x_2, T-t_1) & \dots & f(x_2, T-t_m) \\ \vdots & \ddots & \vdots \\ f(x_n, T-t_1) & \dots & f(x_n, T-t_m) \end{bmatrix}$$

We shall assume that $\boldsymbol{\varepsilon}$ has zero mean and known covariance matrix \mathbf{R} . The covariance of

the measurement errors used is

$$\mathbf{R} = \sigma_R^2 \mathbf{I}$$

where σ_R^2 is the variance of the measurement error, and \mathbf{I} is an $n \times n$ identity matrix.

Furthermore, we will model \mathbf{s} , the unknown, as a random vector with expected value

$$E[\mathbf{s}] = \mathbf{Y}\boldsymbol{\beta}$$

where \mathbf{Y} is a known $m \times p$ matrix and $\boldsymbol{\beta}$ are p unknown drift coefficients. For this problem, a linear but unknown trend in the contaminant release concentration was assumed. Thus

$$\mathbf{Y} = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}$$

The covariance function of \mathbf{s} is

$$\mathbf{Q}(\boldsymbol{\theta}) = E[(\mathbf{s} - \mathbf{Y}\boldsymbol{\beta})(\mathbf{s} - \mathbf{Y}\boldsymbol{\beta})^T]$$

where $\mathbf{Q}(\boldsymbol{\theta})$ is a known function of unknown parameters $\boldsymbol{\theta}$. The generalized covariance matrix \mathbf{Q} was assumed to have cubic form:

$$Q(t_i, t_j | \boldsymbol{\theta}) = \theta |t_i - t_j|^3$$

where $|t_i - t_j|$ is the separation distance (in units of time). Such a covariance results in smooth estimates, in the sense that the second derivatives are minimized.

The approach used to obtain the structural parameters is detailed by Kitanidis (1995). In short, the parameters, in this case $\boldsymbol{\theta}$ and σ_R^2 , are estimated by maximizing the probability of the measurements given these parameters:

$$p(\mathbf{z} | \boldsymbol{\theta}, \sigma_R^2) \propto$$

$$|\Sigma|^{-1/2} |\mathbf{Y}^T \mathbf{H}^T \Sigma^{-1} \mathbf{H} \mathbf{Y}|^{-1/2} \exp\left[-\frac{1}{2} \mathbf{z}^T \boldsymbol{\Xi} \mathbf{z} \right]$$

where

$$\Sigma = \mathbf{H} \mathbf{Q} \mathbf{H}^T + \mathbf{R}$$

$$\boldsymbol{\Xi} = \Sigma^{-1} - \Sigma^{-1} \mathbf{H} \mathbf{Y} (\mathbf{Y}^T \mathbf{H}^T \Sigma^{-1} \mathbf{H} \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{H}^T \Sigma^{-1}$$

and $|\cdot|$ denotes matrix determinant.

Once these parameters have been estimated, the minimization problem for estimating the concentration history in the aquifer overlying the aquitard is:

$$J = (\mathbf{z} - \mathbf{h}(\mathbf{s}, \mathbf{r}))^T \mathbf{R}^{-1} (\mathbf{z} - \mathbf{h}(\mathbf{s}, \mathbf{r})) + (\mathbf{s} - \mathbf{Y}\boldsymbol{\beta})^T \mathbf{Q}^{-1} (\mathbf{s} - \mathbf{Y}\boldsymbol{\beta})$$

In this case, the vector of observations \mathbf{z} and that of the unknown function \mathbf{s} are:

$$\mathbf{z} = \begin{bmatrix} z(x_1, T) \\ z(x_2, T) \\ \vdots \\ z(x_n, T) \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} s(t_1) \\ s(t_2) \\ \vdots \\ s(t_m) \end{bmatrix}$$

