Inverse Modeling of Subsurface Flow and Transport Properties: A Review with New Developments

Jasper A. Vrugt,* Philip H. Stauffer, Th. Wöhling, Bruce A. Robinson, and Velimir V. Vesselinov

Many of the parameters in subsurface flow and transport models cannot be estimated directly at the scale of interest, but can only be derived through inverse modeling. During this process, the parameters are adjusted in such a way that the behavior of the model approximates, as closely and consistently as possible, the observed response of the system under study for some historical period of time. We briefly review the current state of the art of inverse modeling for estimating unsaturated flow and transport processes. We summariz how the inverse method works, discuss the historical background that led to the current perspectives on inverse modeling, and review the solution algorithms used to solve the parameter estimation problem. We then highlight our recent work at Los Alamos related to the development and implementation of improved optimization and data assimilation methods for computationally efficient calibration and uncertainty estimation in complex, distributed flow and transport models using parallel computing capabilities. Finally, we illustrate these developments with three different case studies, including (i) the calibration of a fully coupled three-dimensional vapor extraction model using measured concentrations of volatile organic compounds in the subsurface near the Los Alamos National Laboratory, (ii) the multiobjective inverse estimation of soil hydraulic properties in the HYDRUS-1D model using observed tensiometric data from an experimental field plot in New Zealand, and (iii) the simultaneous estimation of parameter and states in a groundwater solute mixture model using data from a multitracer experiment at Yucca Mountain, Nevada.

ABBREVIATIONS: CPU, central processing unit; EnKF, ensemble Kalman filter; GDPM, generalized dual-porosity model; KF, Kalman filter; LANL, Los Alamos National Laboratory; MDA, Material Disposal Area; MPI, message passing interface; MVG, Mualem–van Genuchten; PFBA, pentafluorobenzoate; RTD, residence time distribution; SLS, simple least squares; SVE, soil vapor extraction; VOC, volatile organic compound.

MODELS SIMULATING flow and transport through the vadose zone require accurate estimates of the soil water retention and hydraulic conductivity function (hereafter referred to as *soil hydraulic properties*) at the application scale of interest. In the past few decades, various laboratory experiments have been developed to facilitate a rapid, reliable, and cost-effective estimation of the soil hydraulic properties. For practical considerations, most of these experiments have focused on relatively small soil cores. Unfortunately, many contributions to the hydrologic literature have demonstrated an inability of these laboratory-scale measurements on small soil cores to accurately characterize flow and transport processes at larger spatial scales. This necessitates the development of alternative methods to derive the soil hydraulic properties at the application scale of the model.

J.A. Vrugt, Center for Nonlinear Studies, Los Alamos National Lab., Los Alamos, NM 87545; P.H. Stauffer and V.V. Vesselinov, Earth and Environmental Sciences Division, Los Alamos National Lab., Los Alamos, NM 87545; Th. Wöhling, Lincoln Environmental Research, Ruakura Research Center, Hamilton, New Zealand; B.A. Robinson, Civilian Nuclear Program Office, Los Alamos National Lab., Los Alamos, NM 87545. Received 23 Apr. 2007. *Corresponding author (vrugt@lanl.gov).

Among the first to suggest the application of computer models to estimate soil hydraulic parameters were Whisler and Whatson (1968), who reported on estimating the saturated hydraulic conductivity of a draining soil by matching observed and simulated drainage. This process of iteratively adjusting the model parameters so that the model approximates, as closely and consistently as possible, the observed response of the system under study during some historical period of time is called *inverse modeling*. Parameter estimation using inverse modeling accommodates more flexible experimental conditions than typically utilized in laboratory experiments and facilitates estimating values of the hydraulic properties that pertain to the scale of interest, and thus is useful for upscaling. When adopting an inverse modeling approach, however, the soil hydraulic properties can no longer be estimated by direct inversion using closed form analytical equations but are determined using an iterative solution, thereby placing a heavy demand on computational resources.

In the field of vadose zone hydrology, inverse modeling usually involves the estimation of the soil water retention and unsaturated soil hydraulic conductivity characteristics using repeated numerical solutions of the governing Richards equation:

$$C(h)\frac{\partial h}{\partial t} = \nabla\left[\mathbf{K}(h)\nabla(h+z)\right] + S(x,y,z,t) \qquad [1]$$

thereby minimizing the difference between observed and model-predicted flow variables such as water content and fluxes. In Eq. [1], $C(h)$ is the so-called soil water capacity, representing the slope of the soil water retention curve [L^{-1}], h denotes the soil water matric head [L], t represents time [T], $\mathbf{K}$ is the unsaturated

hydraulic conductivity tensor [L T^{-1}], z denotes the gravitational head to be included for the vertical flow component only [L], and S is the volumetric sink term representing sources and or sinks of water [L^3 L^{-3} T^{-1}]. To solve Eq. [1] for the considered soil domain, appropriate initial and boundary conditions need to be specified, and time-varying sinks and sources need to be included.

We will briefly review the current state of the art of inverse modeling for estimating unsaturated flow and transport processes. We will explain how the inverse method works, discuss the historical background that led to the current perspectives on inverse modeling, and review the solution algorithms used to solve the parameter estimation problem. We will then highlight and illustrate some of our recent work at Los Alamos on self-adaptive multimethod global optimization, parallel computing, and sequential data assimilation to improve efficiency of parameter estimation in complex flow and transport applications and help assess parameter and model output prediction uncertainty.

Next, we present and discuss three different case studies that illustrate these various summarized developments. The first case study considers the calibration of a fully integrated three-dimensional vapor extraction model using observed data from a volatile organic compound in the subsurface at the Material Disposal Area near the Los Alamos National Laboratory (LANL). The second case study considers a multicriteria calibration of the Mualem–van Genuchten (MVG) parameters in the HYDRUS-1D model using observed tensiometric data from the Spydia field site in New Zealand. The last case study describes the application of a combined parameter and state estimation method to the interpretation of a multitracer experiment conducted at the Yucca Mountain field site in Nevada. This method, entitled Simultaneous Optimization and Data Assimilation (SODA) improves the treatment of input, model structural, and output error by combining the strengths of stochastic global optimization and recursive state updating using an ensemble Kalman filter.

Inverse Modeling: A Review

We provide a short description of the underlying concept of inverse modeling and discuss the progress that has been made with particular emphasis on laboratory experiments. Excellent reviews on inverse modeling of soil hydraulic properties have been presented in Hopmans and Šimůnek (1999) and Hopmans et al. (2002). We suggest reading this material to get a more complete overview and background.

Mathematical Development

A schematic overview of the inverse modeling procedure appears in Fig. 1. Consider a model Φ in which the discrete time evolution of the state vector ψ is described by

$$\psi_{t+1} = \Phi\left(\psi_t, \tilde{X}_t, \beta\right) \qquad [2]$$

where $\tilde{X}$ represents the observed forcing (e.g., boundary conditions), β is the vector of parameter values, and t denotes time. In
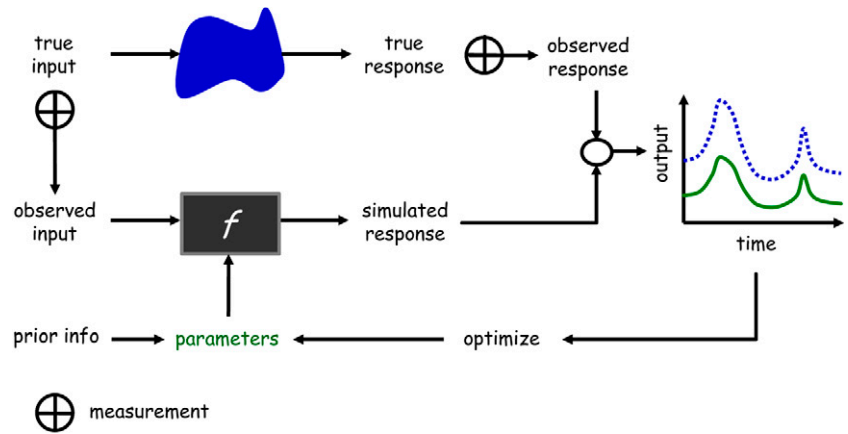

FIG. 1. Schematic overview of inverse modeling: The model parameters are iteratively adjusted so that the predictions of the model, f, (represented with the solid line) approximate as closely and consistently as possible the observed response (represented with the dotted line). The symbol $\oplus$ represents observations of the forcing terms and responses that are subject to measurement erroros and uncertainty, and therefore may be different than the true values. Similarly, f represents the model with functional response of a rectangular box to indicate that the model is, at best, only an approximation of the underlying system. The label "output" on the y axis of the plot on the right-hand side can represent any time series of data.

the current context, Φ represents Richards' equation augmented with parametric forms of the soil water retention and unsaturated hydraulic conductivity functions, whose exact shapes are defined by the values in β. Assume that realistic upper and lower bounds on each of the p model parameters, $\beta = \{\beta_1, ..., \beta_p\}$ can be specified a priori, thereby defining the feasible space of solutions:

$$\beta \in B \subseteq \Re^p \qquad [3]$$

If B is not the entire domain space, $\Re^p$, the problem is said to be constrained.

A common approach to estimating the values of β for a particular site is to take a small soil sample from the field and conduct a transient experiment under controlled conditions with prescribed initial and boundary conditions. During this experiment, one or more flow-controlled variables are measured and collected in the vector $\tilde{Y} = \{\tilde{y}_1, ..., \tilde{y}_n\}$, where n denotes the total number of observations. After this, a simulation is performed with Φ given some initial guess for β (usually based on some prior information), and the vector of model predictions $Y(\beta) = \{y_1(\beta), ..., y_n(\beta)\}$ is computed. These output predictions are directly related to the model state according to

$$y_t = \Omega\left(\psi_t\right) \qquad [4]$$

where the measurement operator $\Omega(\ )$ maps the state space into the measurement or model output space. Finally, the difference between the model-simulated output and measured data is subsequently computed and stored in the residual vector, E:

$$E(\beta) = G\left[Y(\beta)\right] - G\left[\tilde{Y}\right] = \left\{e_1(\beta), ..., e_n(\beta)\right\} \qquad [5]$$

where the function $G(\ )$ allows for various user-selected linear or nonlinear transformations of the simulated and observed data.

The aim of parameter estimation or model calibration now becomes finding those values of β such that the measure E is in

some sense forced to be as close to zero as possible. The formulation of a criterion that mathematically measures the size of $\mathbf{E}(\boldsymbol{\beta})$ is typically based on assumptions regarding the distribution of the measurement errors present in the observational data. The classical approach to estimating the parameters in Eq. [4] is to ignore input data uncertainty and to assume that the predictive model Φ is a correct representation of the underlying physical data-generating system. In line with classical statistical estimation theory, the residuals in Eq. [5] are then assumed to be mutually independent (uncorrelated) and normally distributed with a constant variance. Under these circumstances, the "best" parameter combination can be found by minimizing the following simple least square (SLS) objective function with respect to $\boldsymbol{\beta}$:

$$F_{\text{SLS}} = \sum_{i=1}^{n} e_i(\boldsymbol{\beta})^2 \qquad [6]$$

In this case, Hollenbeck and Jensen (1998) stretched the importance of model adequacy before sound statements can be made about the final parameter estimates and their uncertainty. Model adequacy is determined from

$$p_{\text{adeq}} = 1 - Q\left(\frac{F_{\text{SLS}}}{\sigma_{\text{T}}^2}, n - p\right) \qquad [7]$$

where σ_{T} denotes the error deviation of the measurements, and $Q(\ )$ is the χ^2 cumulative distribution with $(n - p)$ degrees of freedom. This adequacy test gives us a measure of how well the optimized model fits the observations relative to their measurement precision. Using the definition in Eq. [7], models are adequate if p_{adeq} is >0.5.

Historical Background

During the last few decades, a great deal of research has been devoted to exploring the applicability and suitability of the inverse approach for the identification of soil hydraulic parameters. Most of this work has focused on laboratory experiments using small soil cores with well-defined boundary conditions to fully understand the prospects and limitations of the method and facilitate easy benchmarking against hydraulic properties derived from static or steady-state methods on similar soil samples. For our discussion, we group the early work on the use of inverse methods in vadose zone hydrology into five different categories, which we will discuss here in sequential order:

1. The type of transient experiment and kind of initial and boundary conditions suited to yield a reliable characterization of the soil hydraulic properties were investigated by van Dam et al. (1992, 1994), Ciollaro and Romano (1995), Santini et al. (1995), Šimůnek and van Genuchten (1996, 1997), Inoue et al. (1998), Šimůnek et al. (1998b), Romano and Santini (1999), Durner et al. (1999), Wildenschild et al. (2001), Si and Bodhinayake (2005), and Zeleke and Si (2005). Various investigations have demonstrated the usefulness of the evaporation method, multistep outflow (MSO), ponded infiltration, tension infiltrometer, multistep soil water extraction, cone penetrometer, and falling head infiltration method for inverse estimation of the hydraulic parameters. These methods typically enable collection of a sufficient experimental range of data to at least be able to estimate some of the hydraulic parameters with good accuracy.

2. The determination of the appropriate boundary conditions and most informative kind of observational data were investigated by Zachmann et al. (1981), Kool et al. (1985), Parker et al. (1985), Kool and Parker (1988), Valiantzas and Kerkides (1990), Toorman et al. (1992), Eching and Hopmans (1993), Eching et al. (1994), Durner et al. (1999), and Vrugt et al. (2002), among others. Early work reported by Kool et al. (1985) and Kool and Parker (1988) suggested that the transient experiments should cover a wide range in water contents and preferably include tensiometer measurements within the soil sample to match the observed $\theta(h)$ data (Eching and Hopmans, 1993). Additionally, in the case of outflow experiments, van Dam et al. (1994) has shown that more reliable parameter estimates can be obtained by incrementally increasing the pneumatic pressure in several steps instead of using a single pressure increment throughout the entire experiment. In Vrugt et al. (2001c, 2002), a global sensitivity analysis (PIMLI) was presented to show that the information content for the various soil hydraulic parameters was separated sequentially with time during a MSO experiment, and that this information could be used to help improve the identification of the global minimum in the search space.

3. The selection of an appropriate model of the soil hydraulic properties was investigated by Zachmann et al. (1982), Russo (1988), and Zurmühl and Durner (1998). Numerous investigations have demonstrated that the MVG (Mualem, 1976; van Genuchten, 1980), Brooks and Corey (Brooks and Corey, 1964), Rossi and Nimmo (Rossi and Nimmo, 1994), and Kosugi (Kosugi, 1996, 1999) models appropriately describe the hydraulic properties of most soils. To further increase flexibility to fit experimental data, various researchers have proposed extensions of these models to describe multimodal pore size distributions, including multilevel spline approximations of the hydraulic properties (Bitterlich et al., 2004; Iden and Durner, 2007) and extensions to dual-porosity and dual-permeability models to simulate preferential flow (Šimůnek et al., 2003, and the many references therein).

4. The implementation of methods to quantify the uncertainty associated with the inversely estimated parameters was investigated by Kool and Parker (1988), Hollenbeck and Jensen (1998), Vrugt and Bouten (2002), and Vrugt et al. (2003a, 2004). This work has focused on the implementation and use of traditional first-order approximation methods to estimate linear parameter uncertainty intervals (Carrera and Neuman, 1986; Kool and Parker, 1988), development of computationally more expensive response surface and Markov chain Monte Carlo sampling approaches to derive exact nonlinear confidence intervals (Toorman et al., 1992; Hollenbeck and Jensen, 1998; Romano and Santini, 1999; Vrugt et al., 2001c, 2003a), pseudo-Bayesian methods (Beven et al., 2006; Zhang et al., 2006), and multicriteria approaches to interpret the Pareto solution set of two or more conflicting objectives (Vrugt and Dane, 2005; Schoups et al., 2005a,b; Mertens et al., 2006; Wöhling et al., 2008).

5. The construction and weighting of multiple sources of information in an objective function was investigated by van Dam et al. (1994), Clausnitzer and Hopmans (1995), Hollenbeck and Jensen (1998), and Vrugt and Bouten (2002) and search methods to efficiently locate the optimal parameters in rough, multimodal response surfaces (objective function mapped out in the parameter space) were developed and applied by Abbaspour et al. (1997, 2001), Pan and Wu (1998), Takeshita (1999), Vrugt et al. (2001c), Lambot et al. (2002), Vrugt and Bouten (2002), and Mertens et al. (2005, 2006). In recent years, Bayesian and pseudo-Bayesian approaches have been developed to weight multiple types of data in one objective function, and local and global optimization methods such as the sequential uncertainty fitting (SUFI), annealing-simplex, shuffled complex, and ant-colony methods have been proposed to find the optimal values of the soil hydraulic properties.

While initial applications of inverse modeling have focused primarily on the estimation of the unsaturated hydraulic

properties from small soil cores, recent progress is also reported in the description of water uptake by plant roots (Vrugt et al., 2001a,b; Hupet et al., 2002), estimation of soil thermal properties (Hopmans et al., 2001), measurement of water content using time domain reflectometry (Huisman et al., 2004; Heimovaara et al., 2004), prediction of pore geometry based on air permeability measurements (Unsal et al., 2005), simultaneous estimation of soil hydraulic and solute transport parameters (Inoue et al., 2000), optimal allocation of electrical resistivity tomography electrodes to maximize measurement quality (Furman et al., 2004), and automated water content reconstruction of zero-offset borehole ground penetrating radar data (Rucker and Ferre, 2005). Moreover, with the ever-increasing pace of computational power and the inability of small-core measurements to adequately characterize larger scale flow and transport properties, application scales of inverse modeling have significantly increased in recent years, from the soil core to the field plot and regional scale, to provide solutions to emerging environmental problems such as salinization and groundwater pollution (Neupaier et al., 2000; Vrugt et al., 2004; Schoups et al., 2005a,b).

Among the first to apply the inverse modeling approach to a field situation were Dane and Hruska (1983), who estimated soil physical parameters using transient drainage data. The application of inverse modeling to estimate soil hydraulic properties across spatial scales is very promising, yielding parameters that represent effective conceptual representations of spatially and temporally heterogeneous watershed properties at the scale of interest. This approach is particularly powerful because it overcomes many of the difficulties associated with conventional upscaling approaches. With few exceptions, however, larger scale models are typically based on complex multidimensional governing equations, requiring significant computational resources for simulation and efficient optimization algorithms for model calibration. Moreover, unlike in small-scale experiments, boundary and initial conditions at the larger spatial scale are not as well defined because direct measurement techniques are mostly not available. Furthermore, the observational data available to characterize large-scale vadose zone processes are sparse, both in space and time. This requires the use of inverse algorithms that are efficient and can derive meaningful uncertainty estimates on the model parameters and associated model predictions. Note that inverse modeling is a powerful method to derive effective values of the soil hydraulic parameters at various spatial scales and circumvents many of the problems associated with conventional upscaling methods. For an excellent review of upscaling approaches for hydraulic properties and soil water flow, see Vereecken et al. (2007).

Parameter Estimation

For linear models, simple analytical solutions exist that conveniently minimize Eq. [6] and derive the optimal value for β at relatively low computational cost. Unfortunately, for most applications in subsurface flow and transport modeling (and many other inverse problems in hydrology), the optimal value of β can no longer be estimated by direct inversion and needs to be estimated by an iterative process. During this process, the parameters are iteratively adjusted so that the F_{SLS} objective function is minimized and the model approximates, as closely and consistently as possible, the observed response of the system under study. Hence, this procedure is a critical component of the inverse modeling process, as the final optimized hydraulic properties might be susceptible to significant error if a wrong search procedure is used.

Because of the subjectivity and time-consuming nature of manual trial-and-error parameter estimation, there has been a great deal of research into the development of automatic methods for model calibration. Automatic methods seek to take advantage of the speed and power of computers, while being objective and easier to implement than manual methods. These algorithms may be classified as local search methodologies, when seeking for systematic improvement of the objective function using an iterative search starting from a single arbitrary initial point in the parameter space, or as global search methods in which multiple concurrent searches from different starting points are conducted within the parameter space.

One of the simplest local-search optimization methods, which is commonly used in the field of soil hydrology, is a Gauss–Newton type of derivative-based search (Marquardt, 1963; Zijlstra and Dane, 1996):

$$\beta^{k+1} = \beta^k - H\left(\beta^k\right)^{-1}\left[\nabla f\left(\beta^k\right)\right]^T \qquad [8]$$

where β^{k+1} is the updated parameter set and $\nabla f(\beta^k)$ and $H(\beta^k)$ denote the gradient and Hessian matrix, respectively, evaluated at $\beta = \beta^k$. From an initial guess of the parameters β^0, a sequence of parameter sets, $\{\beta^1, \beta^2, \beta^{k+1}\}$, is generated that is intended to converge to the global minimum of $\mathbf{E}(\beta)$ in the parameter space. Doherty (2004) and Clausnitzer and Hopmans (1995) presented general-purpose optimization software that implement this Gauss–Newton or Levenberg–Marquardt (Marquardt, 1963) type of search strategy.

The derivative-based search method defined in Eq. [8] will evolve toward the true optimum in the search space in situations where the objective function exhibits a convex response surface in the entire parameter domain. Unfortunately, numerous contributions to the hydrologic literature have demonstrated that the response surface seldom satisfies these conditions, but instead exhibits multiple optima in the parameter space with both small and large domains of attraction, discontinuous first derivatives, and curved multidimensional ridges. Local gradient-based search algorithms are not designed to handle these peculiarities, and therefore they often prematurely terminate their search with their final solution essentially being dependent on the starting point in the parameter space. Another emerging problem is that many of the hydraulic parameters typically demonstrate significant interaction because of an inability of the observed experimental data to properly constrain all of the calibration parameters. This further lowers the chance of finding a single unique solution with local search methodologies.

These considerations inspired researchers in many fields of study to develop robust global optimization methods that use multiple concurrent searches from different starting points to reduce the chance of getting stuck in a single basin of attraction. Global optimization methods that have been used for the estimation of the unsaturated soil hydraulic properties include the annealing–simplex method (Pan and Wu, 1998), genetic algorithms (Takeshita, 1999; Vrugt et al., 2001c), multilevel grid sampling strategies (Abbaspour et al., 2001; Lambot et al., 2002), ant-colony optimization (Abbaspour et al., 2001), and shuffled

complex methods (Vrugt and Bouten, 2002; Vrugt et al., 2003c; Mertens et al., 2005, 2006).

Recent Advances in Inverse Modeling

We highlight and illustrate here some of our recent work at Los Alamos on self-adaptive multimethod global optimization, parallel computing, and sequential data assimilation to improve efficiency of parameter estimation in complex flow and transport applications, and help assess parameter and model output prediction uncertainty. These methods are illustrated using three different case studies whose details and results are presented below.

Multimethod Global Optimization

The current generation of optimization algorithms typically implements a single operator for population evolution. Reliance on a single model of natural selection and adaptation presumes that a single method exists that can efficiently evolve a population of potential solutions through the parameter space and work well for a large range of problems. Existing theory and numerical benchmark experiments have demonstrated, however, that it is impossible to develop a single universal algorithm for population evolution that is always efficient for a diverse set of optimization problems (Wolpert and Macready, 1999). This is because the nature of the fitness landscape (objective function mapped out as function of β) often varies considerably between different optimization problems and often dynamically changes en route to the global optimal solution. It therefore seems productive to develop a search strategy that adaptively updates the way it generates offspring based on the local peculiarities of the response surface.

In light of these considerations, Vrugt and Robinson (2007a) and Vrugt et al. (2008) recently introduced a new concept of self-adaptive multimethod evolutionary search. This approach, entitled A MultiAlgorithm Genetically Adaptive Method (AMALGAM), runs a diverse set of optimization algorithms simultaneously for population evolution and adaptively favors individual algorithms that exhibit the highest reproductive success during the search. By adaptively changing preference to individual algorithms during the course of the optimization, AMALGAM has the ability to quickly adapt to the specific peculiarities and difficulties of the optimization problem at hand. Synthetic single- and multiobjective benchmark studies covering a diverse set of problem features, including multimodality, ruggedness, ill conditioning, nonseparability, interdependence (rotation), and high dimensionality, have demonstrated that AMALGAM significantly improves the efficiency of evolutionary search (Vrugt and Robinson, 2007a; Vrugt et al., 2008). An additional advantage of self-adaptive search is that the need for algorithmic parameter tuning is reduced, increasing the applicability to solving search and optimization problems in many different fields of study. An extensive algorithmic description and outline of AMALGAM, including comparison against other state-of-the-art single- and multiobjective optimization methods can be found in Vrugt and Robinson (2007a) and Vrugt et al. (2008). In our case studies, we illustrate the application of AMALGAM to inverse modeling of subsurface flow and transport properties. Before doing so, we first highlight recent advances in parallel computing to facilitate an efficient solution to the inverse problem for complex subsurface flow and transport models requiring significant computational time.

Parallel Computing Using Distributed Networks

The traditional implementation and application of many local and global optimization methods involves sequential execution of the algorithm using the computational power of a single central processing unit (CPU). Such an implementation works acceptably well for relatively simple optimization problems and those optimization problems with models that do not require much computational time to execute. For high-dimensional optimization problems involving complex spatially distributed models, such as are frequently used in the field of earth science, however, this sequential implementation needs to be revisited. Most computational time required for calibrating parameters in subsurface flow and transport models is spent running the model code and generating the desired output. Thus, there should be large computational efficiency gains from parallelizing the algorithm so that independent model simulations are run on different nodes in a distributed computer system.

Distributed computers have the potential to provide an enormous computational resource for solving complex environmental problems, and there is active research in this area to take better advantage of parallel computing resources. For example, in hydrology applications, parallel computing is being exploited to improve computational efficiency of individual, large-scale groundwater flow (Wu et al., 2002) and reactive transport (Hammond et al., 2005) models. Efforts to couple hydrologic models consisting of a network of individual submodels (groundwater, surface water, and atmospheric models) also are being designed in a way that submodels can be partitioned to different processors (Winter et al., 2004). Finally, parallel versions of model inversion and sensitivity analysis software such as PEST (Doherty, 2004) or the Shuffled Complex Evolution Metropolis (SCEM-UA) optimization algorithm (Vrugt et al., 2006b) have been developed.