The corresponding system that needs to be solved is:

$$\begin{bmatrix} \Sigma & \vdots & \mathbf{HY} \\ \dots & \dots & \dots \\ (\mathbf{HY})^T & \vdots & 0 \end{bmatrix} \begin{bmatrix} \Lambda^T \\ \dots \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{HQ} \\ \dots \\ \mathbf{Y}^T \end{bmatrix}$$

where Λ is a $m \times n$ matrix of coefficients and \mathbf{M} is a $p \times m$ matrix of multipliers. The best estimate of the function is

$$\hat{\mathbf{s}} = \Lambda \mathbf{z}$$

and its covariance is

$$\mathbf{V} = -\mathbf{YM} + \mathbf{Q} - \mathbf{QH}^T \Lambda^T$$

Using geostatistical methodology, it is also possible to generate realizations of the unknown contamination history that are conditional on all the observations. Viewing a number of conditional realizations can aid in visualizing the unknown function and the uncertainty about the best estimate. The procedure for generating conditional realizations is discussed by Gutjahr et al. (1994) and Kitanidis (1995). First, an unconditional unconstrained realization $\mathbf{s}_{uu,l}$ is generated. A realization of the error vector $\boldsymbol{\varepsilon}_l$ must also be independently generated with zero mean and covariance \mathbf{R} . Then, the conditional unconstrained realization $\mathbf{s}_{cu,l}$ may be found by minimizing

$$\begin{aligned} & (\mathbf{s}_{cu,l} - \mathbf{s}_{uu,l})^T \mathbf{G} (\mathbf{s}_{cu,l} - \mathbf{s}_{uu,l}) \\ & + (\mathbf{z} + \boldsymbol{\varepsilon}_l - \mathbf{h}(\mathbf{s}_{cu,l}))^T \mathbf{R}^{-1} (\mathbf{z} + \boldsymbol{\varepsilon}_l - \mathbf{h}(\mathbf{s}_{cu,l})) \end{aligned}$$

with respect to $\mathbf{s}_{cu,l}$. Here

$$\mathbf{G} = \mathbf{Q}^{-1} - \mathbf{Q}^{-1} \mathbf{Y} (\mathbf{Y}^T \mathbf{Q}^{-1} \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{Q}^{-1}$$

Enforcing Concentration Non-Negativity

The advantage of using a stochastic approach to the inverse problem is that physically significant confidence intervals and conditional realizations can be obtained in addition to a best estimate. We wish to enforce

concentration non-negativity without jeopardizing the statistical rigor of the methodology. The traditional approaches to enforcing constraints have some limitations. The use of Lagrange multipliers (Gill et al., 1986) is not sufficient to guarantee that the conditional realizations are equiprobable, whereas the use of data transformations (Kitanidis, 1997) may result in confidence intervals that are highly asymmetrical and absolute value dependent, and may exhibit convergence problems.

A Markov Chain Monte Carlo (MCMC) method is used here in combination with the application of Lagrange multipliers. Specifically, a Metropolis-Hastings algorithm (Chib and Greenberg, 1995) is applied to conditional realizations generated using a cubic semivariogram and modified to be nonnegative using Lagrange multipliers (Gill et al., 1986). Although MCMC algorithms have traditionally been used to estimate model parameters, they are being applied here to the estimation of the time-dependent boundary condition.

The unconditional realizations used to obtain the candidate conditional realizations are sequentially correlated:

$$\mathbf{s}_{uu,c} = \phi \mathbf{s}_{uu,l} + \alpha \mathbf{u}_c$$

where \mathbf{u}_c is an independently generated unconditional realization, $\mathbf{s}_{uu,l}$ is the unconditional realization used in the generation of the last accepted realization, and

$$0 < \phi < 1, \quad \alpha = \sqrt{1 - \phi^2}$$

An conditional unconstrained realization $\mathbf{s}_{cu,c}$ is generated from $\mathbf{s}_{uu,c}$ using the geostatistical procedure described earlier, and the candidate conditional constrained realization, $\mathbf{s}_{cc,c}$, is then obtained by applying the method of Lagrange multipliers to $\mathbf{s}_{cu,c}$.