The parallel implementation of AMALGAM is presented in Fig. 2. In short, the master computer runs the algorithmic part of AMALGAM and generates an offspring population from the parent population using various genetic operators. This new population is distributed across a predefined number of computational nodes. These nodes (also referred to as *slaves*) execute the simulation model and compute the objective function of the points received. After this, the master computer collects the results and follows the various algorithmic steps to generate the next generation of points. This iterative process continues until convergence has been achieved. Various case studies presented in Vrugt et al. (2006b) have demonstrated that this setup results in an almost linear increase in speed for more complex simulation models, suggesting that the communication time between master and nodes is typically small compared with the time needed to run the model.

The most important message passing interface (MPI) calls that are used to facilitate communication between the master and slave computers are: (i) MPI_Send—to send a package with parameter combinations (master) or model simulation outputs (slave); (ii) MPI_Prob—to check whether there are any incoming messages; (iii) MPI_Get_elements—to track the number of basic elements in the package; and (iv) MPI_Recv—to retrieve the information in the package. A detailed description of each of these functions appears in tutorials from the LAM Team (2006) and so will not be repeated here. An example implementation of these functions for Markov chain Monte Carlo simulation

using the MPITB toolbox developed by Fernández et al. (2004) is presented in Vrugt et al. (2006b).

Bayesian and Multiobjective Inverse Modeling

Classical parameter optimization algorithms focus on the identification of a single best parameter combination, without recourse to uncertainty estimation. This seems unjustified given the presence of input (boundary conditions), output (calibration data), and model structural error (our model is only an approximation of reality) in most of our modeling efforts. Hence the assumptions underlying the classical approach to parameter estimation need to be revisited.

One response to directly confront the problem of overconditioning is to abandon the search for a single "best" parameter combination and adopt a Bayesian viewpoint, which allows the identification of a distribution of model parameters. The Bayesian approach treats the model parameters in Eq. [6] as probabilistic variables having a joint posterior probability density function, which summarizes our belief about the parameters β in light of the observed data $\tilde{Y}$. An example of this approach is the SCEM-UA optimization algorithm for simultaneously estimating the traditional "best" parameter set and its underlying probability distribution within a single optimization run. Detailed investigations of the SCEM-UA-derived posterior mean, standard deviation, and Pearson correlation coefficients between the samples in the high probability density region of the parameter space facilitates the selection of an adequate model structure (Vrugt et al., 2003a), helps to assess how much complexity is warranted by the available calibration data (Vrugt et al., 2003b), and guides the development of optimal experimental design strategies (Vrugt et al., 2002). Applications of the SCEM-UA algorithms to inverse modeling of subsurface flow and transport properties are presented in Vrugt et al. (2003a, 2004), Vrugt and Dane (2005), and Schoups et al. (2005a,b).

Another response, highlighted here and illustrated in the second case study, is to pose the optimization problem in a multiobjective context (Gupta et al., 1998; Neuman, 1973). By simultaneously using a number of complementary criteria in the optimization procedure and analyzing the tradeoffs in the fitting of these criteria, the modeler is able to better understand the limitations of current model structures and gain insights into possible model improvements. As a commonplace illustration, consider the development of a personal investment strategy that simultaneously considers the objectives of high rate of return and low volatility. For this situation, there is no single optimal solution. Rather, there is a family of tradeoff solutions along a curve called the "Pareto-optimal front" in which improvement in one objective (say, high rate of return) comes only at the expense of a degradation of another objective (volatility). Development of robust algorithms to solve such optimization problems is a research direction that is currently attracting great interest.

The multiobjective AMALGAM method discussed above uses the Nondominated Sorted Genetic Algorithm (NSGA-II, Deb et al., 2002), Particle Swarm Optimizer
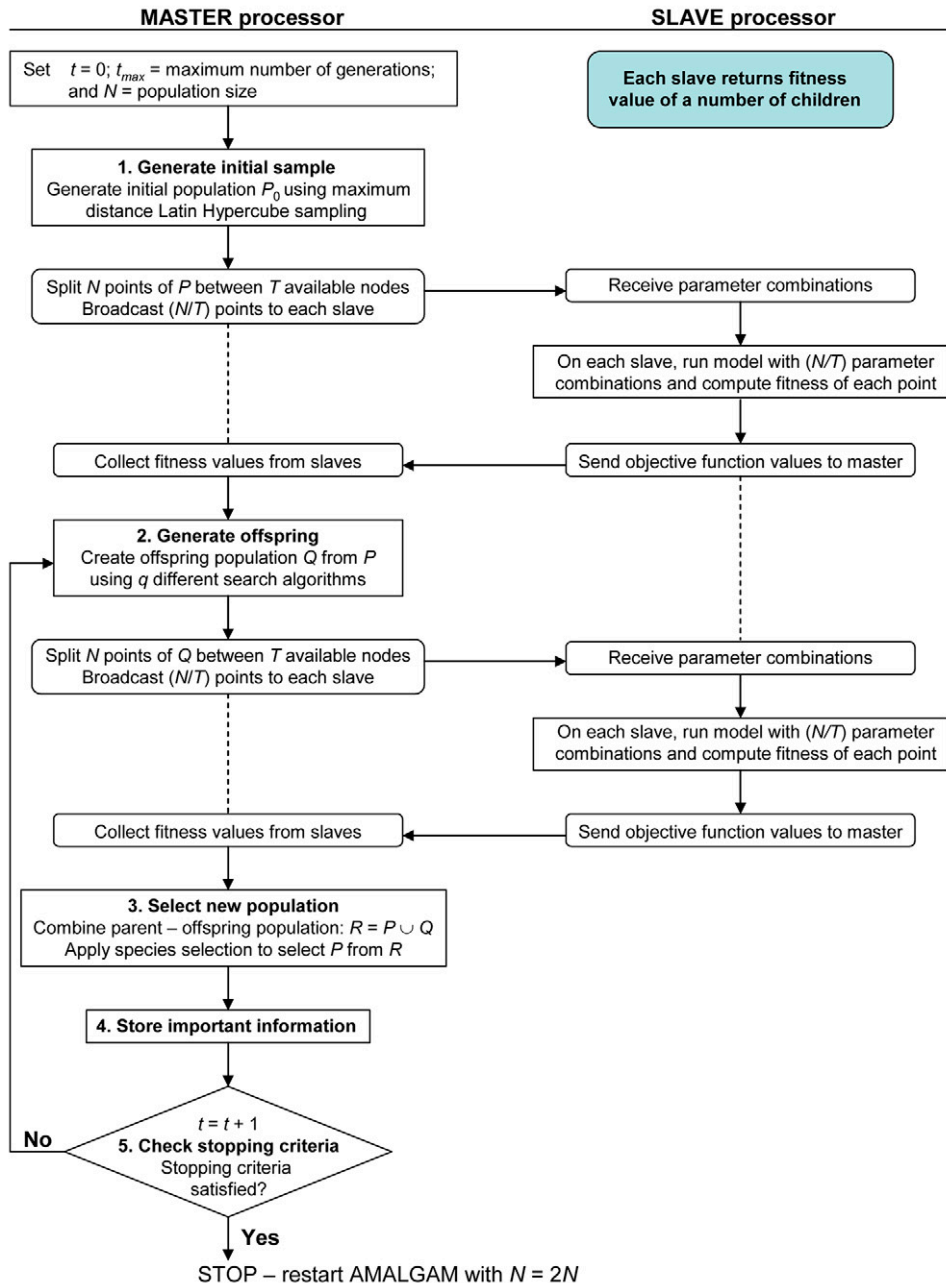

FIG. 2. Flowchart of a parallel implementation of AMALGAM. The master computer performs the various algorithmic steps in AMALGAM (on the left-hand side of the flowchart), while slave computers run the simulation model (on right-hand side of the flowchart).

(Kennedy et al., 2001), Adaptive Metropolis Search (Haario et al., 2001), and Different Evolution (Storn and Price, 1997) for population evolution and is significantly more robust and efficient than current Pareto optimization methods, with improvement approaching a factor of 10 for more complex, higher dimensional problems. We will illustrate the multiobjective version of AMALGAM in Case Study 2 below.

Improved Treatment of Uncertainty: Sequential Data Assimilation

Considerable progress has been made in the development and application of automated calibration methods for estimating the unknown parameters during inverse modeling. Nevertheless, major weaknesses of these calibration methods include their underlying treatment of the uncertainty as being primarily attributable to the model parameters, without explicit treatment of input, output, and model structural uncertainties. In subsurface flow and transport modeling, this assumption can be easily contested. Hence, uncertainties in the modeling procedure stem not only from uncertainties in the parameter estimates, but also from measurement errors associated with the system input and outputs, and from model structural errors arising from the aggregation of spatially distributed real-world processes into a relatively simple mathematical model. Not properly accounting for these errors results in error residuals that exhibit considerable variation in bias (nonstationarity), variance (heteroscedasticity), and correlation structures under different hydrologic conditions. If our goal is to derive meaningful parameter estimates that mimic the intrinsic properties of our underlying transport system, a more robust approach to the optimization problem is required. A few studies have discussed the treatment of input, state, and model structural uncertainties for subsurface flow and transport modeling (Valstar et al., 2004; McLaughlin and Townley, 1996; Katul et al., 1993; Parlange et al., 1993), but such approaches have not become common practice in current inverse modeling practices.

In the past few years, ensemble-forecasting techniques based on sequential data assimilation methods have become increasingly popular due to their potential ability to explicitly handle the various sources of uncertainty in geophysical modeling. Techniques based on the ensemble Kalman filter (EnKF, Evensen, 1994) have been suggested as having the power and flexibility required for data assimilation using nonlinear models. In particular, Vrugt et al. (2005a) recently presented the Simultaneous Optimization and Data Assimilation (SODA) method, which uses the EnKF to recursively update model states while estimating time-invariant values for the model parameters using the SCEM-UA (Vrugt et al., 2003d) optimization algorithm. A novel feature of SODA is its explicit treatment of errors due to parameter uncertainty, uncertainty in the initialization and propagation of state variables, model structural error, and output measurement errors. The development below closely follows that of Vrugt et al. (2005a), after which the application of this method to multitracer experiments is described.

To help facilitate the description of the classical Kalman filter (KF), we start by writing the model dynamics in Eq. [2] as a stochastic equation:

$$\psi_{t+1} = \Phi(\psi_t, \tilde{X}_t, \boldsymbol{\beta}) + q_t \qquad [9]$$

where q_t is a dynamical noise term representing errors in the conceptual model formulation. This stochastic forcing term flattens the probability density function of the states during the integration. We assume that the observation Eq. [4] also has a random additive error ε_t, called the measurement error:

$$\tilde{y}_t = H(\psi_t{}^*) + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_t^0) \qquad [10]$$

where σ^0 denotes the error deviation of the observations, and ψ_t^* denotes the true model states at time t. At each measurement time, when an output observation becomes available, the output forecast error z_t is computed:

$$z_t = \tilde{y}_t - H(\psi_t^{\mathrm{f}}) \qquad [11]$$

and the forecast states, ψ_t^{f}, are updated using the standard KF analysis equation:

$$\psi_t^{\mathrm{u}} = \psi_t^{\mathrm{f}} + K_t \left[\tilde{y}_t - H(\psi_t^{\mathrm{f}}) \right] \qquad [12]$$

where ψ_t^{u} is the updated or analyzed state and K_t denotes the Kalman gain. The size of the gain depends directly on the size of the measurement and model error.

The analyzed state then recursively feeds the next state propagation step in the model:

$$\psi_{t+1}^{\mathrm{f}} = \Phi(\psi_t^{\mathrm{u}}, \tilde{X}_t, \boldsymbol{\beta}) \qquad [13]$$

The virtue of the KF method is that it offers a very general framework for segregating and quantifying the effects of input, output, and model structural error in flow and transport modeling. Specifically, uncertainty in the model formulation and observational data are specified through the stochastic forcing terms q and ε, whereas errors in the input data are quantified by stochastically perturbing the elements of X.

The SODA method is an extension of traditional techniques in that it uses the EnKF to solve Eq. [9–13]. The EnKF uses a Monte Carlo method to generate an ensemble of model trajectories from which the time evolution of the probability density of the model states and related error covariances are estimated (Evensen, 1994). The EnKF avoids many of the problems associated with the traditional extended Kalman Filter (EKF) method. For example, there is no closure problem, as is introduced in the EKF by neglecting contributions from higher order statistical moments in the error covariance evolution. Moreover, the conceptual simplicity, relative ease of implementation, and computational efficiency of the EnKF make the method an attractive option for data assimilation in the meteorologic, oceanographic, and hydrologic sciences.