The candidate constrained conditional realizations generated in this fashion are accepted or rejected based on their posterior probability relative to that of the last accepted realization $\mathbf{s}_{cc,l}$. The probability of acceptance is defined as:

$$\zeta(\mathbf{s}_{cc,c} | \mathbf{s}_{cc,l}) = \min \left\{ \frac{p''(\mathbf{s}_{cc,c})q(\mathbf{s}_{cc,l} | \mathbf{s}_{cc,c})}{p''(\mathbf{s}_{cc,l})q(\mathbf{s}_{cc,c} | \mathbf{s}_{cc,l})}, 1 \right\},$$

where ζ is the probability of acceptance of the new candidate realization $\mathbf{s}_{cc,c}$, $p''(\mathbf{s}_{cc,\cdot})$ is the posterior probability distribution:

$$p''(\mathbf{s}_{cc,\cdot}) \propto \exp \left[-\frac{1}{2} (\mathbf{z} - \mathbf{H}\mathbf{s}_{cc,\cdot})^T \mathbf{R}^{-1} (\mathbf{z} - \mathbf{H}\mathbf{s}_{cc,\cdot}) - \frac{1}{2} (\mathbf{s}_{cc,\cdot} - \mathbf{Y}\boldsymbol{\beta})^T \mathbf{Q}^{-1} (\mathbf{s}_{cc,\cdot} - \mathbf{Y}\boldsymbol{\beta}) \right]$$

and $q(\mathbf{s}_{cc,\cdot} | \mathbf{s}_{cc,\cdot})$ is the transition probability from one realization to the other. This probability is approximated by $q(\mathbf{s}_{uu,\cdot} | \mathbf{s}_{uu,\cdot})$, the transition probability from one unconditional realization to the other:

$$q(\mathbf{s}_{uu,c} | \mathbf{s}_{uu,l}) \propto \exp \left[-\frac{1}{2} (\mathbf{s}_{uu,c} - \phi \mathbf{s}_{uu,l})^T \frac{\mathbf{Q}^{-1}}{\alpha^2} (\mathbf{s}_{uu,c} - \phi \mathbf{s}_{uu,l}) \right]$$

A number u is sampled from a uniform distribution in the range $[0,1]$. If this number is less than $\zeta(\mathbf{s}_{cc,c} | \mathbf{s}_{cc,l})$ then $\mathbf{s}_{cc,l+1} = \mathbf{s}_{cc,c}$,

otherwise, $\mathbf{s}_{cc,l+1} = \mathbf{s}_{cc,l}$.

The chain is run until the probability space has been appropriately sampled. The chain exhibits stationary properties in terms of the posterior probability of its members from the start, and does not require a wind up period for which all resulting realizations are ultimately discarded.

4. Results and Discussion

PCE

The results presented in Figures 1 and 2 were generated by applying the unconstrained methodology to the PCE concentrations profiles measured at PPC11 and PPC13, respectively. Figures 1a and 2a show the estimated boundary concentration with 95% confidence intervals, Figures 1b and 2b show sample conditional realizations, and Figures 1c and 2c show the actual concentration profiles in the aquitard cores taken at these locations along with the fitted concentrations resulting from the best estimate of the boundary concentration.

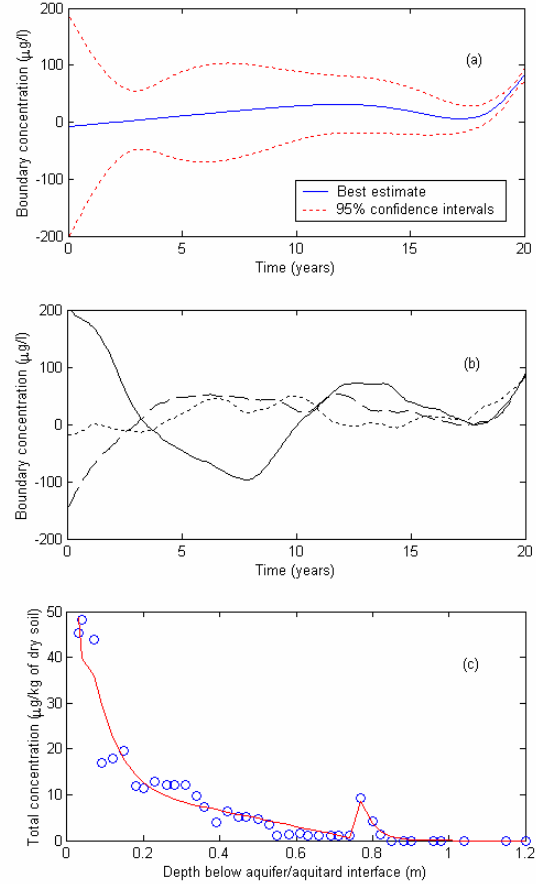


Figure 1. Results of source estimation from PCE data at location PPC11. (a) Estimated time variation of boundary concentration at the interface between the aquifer and aquitard. The end time represents the sampling date (October 27, 1994). (b) Conditional realizations of boundary concentrations. (c) Measurement data and fitted concentrations resulting from the estimated boundary conditions.

Figure 3 and 4 show the same information for the case where concentration non-negativity is enforced.

For all the data sets, the structural parameters θ in the generalized covariance function and the variance of the measurement error σ_R^2 were optimized using the method described for the unconstrained geostatistical approach. The optimal parameters are presented in Table 2, and were used for both the unconstrained and constrained solutions. Time was discretized at one-month intervals. The constrained best estimates and confidence intervals were calculated from chains of 50000 conditional realizations ($\phi = 0.99$).

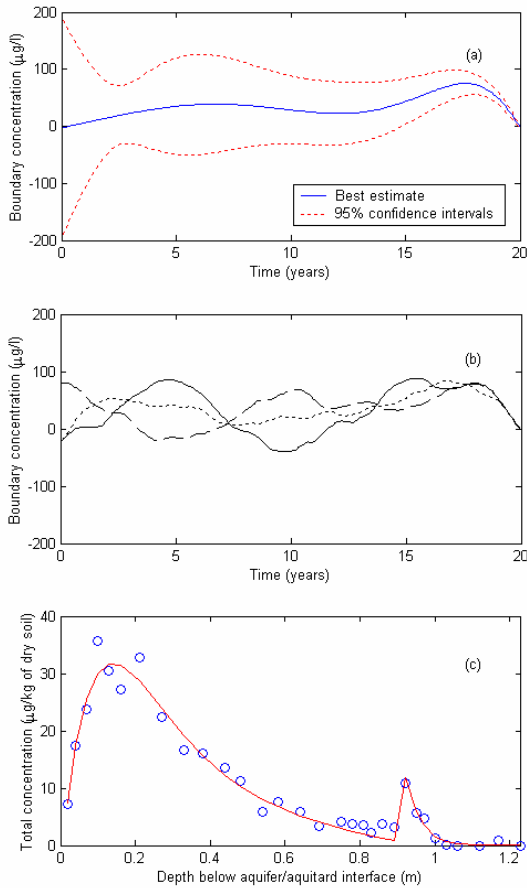


Figure 2. Results of source estimation from PCE data at location PPC13. (a) Estimated time variation of boundary concentration at the interface between the aquifer and aquitard. The end time represents the sampling date (June 6, 1996). (b) Conditional realizations of boundary concentrations. (c) Measurement data and fitted concentrations resulting from the estimated boundary conditions.

Our results for location PPC11 are consistent with those presented by Liu and Ball (1999), indicating an increase in PCE concentrations in the aquifer in recent times, with a double-peaked distribution for the best estimate when non-negativity is enforced, although only a single-peaked is present when the solution is unconstrained. However, as can be seen from the conditional realizations presented in Figure 1a and 3a, a double peaked distribution is not the only possible explanation for the observed data.