In summary, the EnKF propagates an ensemble of state vector trajectories in parallel, such that each trajectory represents one realization of generated model replicates. When an output measurement is available, each forecasted ensemble state vector ψ_t^{f} is updated by means of a linear updating rule in a manner analogous to the Kalman filter. A detailed description of the EnKF method, including the algorithmic details, is found in Evensen (1994) and so will not be repeated here.

Case Studies

To illustrate the various methods described above, we present three different case studies. The first case study considers the calibration of a fully coupled three-dimensional vapor extraction model using measured concentrations of volatile organic compounds in the subsurface near the LANL. The second study involves the multiobjective inverse estimation of soil hydraulic properties in the HYDRUS-1D model using observed tensiometric data from an experimental field plot in New Zealand. The last case study presents a simultaneous estimation of parameter and states in a groundwater solute mixture model using data from a multitracer experiment at Yucca Mountain, Nevada.

Case Study 1: Global Optimization of a Three-Dimensional Soil Vapor Extraction Model

Thousands of sites across the United States are contaminated with volatile organic compounds (VOCs) including trichloroethane, trichloroethene, and tetrachloroethene. These industrial solvents have high vapor pressure and low solubility and generally migrate faster in the vapor phase than in the liquid phase (Jury et al., 1990). Volatile organic compound plumes in the vadose zone can grow rapidly and are more likely to spread laterally because vapor diffusion is typically four orders of magnitude greater than liquid diffusion (Fetter, 1999). Remediation of vadose zone VOC plumes is required to protect human health, the environment, and deeper groundwater resources. The VOC plumes in deep vadose zones rely on remediation techniques that differ substantially from the techniques used to remediate plumes in the saturated zone.

One of the primary remediation techniques currently used on VOCs in the vadose zone is soil vapor extraction (SVE). Soil vapor extraction is appealing because of the relatively low costs associated with installation and operation, the effectiveness of remediation, and its widespread use at contaminated sites (Lehr, 2004). This technique uses an applied vacuum to draw pore gas toward an extraction hole. In a contaminated area, the extracted soil gas will contain some fraction of the VOC in addition to air, water vapor, and CO_2. As the VOC is removed from the subsurface pore gas, any dissolved, adsorbed, or liquid-phase VOC will tend to move into the vapor phase, thus reducing the total VOC plume (Hoeg et al., 2004). Depending on the off-gas concentrations and local regulations, the gas stream may need to be treated by methods such as C adsorption or burning with a catalytic agent.

The focus of the first case study is a VOC plume located in the vadose zone surrounding the primary liquid material disposal area (MDA L) at LANL. This area is located on Mesita del Buey, a narrow finger mesa situated near the eastern edge of the Pajarito Plateau (Broxton and Vaniman, 2005; McLin et al., 2005). During operations, liquid chemical waste was emplaced in 20-m-deep shafts on the mesa top. The VOCs from the shafts have subsequently leaked into the subsurface and created a vadose zone plume that extends laterally beyond the boundaries of the site and vertically to depths of more than 70 m below the ground surface. The vadose zone at this site is quite deep and SVE is being investigated as a possible corrective measure to remediate the plume. The initial investigation consisted of short-duration (~22 d) SVE tests on two extraction holes (indicated as SVE West and East) that were designed to collect a range of data including pressure responses and concentration measurements from both the extraction holes and surrounding monitoring holes. A detailed description of the VOC, pressure response data, and SVE numerical model appears in Stauffer et al. (2005, 2007a,b). Here, we focus on the calibration of a fully integrated three-dimensional SVE model using observed VOC extraction concentrations during the initial SVE test. Note that this case study of gas-phase transport is different than the typical applications of inverse modeling of soil hydraulic properties discussed above. Moreover, our numerical grid is made up of soil and rock mass, representing the geologic setting at LANL. Below, we shortly summarize the details relevant to our calibration.

Geologic Setting

The Pajarito plateau is located on the eastern flank of the Jemez volcanic center, and the rocks that form the plateau were created in two main ignimbrite eruptions that occurred at approximately 1.61 and 1.22 million yr (Izett and Obradovich, 1994). The plateau has been incised by canyons that drain into the Rio Grande. The regional aquifer is located approximately 300 m below the surface of MDA L, and no perched water was encountered during drilling at the site.

Figure 3 gives the general geography for MDA L and the surrounding area, and Fig. 4 shows the approximate site stratigraphy on a north–south cross-section of the mesa with several of the boreholes from Fig. 3 projected onto the plane. The most important rocks with respect to the current study are the ash-flow Units 1 and 2 of the Tshirege member of the Bandelier Tuff, Qbt1 and Qbt2, respectively. The uppermost unit at the site, Qbt2, is relative welded with ubiquitous vertical cooling joints. Many of the joints in this unit are filled with clay in the upper few meters, but are either open at depth or filled with powdered tuff (Neeper and Gilkeson, 1996). Unit Qbt1 is subdivided into upper and lower units, Qbt1-v and Qbt1-g, respectively; Qbt1-v is further subdivided into a nonwelded upper unit and a welded lower unit, Qbt-1vu and Qbt-1vc, respectively. Subunit Qbt-1vu is less welded than Qbt2 and has fewer cooling joints. Subunit Qbt-1vc is a welded tuff with many open vertical joints that provide rapid equilibration of pressure changes during pump tests (Neeper, 2002). Unit Qbt1-g is a nonwelded, glass-bearing ash flow that contains a few joints that appear to be continuations of joints formed in Qbt1-v that die out at depth. For a more complete description of the geologic framework of the Pajarito plateau, see Broxton and Vaniman (2005).

Model Formulation

Our SVE model builds on the Los Alamos porous flow simulator, FEHM, a one-, two-, and three-dimensional finite-volume heat and mass transfer code (Zyvoloski et al., 1997). This model has been used extensively for simulation of multiphase transport and has extensive capabilities to simulate subsurface flow and transport systems with complicated geometries in multiple dimensions (Stauffer et al., 1997, 2005; Stauffer and Rosenberg, 2000; Wolfsberg and Stauffer, 2004; Neeper and Stauffer, 2005; Kwicklis et al., 2006; and many others). Equations governing the conservation of phase mass, contaminant moles, and energy are solved numerically using a fully implicit Newton–Raphson scheme. A detailed description of FEHM appears in Zyvoloski et al. (1997).

FIG. 3. Areal overview of Material Disposal Area L at the Los Alamos National Laboratory.

The primary assumptions governing vapor-phase flow and transport are as follows. First, we assume that the vapor phase is composed solely of air that obeys the ideal gas law and calculate vapor-phase density (kg m^{-3}) as a function of vapor pressure and temperature. Density differences due to spatial variations in VOC concentrations hardly affected gaseous movement. For example, at the maximum plume concentration of 3000 μm^3 m^{-3}, the pore gas changes in density from about 1 to 1.01 kg m^{-3}. We use Darcy's law to calculate the advective volume flux. Vapor-phase contaminant conservation is governed by the advection–dispersion equation (Fetter, 1999) where the contaminant flux (mol m^{-2} s^{-1}) is given by

$$q_v = v_v C_v + \phi S_v D_{cv}^i \nabla C_v \qquad [14]$$

where v_v represents the advective volume flux, ϕ denotes the porosity, S_v is vapor saturation defined as air-filled porosity divided by total porosity, C_v is the molar concentration (mol m^{-3}) and the dispersion coefficient, D_{cv}, includes contributions from both dispersivity and molecular diffusion as

$$D_{cv}^i = \alpha^i v_v^i + D_v {}^* \qquad [15]$$

where the molecular diffusion coefficient in FEHM is a function of the free-air diffusion coefficient (D_{free}) and the tortuosity (τ) as

$$D_v {}^* = \tau D_{free} \qquad [16]$$

The dispersivity tensor (α^i) is directional; however, in FEHM we keep only the diagonal terms of this tensor. The superscript i implies that the equation is solved for

the principle directions. For example, in three dimensions, the volume flux at any point can be decomposed into three principle components, v_x, v_y, and v_z. An additional constraint is imposed by Henry's Law equilibrium partitioning, which requires a constant ratio between concentrations in the liquid and vapor phase as

$$C_v = H_{TCA} C_l \qquad [17]$$

where Henry's Law value and other properties relevant for trichloroethane (TCA) transport can be found in Stauffer et al. (2005, Table 2). In summary, the model is a molar-based solution to the advection–dispersion equation using Fickian transport theory. We do not account for the effects of non-Fickian diffusion; however, corrections for non-equimolar behavior are relatively small (<3%) (Fen and Abriola, 2004).
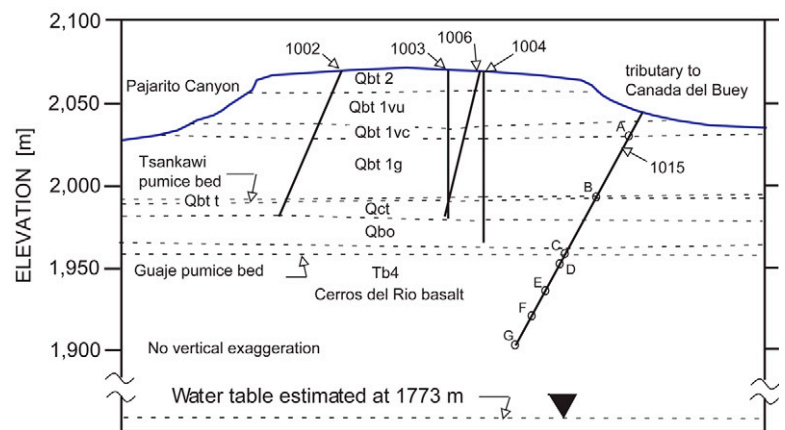

FIG. 4. Site stratigraphy with wells lying immediately to the east of Material Disposal Area L (after Neeper, 2002).

Three-Dimensional Model Domain and Computational Grid

The three-dimensional simulation domain is approximately centered on MDA L and includes the surrounding mesa and canyon environment from the land surface to the water table. Figure 5 shows a portion of the computational domain, the site boundary as a heavy black polygon, and an orthophoto showing roads and buildings. Material Disposal Area L is approximately 180 m east–west by 120 m north–south and the simulated domain extends beyond the site on all sides by a minimum of 100 m to minimize boundary effects. The computational grid is made up of >140,000 nodes and nearly 800,000 volume elements. The lateral extent is 410 m east–west by 370 m north–south. The grid extends vertically from an elevation of 1737 m above sea level (ASL) at the water table to 2074 m ASL on the northwestern corner of Mesita del Buey. The grid has a vertical resolution of 1 m in the top 90 m and stretches to a resolution of 25 m at the water table. The horizontal resolution is everywhere 10 m in both x and y directions. The grid captures the topography of the site and extends to the water table, >300 m below the surface of the mesa on which MDA L is situated. The deeper part of the grid, 90 m below the mesa top, has little impact on the simulations and is included with a vertical spacing of 10 m to address questions concerning plume impacts on the regional water table. The three-dimensional grid used here is an extension and refinement of the grid used in Stauffer et al. (2005) and images from that study will be helpful for visualizing the current domain and grid.

The computer code FEHM has a new capability that allows us to embed radial boreholes within an existing three-dimensional site-scale mesh (Pawar and Zyvoloski, 2006). This capability is used to reduce the total number of nodes required to capture the radial flow near the simulated SVE holes while also capturing the topography and stratigraphy at the site scale. Without this capability, we would have had to embed two three-dimensional extraction borehole meshes and all the necessary extra nodes to allow the borehole meshes to correctly connect to the existing three-dimensional grid while maintaining the Voronoi volume constraints that are required for computational accuracy. The wellbores used in the simulations each have an inner radius of 0.08 m and an outer shell radius of 2 m, with four nodes spanning this distance. Therefore, each well has one vertical line of nodes representing the open hole and four onion skins surrounding this. Both SVE holes have a total depth of 66 m with 67 nodes along the vertical. The nodes representing the open hole are assigned a permeability of 5×10^{-7} m^2, providing little resistance to flow in the open hole. The first onion skins in the upper 20 m of each hole are assigned a permeability of 5×10^{-19} m^2 and a diffusion coefficient (D^*) of 5×10^{-19} m^2 s^{-1} to simulate the effects of the steel casing. Nodes in the second and third onion skins are assigned the rock and tracer properties specified in a given simulation for the geologic unit in which they reside.
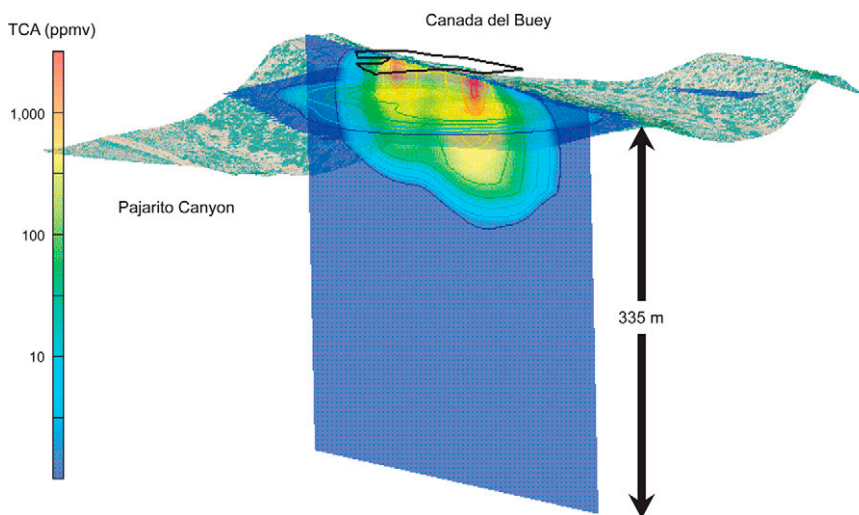

FIG. 5. Slice plane of the three-dimensional numerical grid. Colors depict the initial volatile organic compound (VOC) concentration before the soil vapor extraction pilot tests. The areal photograph is draped onto the digital elevation model of the site and shows the canyons on either side of the mesa. The Material Disposal Area L site boundary is indicated with the black polygon.

Calibration Details: Parameters and Initial and Boundary Conditions

In situ measured VOC concentrations at the extraction and surrounding bore holes were used to initialize the respective concentrations in the grid domain. A snapshot of the initial VOC concentration at the onset of the SVE pilot tests is presented in Fig. 5. The extraction flow rate during the test was assumed to be constant and calculated from an equation provided by the manufacturer of the orifice plate used to measure the pressure drop across a slight decrease in the diameter of the extraction pipe. We further simplified the modeling analysis by assuming no movement of the liquid phase. Furthermore, the atmospheric boundary pressure was held constant at 80 kPa, and the temperature was fixed to the yearly average of about 10°C (derived from local stations). The vertical side boundaries of the domain were no flow with respect to both heat and mass, and no flow of water or vapor was permitted across the bottom boundary, with its temperature being held constant at 25°C (Griggs, 1964). Finally, the porosity and saturation fraction of the different geologic units was fixed to measured values presented in Birdsell et al. (2002) and Springer (2005). Justification for all these assumptions in the context of the Pajarito Plateau was given in Stauffer et al. (2005).

Observed VOC concentrations in the vapor phase at the wellhead of the west and east extraction holes during the initial 22-d pilot experiment were used for SVE model calibration. Because of the significant spatial variability in permeability observed from straddle packer tests, the west and east pilot tests were calibrated separately. As calibration parameters, we selected the permeability of the rocks of the upper four geologic units depicted in Fig. 4. Initial sensitivity analysis demonstrated that deeper geologic units had marginal impact on the simulation results. To simulate the effects of vertical cracks, separate values for the horizontal and vertical permeability were optimized for each rock type. Moreover, we also optimized the permeability of asphalt because this material acts as an important diffusive barrier over almost the entire grid surface. Table 1 provides an overview of the calibration parameters and their initial uncertainty ranges. These ranges were based on straddle packer data

and core permeability measurements presented in Neeper (2002) and Stauffer et al. (2007a,b).

A distributed computing implementation of AMALGAM was used to optimize the SVE model parameters for the west and east pilot tests using a simple F_{SLS} objective function. We used a population size of 10 points, and hence 10 different slave computers, in combination with 120 computing hours on the LISA cluster at the SARA parallel computing center (University of Amsterdam, the Netherlands). Each of these nodes is equipped with a dual-core Intel Xeon 3.4-GHz processor with 4 GB of memory. The results of this single objective optimization are summarized in Table 1 and Fig. 6 and discussed below.

Figure 6 presents a time series plot of observed (solid circles) and simulated (solid line) VOC vapor-phase outlet concentrations at the western (Fig. 6A) and eastern (Fig. 6B) SVE wells. The corresponding optimized permeability values for the individual rock types are listed in Table 1. In general, the fit to the observed data can be considered quite good for both pilot tests after about 2 d, successfully capturing and simulating the process of matrix flow. During the first 2 d, however, the SVE model significantly underestimated observed VOC concentrations. This initial misfit was caused by flow through joints and fractures, a process widely observed throughout the Pajarito Plateau and the experimental site, but not explicitly included in our model. We represent fracture flow by allowing AMALGAM to optimize the horizontal and vertical permeability in ranges above measured matrix values. This implementation combines the effect of matrix and fracture flow, and does not allow us to explicitly simulate flow through fractures, which would be required to match the early-time data. Hence, much better predictions at the initial time steps are possible if we explicitly incorporate fracture flow in the model. One approach to doing this would be to augment the current SVE model with the generalized dual porosity model (GDPM)
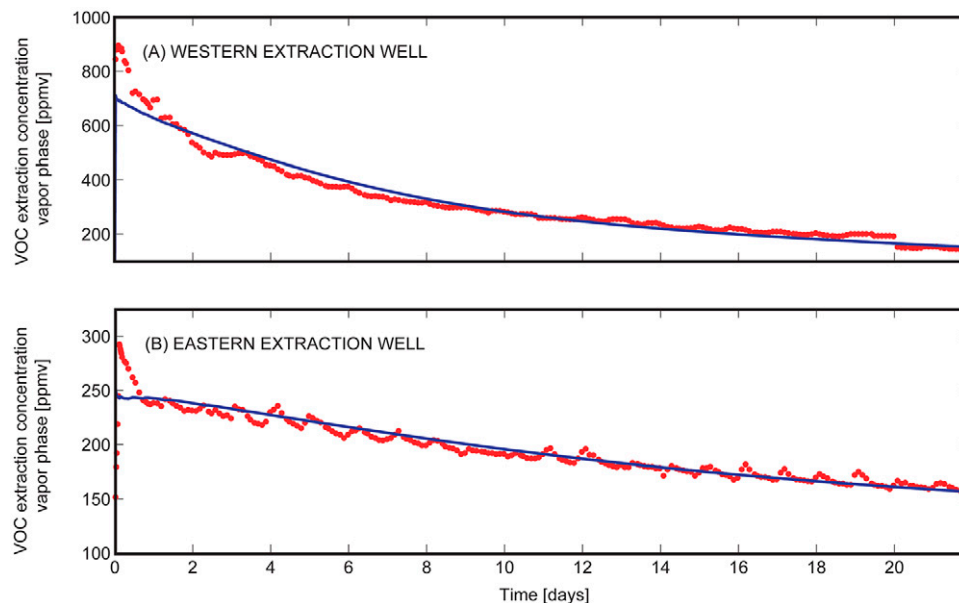

FIG. 6. Comparison of simulated (line) and observed (solid circles) volatile organic compound (VOC) outlet concentrations (ppmv is parts per million volume or $\mu m^3\ m^{-3}$) as a function of time at the (A) western and (B) eastern extraction wells.

presented in Zyvoloski et al. (2008). This method assumes one-dimensional transport into and out of the matrix using multiple closely spaced nodes connected to the primary fracture nodes. This setup can capture preferential flow and transport processes and therefore will probably simulate the high initial extraction of the VOC observed in the experimental data. An example of the implementation of GDPM is discussed below in Case Study 3. We are currently in the process of including this in our SVE model as well. Note that the diurnal variations in the VOC concentrations are due to the measurements being calibrated at a single temperature. As temperature varies during a 24-h cycle, a constant concentration in the outflowing gas would thus lead to a sinusoidal measurement time series. Nevertheless, these changes appear small, and therefore we have attempted to calibrate the SVE model to the mean signal.

The optimized permeabilities for the west and east pilot tests are in good agreement, generally within an order of magnitude difference. The optimized permeabilities appear reasonable and demonstrate the appropriate variability as observed in the field (Neeper, 2002; Stauffer et al., 2007a,b). Moreover, their optimized values are typically within the middle of the prior uncertainty ranges, only approaching the outer bounds for a few parameters. We hypothesize that more realistic values for these particular parameters will be obtained when fracture flow is explicitly included in the SVE model. It seems unjustified, however, to overcondition the calibration process to a single "best" parameter combination as done here, given the presence of systematic errors in our model predictions so apparent during the first 2 d for both pilot tests. In the next two case studies, therefore, we will provide a better treatment of parameter, model, and output uncertainty.

Case Study 2: Multiobjective Calibration of Soil Hydraulic Properties

We conducted a multiobjective calibration of the unsaturated soil hydraulic properties using observed tensiometric pressure data from three different depth intervals at the Spydia field site

TABLE 1. Case Study 1: Calibration parameters, their initial uncertainty ranges, and their final optimized values using AMALGAM for the western and eastern soil vapor extraction pilot tests.

Geologic unit[†]	Description	Prior[‡][§]	Western[§]	Eastern[§]
			m²	
Qbt 2	Permeability x, y direction	11–13	12.19	12.96
	Permeability z direction	11–13	12.34	13.00
Qbt 1vu	Horizontal permeability	11–13	11.22	11.46
	Vertical permeability	11–13	11.70	11.80
Obt 1vc	Horizontal permeability	11–13	12.99	12.05
	Vertical permeability	11–13	12.24	11.22
Obt 1g	Horizontal permeability	11–13	13.00	13.00
	Vertical permeability	11–13	12.29	12.99
Asphalt	Permeability	10–16	12.73	11.57

† Correspond to the various geological units identified and depicted in Fig. 4.
‡ Ranges based on packer and core permeability measurements presented in Stauffer et al. (2007a,b).
§ Negative $\log_{10}$ values are given.

in the Lake Taupo catchment in New Zealand. The vadose zone materials at this site consist of a loamy sand to sand material with gravel fractions up to 34%. The soil encompasses a young volcanic soil (0–1.6 m depth) on top, followed by the unwelded Taupo Ignimbrite (1.6–4.4 m) and two older Paleosols (4.5–5.8 m) having a late Pleistocene to Holocene age. Their parent material was inferred to be weathered tephra or tephric loess. A detailed physical and textural analysis of the vadose zone at the experimental site is given in Wöhling et al. (2008) and so will not be repeated here.

Altogether, 15 tensiometer probes (Type UMS T4e, UMS Umweltanalytische Mess-Systeme GmbH, Munich) were installed at five depths (0.4, 1.0, 2.6, 4.2, and 5.1 m below the soil surface) using three different replicates at each depth. The pressure head was recorded at 15-min intervals using a compact field point controller (cFP2010, National Instruments, Austin, TX). In addition, precipitation was recorded by event using a 0.2-mm bucket gauge and upscaled to hourly values for use in our calculations. Daily values of potential evapotranspiration were calculated with the Penman–Monteith equation (Allen et al., 1998) using observed meteorological data from the nearby Waihora weather station. For our inverse modeling exercise, we used the HYDRUS-1D model (Šimůnek et al., 1998a). This model simulates water flow in variably saturated porous media and uses the Galerkin finite element method based on the mass conservative iterative scheme proposed by Celia et al. (1990). An extensive description of this model appears in Šimůnek et al. (1998a). In this study, the unsaturated soil hydraulic properties are described with the MVG model (van Genuchten, 1980; Mualem, 1976):

$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r} = \left[1 + \left(\alpha_{VG}\,|h|\right)^{n_{VG}}\right]^{m_{VG}}$$

$$m_{VG} = \frac{n_{VG} - 1}{n_{VG}}$$

[18]

and

$$K\left(S_e\right) = K_s S_e^l \left[1 - \left(1 - S_e^{1/m}\right)^m\right]^2$$

[19]

where θ_s is the saturated water content [$L^3\,L^{-3}$], θ_r is the residual water content [$L^3\,L^{-3}$], α_{VG} [L^{-1}] and n_{VG} (dimensionless) are curve shape parameters, l is the pore-connectivity parameter of Mualem (1976), and K_s is the saturated hydraulic conductivity [$L\,T^{-1}$].

In the preprocessing phase, the soil domain was discretized into a rectangular grid of finite elements using a uniform nodal distance of 0.02 m and maximum depth of 4.2 m. The numerical solution with this spacing was generally found to be accurate to within 0.5% of a much smaller nodal spacing but required far less computational time for a single forward run of the model. The initial pressure head throughout the soil profile was derived from observed tensiometric data at the onset of the simulation. Moreover, all our simulations were run with an atmospheric upper boundary condition (switching between a prescribed flux and prescribed head condition depending on the pressure head at the soil–air interface) and prescribed lower boundary pressure head derived from observed tensiometric data at the bottom of the profile (4.2-m depth).

Simulations were conducted for a 282-d period between 11 Apr. 2006 and 18 Jan. 2007. In this calibration study, we focus on the tensiometer data from the unsaturated zone only, encompassing

the 0.4-, 1.0-, and 2.6-m depth intervals. Deeper intervals were influenced by lateral groundwater flow. Consistent with our knowledge of the soil profile, we used three different layers to properly characterize the unsaturated soil hydraulic properties. For each layer, we optimized the MVG parameters θ_s, α_{VG}, n_{VG}, K_s, and l. The upper and lower bounds of these parameters that define the prior uncertainty ranges are listed in Table 2. Because of poor sensitivity, the residual water content was set to zero. This is a common assumption, reducing the number of calibration parameters and increasing the computational efficiency of the calibration.