As noted by Liu and Ball (1999), the boundary conditions estimated using Tikhonov regularization and those estimated using a two-

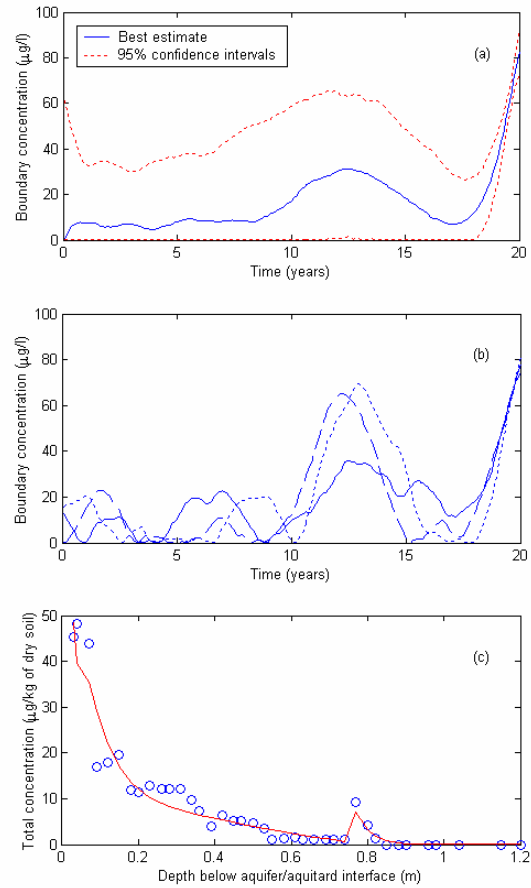


Figure 3. Results of source estimation from PCE data at location PPC11 with non-negativity constraint. (a) Estimated time variation of boundary concentration at the interface between the aquifer and aquitard. The end time represents the sampling date (October 27, 1994). (b) Conditional realizations of boundary concentrations. (c) Measurement data and fitted concentrations resulting from the estimated boundary conditions.

step approach were quite different, and yet reproduced the observed concentration profiles equally well. This point emphasizes the advantages of a stochastic approach, because the uncertainty associated with the estimated boundary conditions can be quantified. The ability to generate confidence intervals and conditional realizations greatly improves the ability to interpret obtained results.

Overall, current results indicate that the diffusive process that lead to the contamination of the aquitard, combined with the significant concentration measurement error, result in relatively wide confidence intervals about the

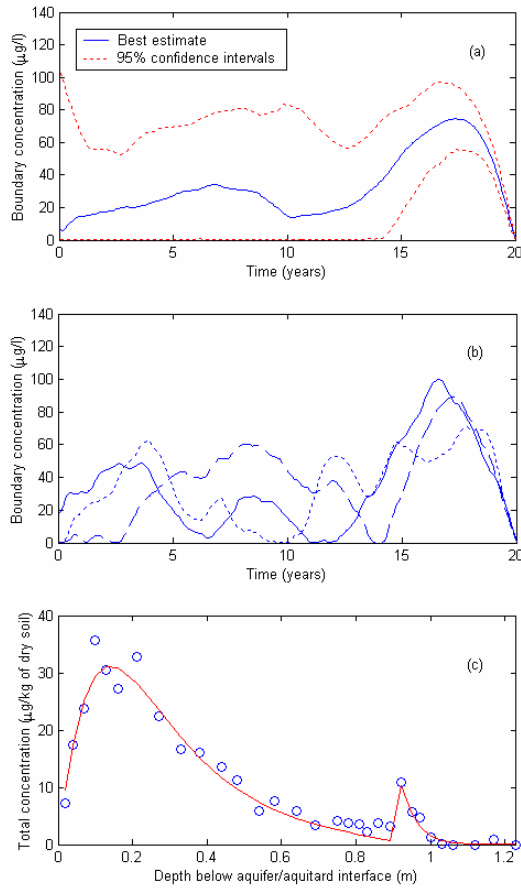


Figure 4. Results of source estimation from PCE data at location PPC13 with non-negativity constraint. (a) Estimated time variation of boundary concentration at the interface between the aquifer and aquitard. The end time represents the sampling date (June 6, 1996). (b) Conditional realizations of boundary concentrations. (c) Measurement data and fitted concentrations resulting from the estimated boundary conditions.