To measure the ability of the HYDRUS-1D model to simulate the tensiometric data at the 0.4-, 1.0-, and 2.6-m depths, we defined separate RMSE values for each interval. The Pareto optimal solution space for the three criteria {F_1, F_2, F_3} and 15 MVG parameters was estimated with AMALGAM using a population size of 100 points in combination with 20,000 function evaluations. The results of this three-criterion calibration are summarized in Fig. 7, 8, and 9, and discussed below. An extensive treatment of this data set and multicriteria optimization using various different optimization algorithms is presented in Wöhling et al. (2008). Here we only show and discuss some initial results.

Figure 7 presents normalized parameter plots for each of the 15 MVG parameters in the HYDRUS-1D model. The various parameters are listed along the x axis, while the y axis corresponds to the parameter values scaled according to their prior ranges defined in Table 2 to yield normalized ranges. Each line across the graph represents one Pareto solution. The lines separately indicated with symbols represent the optimal solutions for the three different criteria. The three-dimensional plot at the right-hand side depicts the Pareto solution surface in objective function space. Notice that there is significant variation in identifiability of the MVG parameters. For instance, there is considerable uncertainty associated with the saturated hydraulic conductivity for each of the three layers with Pareto solutions that span the entire prior range. On the contrary, the shape parameter n_{VG} tends to more closely cluster in the Pareto space, with values that are consistent with the soil type found at the field site. Significant tradeoff in the fitting of the various criteria is also observed in the Pareto surface of solutions on the right-hand side. If the model would be a perfect description of reality and the input and output data were observed without error, than a single combination of the 15 MVG parameters would exist that perfectly fits the observations at the various depths. In that case, the Pareto surface would consist of a single point, with values of zero for the individual objectives. The presence of various sources of error results in an inability of the HYDRUS-1D model to perfectly describe the data with a single parameter combination and therefore results in a tradeoff in the fitting of different objectives. It seems logical that there is a strong connection between Pareto uncertainty and the various error sources, although it is not particularly clear what this relationship is.

TABLE 2. Case Study 2: Lower and upper bounds for each of the HYDRUS-1D parameters used in the multicriteria optimization.

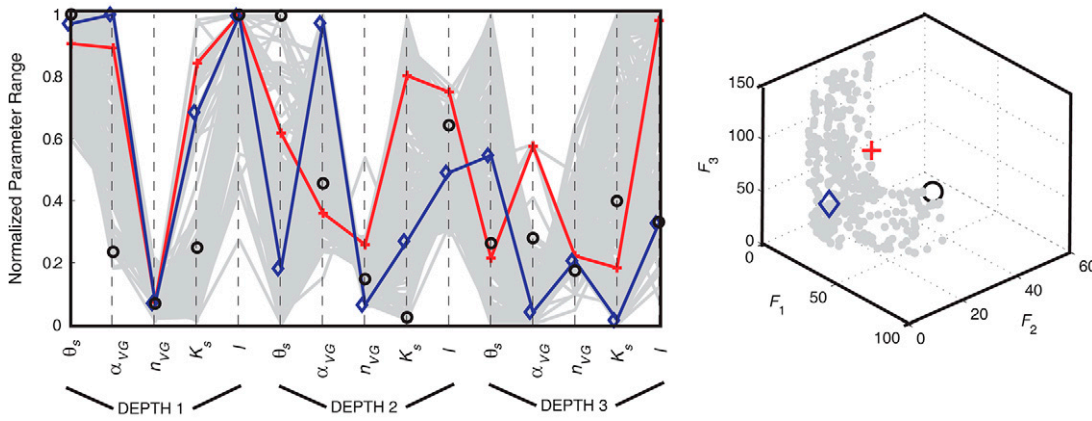| Parameter	Description	Prior
θ_s, m^3 m^{-3}	Saturated water content	0.3–0.7
α_{VG}, m^{-1}	Mualem–van Genuchten shape factor	1.0–20.0
n_{VG}	Mualem–van Genuchten shape factor	1.1–9.0
K_s, m s^{-1}	Saturated hydraulic conductivity	10^{-7}–10^{-3}
l	Pore-connectivity parameter	0.1–1.0

FIG. 7. Normalized parameter ranges for each of the Mualem–van Genuchten parameters in the HYDRUS-1D model using the three-criteria {$F_{0.4}$, $F_{1.0}$, $F_{2.6}$} calibration with AMALGAM. Each line going from left to right across the plot denotes a single Pareto solution. The squared plot at the right-hand side represents a three-dimensional projection of the objective space of the Pareto solution set. The single-objective solutions for $F_{0.4}$, $F_{1.0}$, and $F_{2.6}$ are indicated in red, blue, and black, respectively.

dashed line. The Pareto prediction uncertainty ranges generally capture the observations, but can be considered quite large, especially during prolonged dry conditions. Perhaps smaller uncertainty bounds would be achievable if we used prior information in our model calibration for at least some of the parameters, such as the saturated water content (see, for instance, Mertens et al., 2004). Note that the fit of the best parameter combination can be considered quite good, with pressure head predictions that nicely fluctuate around the measured tensiometric values. A systematic delay in modeled response is found, however, after most of the rain events. This suggests preferential flow, a mechanism that was not explicitly incorporated in our model formulation; however, HYDRUS-1D hardly improved our results by mimicking this process through the use of a dual-porosity model.

Case Study 3: Parameter and State Estimation in Subsurface Media

Interwell tracer experiments provide the best possible method at hand to study the complex flow and transport properties of a subsurface system and to measure field-scale flow and transport parameters of aquifers and reservoirs. Tracer tests bridge theory and practice by providing access to the transport characteristics of a heterogeneous porous medium. Interpretation of field tracer experiments is notoriously difficult, however, due to a

To illustrate the relative efficiency of AMALGAM, Fig. 8 presents a comparative efficiency analysis of different multiobjective optimization algorithms. This graph depicts the evolution of the sum of the best individual found objective function values for the three criteria as a function of the number of HYDRUS-1D model evaluations using the NSGA-II (Deb et al., 2002), MOSCEM-UA (Vrugt et al., 2003c), and AMALGAM (Vrugt and Robinson, 2007a; Wöhling et al., 2008) optimization algorithms. This graph is reproduced from results presented in Wöhling et al. (2008) and more details regarding this comparison can be found there. In general, results demonstrate that the AMALGAM method finds significantly better solutions than the other two nonlinear global optimization algorithms. Specifically, AMALGAM finds an overall objective function value after about 3000 runs that is better than the other two algorithms after 20,000 HYDRUS-1D model evaluations. It is probable that the performance of the other optimization algorithms would have been more similar in terms of parameter estimates and final objective function values if they would have been used in conjunction with a larger population size; however, running them with a larger population size would have significantly deteriorated their efficiency. Moreover, in this study we used commonly used values of the population. Despite these findings, the strength of AMALGAM is its good and reliable performance with a relatively small population, inspiring confidence in the efficiency of this multimethod search algorithm.

The hydraulic head predictions at the 0.4-m depth associated with the Pareto solution set is illustrated in Fig. 9 for a portion of the 228-d calibration period. Model predictions are indicated with solid lines, while the solid circles denote observations. The overall "best" prediction, having an average RMSE of about 0.14 m for the three different depths, is indicated separately with a
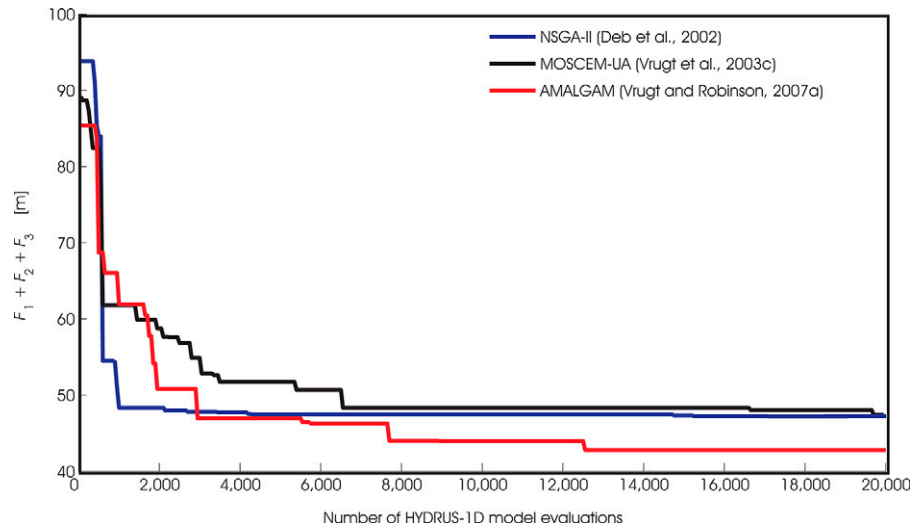

FIG. 8. Evolution of the best overall RMSE value for the three different depths as a function of the number of HYDRUS-1D model evaluations with the NSGA-II (Deb et al. 2002), MOSCEM-UA (Vrugt et al., 2003c), and AMALGAM multiobjective optimization algorithms. An extensive comparison of these algorithms appears in Wöhling et al. (2008).
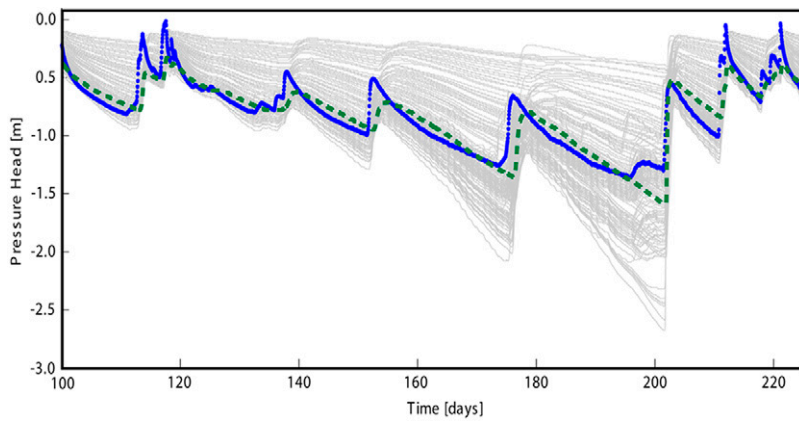
FIG. 9. Time series plots of observed and simulated pressure heads with the HYDRUS-1D model at three different depths using a representative 125-d portion of the calibration period. The solid circles denote observed data and the gray lines represent the Pareto solutions.

lack of detailed information on subsurface flow paths and mixing between pathways, incomplete knowledge about subsurface heterogeneity, and problems with parameter estimation.

Multitracer experiments are an extension of the tracer technique that in principle dramatically increases the information content relative to that of a single tracer breakthrough curve. For example, multiple tracers with different diffusion coefficients are useful to distinguish between the effects of diffusion and hydrodynamic dispersion for transport in fractured rock. Similarly, a comparison of the breakthrough curves of sorbing and nonsorbing tracers provides a direct estimate of the sorption properties along the flow path between the injection and production well. Thus, multitracer experiments help to reduce conceptual model uncertainty, providing far less ambiguous information on transport processes.

Here, we apply the SODA method to the interpretation of a multitracer experiment conducted in fractured volcanic tuffs at the C wells, near Yucca Mountain, Nevada. Cross-hole tracer tests were conducted between injection well C3 and production well C2. The purpose of this and other similar tests at this site was to establish the validity of transport models and laboratory-determined sorption parameters when applied to field-scale transport in the saturated zone beneath the proposed Yucca Mountain nuclear waste repository. A detailed description of the tracer test methods and initial interpretation was given by Reimus et al. (2003); only the aspects relevant to our modeling are repeated below. Our application is an extension of the recently published work by Vrugt et al. (2005b), which discussed the parameter estimation of a single tracer within a recursive data assimilation framework. Our current application extends this previous work by considering the measured breakthrough curves of different tracers simultaneously to estimate model parameters and states.

MultiTracer Breakthrough Curve Measurements

Breakthrough curves under forced-gradient conditions were measured for two nonsorbing solute tracers with different diffusion coefficients (Br⁻ and pentafluorobenzoate [PFBA]) and one weakly sorbing tracer (Li⁺). Tracers were mixed in 12 m^3 of water and injected simultaneously. Figure 10 shows the measured normalized concentrations of the three solute tracers at the production well as a function of time. By scaling the breakthrough curves to the relative mass of each tracer injected, differences between the curves can be attributed to differences in transport properties of the tracers. The attenuation of Br⁻ relative to PFBA is evidence of diffusion from fractures into the rock matrix, and the more attenuated Li⁺ response is attributable to sorption (Reimus et al., 2003). Additionally, the test featured three flow interruptions (one unintentional and two intentional) during the tailing portion of the experiment. The increase in tracer concentration on resumption of flow is another signature of dual-porosity behavior, with fracture flow and diffusion into the rock matrix (Reimus et al., 2003).