Table 2. Optimal structural parameter values

Structural Parameter	PCE		TCE	
	PPC11	PPC13	PPC11	PPC13
θ [month ⁻³]	2.43×10^{-2}	3.31×10^{-2}	2.12×10^3	7.56×10^2
σ_R^2 [($\mu\text{g}/\text{kg}$) ²]	11.7	4.19	2.10×10^4	1.67×10^4

estimated contamination history in the overlying aquifer. However, the introduction of additional information into the system in the form of a non-negativity constraint greatly reduced the width of the confidence intervals. The results obtained by Liu and Ball (1999) fall within the obtained confidence intervals, but the conditional realizations presented in Figures 3 and 4 show the variety of concentration histories that may have lead to the observed

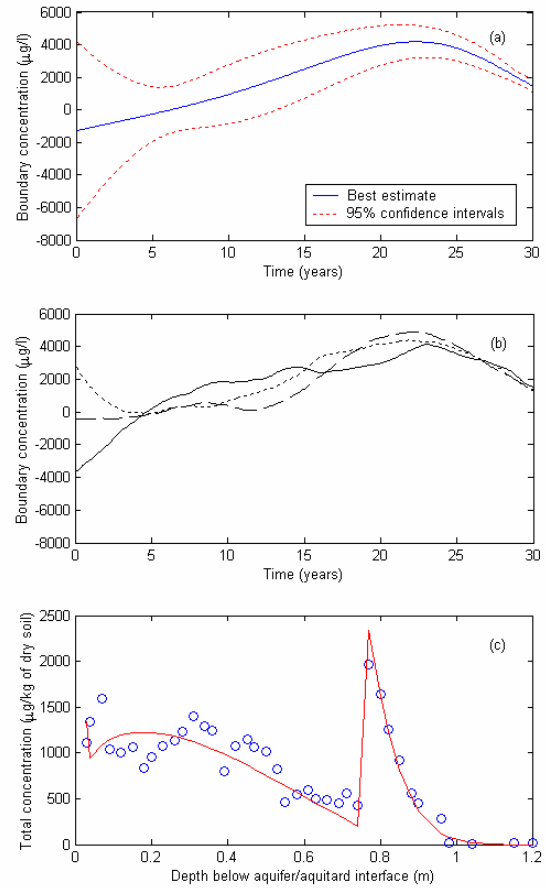


Figure 5. Results of source estimation from TCE data at location PPC11. (a) Estimated time variation of boundary concentration at the interface between the aquifer and aquitard. The end time represents the sampling date (October 27, 1994). (b) Conditional realizations of boundary concentrations. (c) Measurement data and fitted concentrations resulting from the estimated boundary conditions.

TCE

The results presented in Figures 5 and 6 were generated by applying the unconstrained methodology to the TCE concentrations profiles measured at PPC11 and PPC13, respectively. Figures 5a and 6a show the estimated boundary concentration with confidence intervals, Figures 5b and 6b show sample realizations, and Figures 5c and 6c show the actual concentration profiles in the aquitard cores taken at these locations along with the fitted concentrations resulting from the best estimate of the boundary concentration.

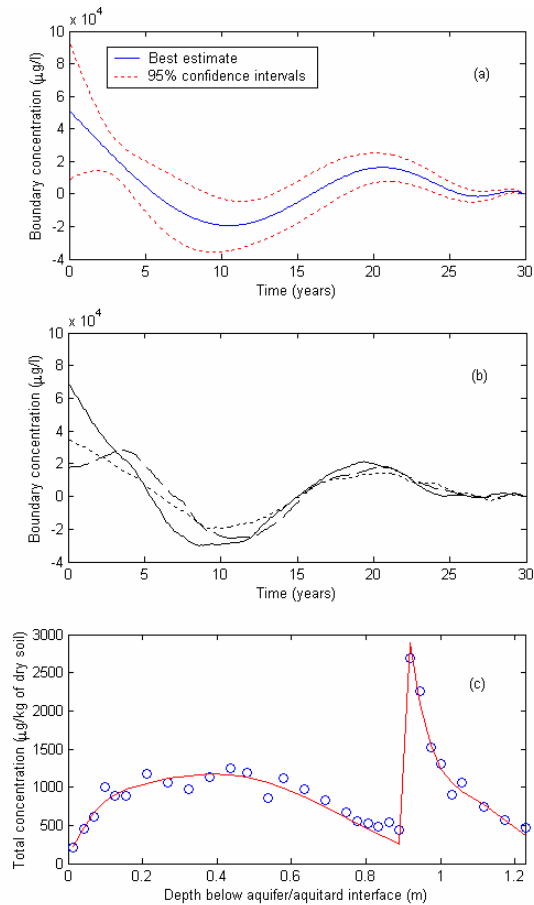


Figure 6. Results of source estimation from TCE data at location PPC13. (a) Estimated time variation of boundary concentration at the interface between the aquifer and aquitard. The end time represents the sampling date (June 6, 1996)(b) Conditional realizations of boundary concentrations. (c) Measurement data and fitted concentrations resulting from the estimated boundary conditions.

Figure 7 and 8 show the same information for the case where concentration non-negativity is enforced.

As with the PCE data, the addition of the non-negativity constraint greatly improves the results, as evidenced by the narrower confidence intervals. The results obtained by Liu and Ball (1999) fall within these intervals for location PPC13, but this earlier work had indicated a double peak for location PPC11. This second peak had been inconsistent with results at location PPC13, and does not appear in the current results.

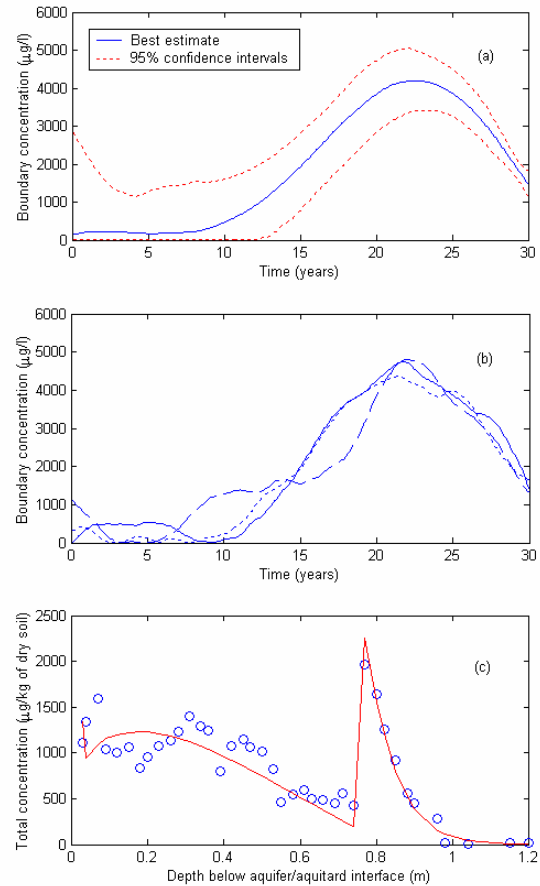


Figure 7. Results of source estimation from TCE data at location PPC11 with non-negativity constraint. (a) Estimated time variation of boundary concentration at the interface between the aquifer and aquitard. The end time represents the sampling date (October 27, 1994). (b) Conditional realizations of boundary concentrations. (c) Measurement data and fitted concentrations resulting from the estimated boundary conditions.

5. Conclusions

This paper demonstrates the applicability to field data of a stochastic inverse modeling technique based on geostatistical principles. Furthermore, the robustness of a geostatistical approach when applied to a non-uniform domain is demonstrated. Finally, a new method for enforcing concentration non-negativity is developed using a Metropolis-Hastings MCMC algorithm combined with the application of Lagrange multipliers.