Conceptual Transport Model

Incomplete knowledge of subsurface heterogeneity and lack of detailed information of subsurface flow paths are complicating uncertainties that place practical limits on the complexities of flow and transport models. Thus, tracer tests are often modeled with idealized one-dimensional conceptual representations based on single-pathway models rather than full three-dimensional numerical models of subsurface flow and transport. Unfortunately, field tracer tests are often affected by complex flow through multiple paths and can exhibit long tails, skewness, and multiple peaks. If the underlying conceptual model for flow, mixing, and dispersion is too simple, transport parameter estimates will be dominated by structural deficiencies in the mixing part of the transport model and the advantage of using multiple tracers will be negated.

Many modeling approaches have been developed to deal with these complexities (e.g., Cvetkovic and Shapiro, 1990; Destouni and Cvetkovic, 1991; Simmons et al., 1995; Ginn, 2001). In our case, the goal of multitracer experimentation was to identify transport processes such as diffusion and sorption, and to estimate field-scale parameters, rather than to describe the detailed pathways in between a given well pair. Such details are not really important to the goal but can stand in the way of obtaining useful results by
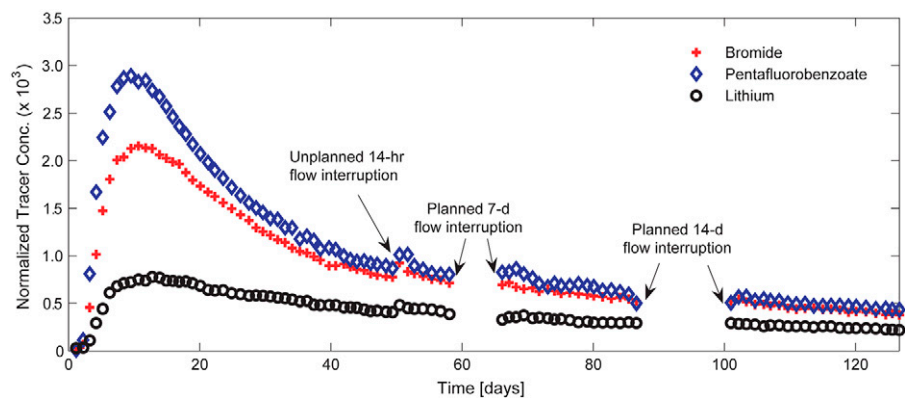

FIG. 10. Measured normalized concentrations of Br⁻, pentafluorobenzoate, and Li⁺ at the production well. The interwell tracer experiment was conducted in the period between September 1998 and January 1999 at the C-wells near Yucca Mountain, Nevada.

introducing complexity in the flow model that is not critical to the task of evaluating diffusive and sorptive processes. The philosophy of our approach is to capture the site-specific details that will inevitably result in a diverse set of advective velocities with a simple yet flexible model so that focus can be on the underlying transport processes other than advection. While other approaches are possible, we have found this simplification to lead to fruitful results.

Similar to our previous work (Vrugt et al., 2005b), we apply the theory of micromixing to the interpretation of interwell tracer tests. The model has its origins in the field of chemical reaction engineering (Zwietering, 1959), and similar approaches have been adopted in the field of groundwater (e.g., Rainwater et al., 1987; Dagan and Cvetkovic, 1996). Instead of requiring the collection of complex, in situ information on the three-dimensional subsurface flow pathways, the method enforces a particular residence time distribution (RTD). Robinson and Viswanathan (2003) showed how to build this advection flow path with side exits and assign the locations and flow rates of the side exits to reproduce an arbitrary RTD. A schematic overview of the model is given in Fig. 11.

In an extension of our previous work, the GDPM briefly discussed in Case Study 1 was applied in the present study to represent diffusion and sorption from the fractured network into the rock matrix. This method assumes one-dimensional transport into and out of a matrix connected to each primary porosity node (representing the fracture domain) in the mixing model. In contrast to other dual-porosity models, the GDPM method allows multiple, closely spaced matrix nodes to be used to capture sharp concentration gradients immediately adjacent to the primary (fracture) flow medium. The model domain of primary porosity and one-dimensional matrix nodes is represented in integrated finite difference form, and the computer code FEHM (Zyvoloski et al., 1997, 2008) is used to perform the transport calculations. The present study represents, to our knowledge, the first attempt to couple a residence-time-based conceptual model for advection to a matrix diffusion mass exchange model.

In this study, we chose the probability density function of stable distributions to represent the RTD. Stable distributions are a rich class of probability laws that can be used to represent the skewness and tailing behavior in the probability density function. The characteristic function of a stably-distributed random variable P is defined as

$$E \exp(iuP) =$$

$$\exp\left[-\gamma^{\alpha_{ST}} |u|^{\alpha_{ST}} \left(1 - i\beta_{ST} \tan\frac{\pi\alpha_{ST}}{2} \text{sign} u\right) + i\delta\delta\right] \quad [20]$$

when $\alpha_{ST} \neq 1$ and where α_{ST} is an index of stability ($\alpha_{ST} \in (0,2]$), β_{ST} is a skewness parameter ($\beta_{ST} \in [-1,1]$), γ is a scale parameter ($\gamma > 0$), and δ is a location parameter ($\delta \in \Re$). Initial sensitivity analysis and optimization runs demonstrated that the implementation of Eq. [20] would be most productive if scaling parameters in the x and y dimensions (s_x and s_y) are introduced, while fixing the values of β_{ST}, γ, and δ to 1, 60, and 60, respectively. To simplify our problem, we assumed $\beta_{ST} = 1$, forcing the stable density function to be skewed to the left, which is consistent with the data.
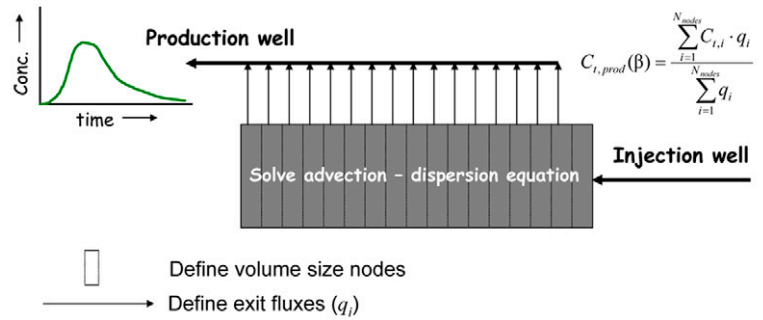

FIG. 11. Illustration of the conceptual transport model used for characterization of the subsurface. The arrows denote the side exits to reproduce a given residence time distribution; $C_{t,prod}$ denotes the simulated concentration at the production well, and N_{nodes} is the total number of nodes in the model.

After definition of the RTD, the sorbing and nonsorbing tracers are modeled in the mixing model by including parameters for matrix diffusion for all tracers and sorption for Li$^+$. It was assumed that all the tracers were transported along the same set of three-dimensional subsurface flow pathways. Consistent with laboratory sorption studies (Anghel et al., 2002), Li$^+$ sorption was modeled using a Freundlich sorption isotherm:

$$S_{t,i} = K_f C_{t,i}^{n_{FR}} \quad [21]$$

where S and C are the sorbed and aqueous concentrations at node i, respectively, and K_f and n_{FR} are sorption parameters that need to be estimated by modeling the Li$^+$ breakthrough curve. The diffusion and sorption parameters were assumed to be constant along the flow paths, an assumption that is tantamount to assuming uniform transport properties in the formation.

Table 3 summarizes the calibration parameters and their prior uncertainty ranges. Note that the parameters α_{ST}, s_x, and s_y will affect the fitting of each of the three different tracer curves as it defines the RTD, whereas the additional diffusion and sorption parameters are tracer specific. This approach is consistent with the fact that all tracers undergo the same advective and mixing processes, while each tracer has unique sorption and diffusion behavior.

State Estimation

Similarly to our previous work (Vrugt et al., 2005b), the mismatch between the measured and simulated tracer concentrations at the production well (left-hand side in Fig. 11) were used to recursively update the FEHM simulated concentrations of the various tracers. Computational details on how to do this are found in Vrugt et al. (2005b). Note that in this study, we did not impose any mass constraints on the proposed sequential state updates, possibly resulting in systematic tracer removal or addition. In principle, we could add this constraint to the KF, but a posteriori analysis has demonstrated that this was not urgent, as the total mass balance was within 5%. Based on recommendations in the classic treatise of the EnKF in Evensen (1994), the time evolution of the model error was modeled as a first-order autoregressive process.

TABLE 3. Case Study 3: Los Alamos porous flow simulator FEHM parameters subject to calibration, with their initial uncertainty and final optimized multicriteria ranges using the AMALGAM algorithm.

Parameter	Description	Prior	Multicriteria
α_{ST}	Index of stability	0.40–0.60	0.49–0.53
s_x, d	Scaling x axis residence time distribution	0.0–2.00	0.41–0.99
s_y, mg kg^{-1}	Scaling y axis residence time distribution	0.10–2000.0	134.5–1989.3
D_{Br}, m^2 s^{-1}	Matrix diffusion coefficient	10^{-15}–10^{-9}	$10^{-11.8}$–$10^{-11.0}$
D_{PFBA}, m^2 s^{-1}	Matrix diffusion coefficient	$(0.50$–$1.00)D_{Br}$†	0.34–0.50
D_{Li}, m^2 s^{-1}	Matrix diffusion coefficient	$(0.50$–$1.00)D_{Br}$†	0.52–0.97
K_f, (g kg^{-1})$^{1-n}$ dm^3 kg^{-1}	Sorption parameter	10^{-2}–10.0	4.45–7.98
n_{FR}	Sorption parameter	0.00–1.00	0.04–0.71

† Diffusion coefficients are constrained based on experimental findings in Reimus et al. (2003).

Parameter Estimation

When interpreting a multitracer experiment, the set of measurements available for parameter estimation is no longer a single outlet concentration, but consists of different transient mass outputs, each representing the measured breakthrough curve of one of the injected tracers. Considered in total, these individual tracer data sets probably contain conflicting information about the underlying transport system, so that a particular set of parameters might result in an excellent fit to one of the tracers while exhibiting poor predictive capabilities for another tracer. In principle, these individual breakthrough curves could be weighted in a single aggregate scalar, and a stochastic global optimization algorithm could be used to find the underlying posterior distribution of the parameters. In the absence of a compelling framework for the assignment of the weights, however, the parameter estimation problem is inherently multiobjective, and any attempt to convert it into a single-objective problem is associated with considerable subjectivity.

Here we implement the multiobjective model calibration approach considered in the previous case study. We use three RMSE objective functions, F_{Br}, F_{PFBA}, and F_{Li}, to trade off the fitting of the various tracer curves. Each of these objective functions is computed using the EnKF-derived time series of forecast

errors, z_t in Eq. [11], for each of the respective breakthrough curves:

$$F_{Br} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(z_t^{Br}\right)^2}$$

$$F_{PFBA} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(z_t^{PFBA}\right)^2} \quad [22]$$

$$F_{Li} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(z_t^{Li}\right)^2}$$

The Pareto optimal solution space for the three criteria specified in Eq. [22] was estimated using a population size of 100 points, in combination with 10,000 trials with a distributed computing implementation of the AMALGAM algorithm. The calculations reported here were implemented using 50 Pentium IV 3.40 GHz processors of the LISA cluster at the SARA parallel computing center (University of Amsterdam, the Netherlands). The CPU time required for joint Pareto optimization and ensemble state estimation of the conceptual mixing model was approximately 22 h. The results of the three-criteria {F_{Br} F_{PFBA}, F_{Li}} optimization are summarized in Table 3 and Fig. 12, 13, and 14 and discussed below.