Results show that the history of contamination overlying the sampled aquitard

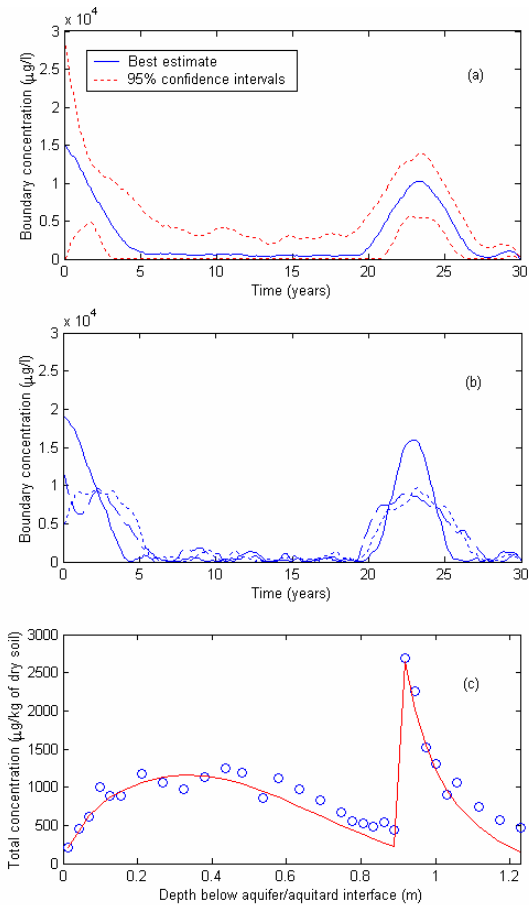


Figure 8. Results of source estimation from TCE data at location PPC13 with non-negativity constraint. (a) Estimated time variation of boundary concentration at the interface between the aquifer and aquitard. The end time represents the sampling date (June 6, 1996). (b) Conditional realizations of boundary concentrations. (c) Measurement data and fitted concentrations resulting from the estimated boundary conditions.

can be estimated with reasonable precision. This precision is greatly improved by the additional information provided by the non-negativity constraint.

6. Acknowledgments

This research was partially funded by the Natural and Accelerated Bioremediation Research (NABIR) program, Biological and Environmental Research (BER), U.S. Department of Energy (grant # DE-FG03-00ER63046).

7. References

- Ball, W. P., C. Liu, et al. (1997). "A diffusion-based interpretation of tetrachloroethene and trichloroethene concentration profiles in a groundwater aquitard." *Water Resour. Res.* **33**(12): 2741-2757.
- Chib, S. and E. Greenberg (1995). "Understanding the Metropolis-Hastings Algorithm." *The American Statistician* **49**(4): 327-335.
- Gill, P. E. (1986). *Practical Optimization*. San Diego, Academic Press.
- Gutjahr, A., B. Bullard, et al. (1994). "Joint conditional simulations and the spectral approach for flow modeling." *Stoch. Hydrol. Hydraul.* **8**(1): 79-108.
- Kitanidis, P. K. (1995). "Quasi-linear geostatistical theory for inversing." *Water Resources Research* **31**(10): 2411-2519.
- Kitanidis, P. K. (1997). *Introduction to Geostatistics Applications in Hydrogeology*. New York, Cambridge University Press.
- Liu, C. and W. P. Ball (1999). "Application of Inverse Methods to Contaminant Source Identification from Aquitard Diffusion Profiles at Dover AFB, Delaware." *Water Resources Research* **35**(7): 1975-1985.
- Mackay, D. M., R. D. Wilson, et al. (1997). Field and laboratory studies of pulsed pumping for cleanup of contaminated aquifers, Final Rep. AL/EQ-TR-1997-0017, Armstrong Lab. Environics Dir., Tyndall AFB, Fla.
- Michalak, A. M. (2001). Feasibility of contaminant source identification for property rights enforcement. *The Technology of Property Rights*. T. L. Anderson and P. J. Hill. Lanham, MD, Rowman and Littlefield Publishers, Inc.: 123-145.
- Snodgrass, M. F. and P. K. Kitanidis (1997). "A geostatistical approach to contaminant source identification." *Water Resources Research* **33**(4): 537-546.
- Tierney, L. (1994). "Markov Chains for Exploring Posterior Distributions" (with discussion), *Annals of Statistics* **22**: 1702-1762.