Figure 12 shows a comparison of measured and predicted breakthrough curves of Br$^-$ (Fig. 12A), PFBA (Fig. 12B), and Li$^+$ (Fig. 12C) using the multiobjective SODA method. The dark gray region represents the uncertainty in the model predictions associated with the Pareto uncertainty of the parameters. Additional uncertainty due to conceptual model errors is indicated by the light gray region. Note that the simulated breakthrough curves representing both parameter and conceptual model uncertainty are reasonable and bracket the measured concentrations at the production well for each of the three tracers. Adopting a strategy of recursive state estimation using the measured breakthrough curves of the various tracers facilitates the matching of the observations, and reduces the bias in the model predictions. Perhaps more importantly, the EnKF-computed time series of temporal and spatial variations in state updates along the flow path contains useful information to improve
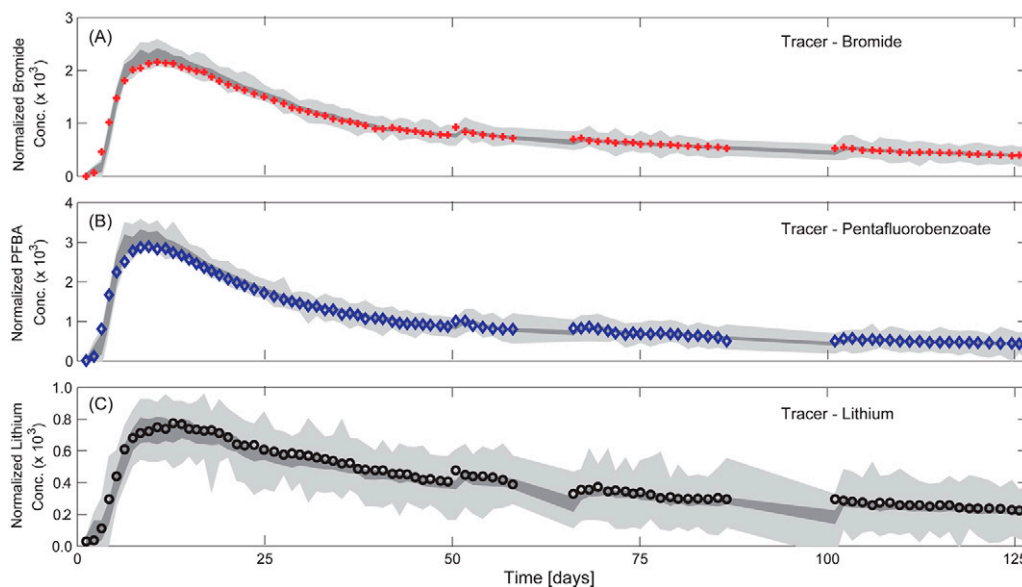

FIG. 12. Normalized breakthrough curve model predictions for (A) Br$^-$, (B) pentafluorobenzoate, and (C) Li$^+$ using the multiobjective SODA method. In each case, the symbols denote the measured data, the dark gray region indicates the prediction uncertainty bounds associated with the Pareto uncertainty in the parameters, and the light gray region represents conceptual model uncertainty.

the mixing model. In addition, the output prediction uncertainty ranges associated with the Pareto uncertainty of the parameters is generally small compared with the model conceptual error. This indicates that most of the parameters in the conceptual mixing model are reasonably well determined by optimization against multiple tracer data and that further improvement in the model would require refinement of the conceptual model. We will further elaborate on this below.

A related observation is that the output prediction uncertainty ranges are much larger for the sorbing tracer Li^+ than for the nonsorbing tracers Br^- and PFBA. This finding demonstrates that the conceptual mixing model as currently formulated is appropriate for predicting the subsurface transport processes of nonsorbing tracers, and that most of the error in the model arises from incorrect specification of the sorption processes along the flow paths. Note that the characteristic jumping behavior of the uncertainty bounds in Fig. 12C is caused by an insufficient ensemble size.

It is instructive to contrast the results obtained using the multiobjective SODA approach with a classical multiobjective AMALGAM optimization without state-variable updating. Figure 13 presents the autocorrelation functions of the computed residuals for the SODA method (Fig. 13A, 13C, and 13E) and the classical approach (Fig. 13B, 13D, and 13F). Autocorrelation is a measure of the serial dependency of the error residuals between model predictions and corresponding data. Ideally, this measure should be centered around zero at all lags, implying a lack of bias (systematic errors). Notice that there is significant autocorrelation between the error residuals for the classic multiobjective optimization with AMALGAM without state updating. Although the classical method captures the general trends in the measured tracer data (not shown), the residuals exhibit considerable variation in bias, variance, and correlation structure. Without state updating, the single-pathway conceptual mixing model is unable to track precisely the time series of measured breakthrough data for each of the individual tracers. This results in significant serial correlation between the error residuals, the effects of which clouds the interpretation of the tests. By comparison, there is considerably less autocorrelation between the residuals when performing a multiobjective SODA calibration, suggesting that recursive state adjustments remove a large part of the bias in the model predictions. When properly acknowledging model conceptual uncertainty by state variable updating, the error residuals become serially independent and exhibit normality. This approach should yield more reliable parameter estimates and uncertainty bounds.
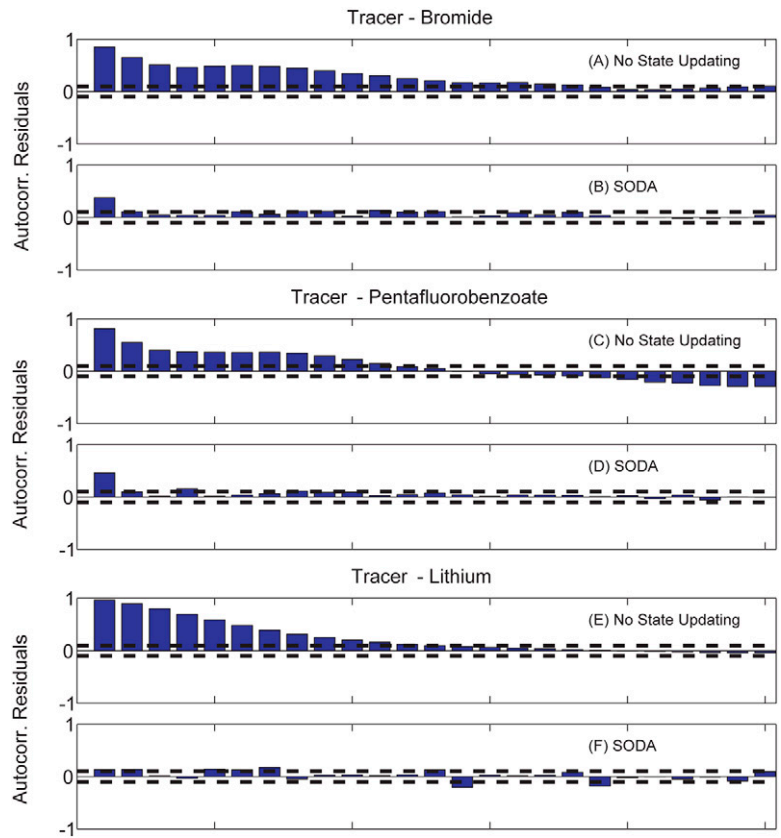

FIG. 13. Autocorrelation functions of the time series of error residuals for the Br^-, pentafluorobenzoate, and Li breakthrough curves using (A, C, and E) a classical parameter estimation algorithm without state variable updating, and (B, D, and F) the multiobjective SODA framework with state updating. The solid lines denote the theoretical upper and lower 99% significance intervals of a time series of Gaussian error residuals.

Finally, Fig. 14 presents normalized parameter plots of the results for the three-criteria $\{F_{Br}, F_{PFBA}, F_{Li}\}$ SODA calibration with the AMALGAM algorithm. Each line going from left to right across the normalized parameter plots correspond to a different parameter combination. The gray lines represent members of the Pareto set, whereas the red, blue, and black symbols refer to the optimal solutions for F_{Br}, F_{PFBA}, and F_{Li}, respectively. The eight model parameters are listed along the x axis, and the
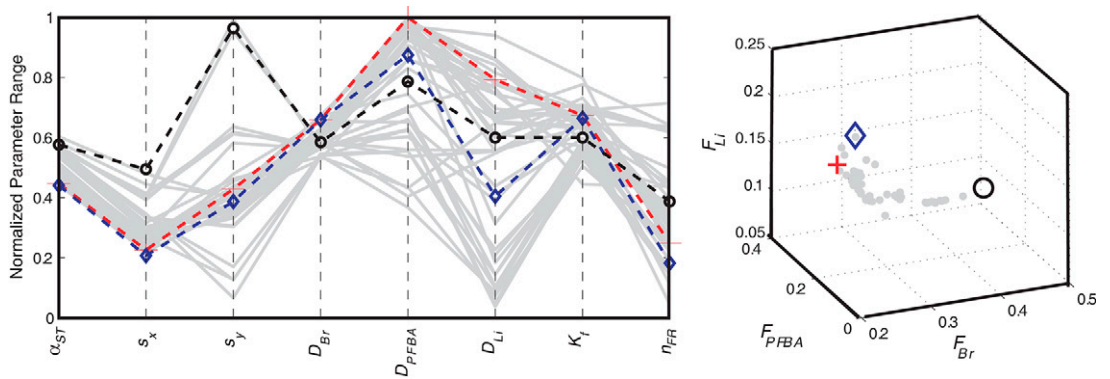

FIG. 14. Normalized parameter plots for each of the model parameters using the three-criteria $\{F_{Br}, F_{PFBA}, F_{Li}\}$ SODA optimization. Each line across the graph denotes a single parameter set: shaded is the Pareto solution set; the dashed lines in red, blue, and black are the optimal solutions for F_{Br}, F_{PFBA}, and F_{Li}, respectively. The plot at the right-hand side represents a three-dimensional projection of the objective space of the Pareto set of solutions. The optimal solutions of the three objectives are indicated with the same colors.

y axis corresponds to the parameter values, normalized by their initial ranges, as defined in Table 3. Additionally, Table 3 lists the Pareto uncertainty intervals of the parameters estimated with the MOSCEM-UA algorithm. The plot at the right-hand side in Fig. 14 depicts a three-dimensional projection of the tradeoff surface of the three objective functions. The Pareto Rank 1 solutions are indicated with gray points, whereas the optimal solutions corresponding to the three objectives are denoted with the prior defined symbols.

The results presented in Fig. 14 demonstrate that most of the mixing-model parameters are well constrained by optimization against multiple tracer data, in that most of the parameters the Pareto uncertainty intervals occupy a small region interior to the prior-defined, physically plausible uncertainty ranges. Most significantly, the diffusion coefficients are in the same general range as laboratory-determined and field-estimated values (Reimus et al., 2003), despite the fact that a very large prior range was chosen. This suggests that there is sufficient information in the time series of measured breakthrough data to recursively estimate the model states and simultaneously estimate the RTD and tracer-specific diffusion parameters. This result lends strong support to the matrix diffusion model. Also note that the stable parameter α_{ST} is estimated to be in vicinity of 0.5. For the special case when α_{ST} = 0.5 and β_{ST} = 1 (fixed a priori), the stable distribution represents the Lévy distribution.

It is interesting to note that for some of the model parameters, the Pareto set is discontinuous, with different well-separated clusters of solutions in the parameter space. This is most obvious for the scaling parameters s_x and s_y, and highlights that different combinations of these parameters result in a similar RTD curve. This conclusion was expected; no particular physical significance can be attached to the parameters of the stable distribution. Rather, the RTD-based approach is designed to allow for flexibility in establishing a nondiffusive RTD to represent all of the site-specific, unknown details associated with the advective flow paths between the wells. The approach provides a sound foundation for identifying the diffusion and sorption processes and parameters in the absence of such detailed flow path information.

The three-dimensional projection of the objective function space of the Pareto set of solutions on the right-hand side of Fig. 14 illustrates that there is significant tradeoff in the fitting of the various breakthrough curves, suggesting that for this model structure, it is not possible to find a single set of parameter values that simultaneously provides an excellent fit to the measured breakthrough curves of each of the three tracers. The most significant tradeoffs occur between the fitting of the sorbing and nonsorbing tracers. This conjecture is consistent with our previous conclusion that most of the error in the conceptual mixing model arises from an insufficiently complex specification of the sorption processes along the flow path. Note that the optimization seeks sorption isotherm parameters that are significantly more nonlinear (lower values of n_{FR}) than would be expected from batch sorption experiments (Anghel et al., 2002). Heterogeneity in sorption parameters within the fractured tuff formation is a possible reason for this result. Concentrations are highest within the short-residence-time paths, and the inverse model is curtailing the degree of sorption in this portion of the mixing model with the only parameter at its disposal, the exponent n_{FR}. A better simultaneous fit to the early and late portions of the Li^+ breakthrough curve would probably be obtained with a conceptual model of sorption heterogeneity, whereby short-residence-time paths have lower sorption coefficients than the longer paths. This concept will be explored in future work.

Summary and Conclusions

The application and use of inverse modeling for estimating subsurface flow and transport parameters has found widespread implementation and use in the last few decades. This has resulted in much experience with the method on small soil cores under controlled boundary conditions in the laboratory. With the ever-increasing pace of computational power and the inability of small-core measurements to adequately characterize larger scale flow and transport properties, however, typical application scales of inverse modeling have significantly increased in recent years from the soil core to the field plot and regional scale. Inverse modeling is particularly appealing to derive effective flow and transport parameters at larger spatial scales, because it overcomes many of the difficulties associated with conventional upscaling methods. Unfortunately, larger spatial scales typically involve complex simulation models, which are based on multidimensional governing equations, requiring significant computational resources for simulation and efficient optimization algorithms for model calibration. Moreover, boundary and initial conditions at larger spatial scales are not as well defined because direct measurement techniques are mostly not available. Furthermore, the observational data available to characterize large-scale vadose zone processes are sparse, both in space and time. These measurement limitations require significant additional methodological developments to appropriately address and communicate uncertainty.

We have reviewed the historical development of the current perspectives on inverse modeling in unsaturated flow and transport modeling, and have highlighted some of our recent work on this topic at LANL. In particular, we have discussed a general-purpose multialgorithm evolutionary search approach, called AMALGAM, that can efficiently handle a high number of parameters in complex, distributed flow and transport models and find preferred parameter solutions quickly in multimodal and noisy response surfaces. In addition, we have discussed the implementation of AMALGAM on a distributed computer network using a standard MPI toolbox in Octave (open-source software). Finally, we have discussed a combined optimization and Monte Carlo based sequential nonlinear filtering method to improve the treatment of uncertainty and provide meaningful prediction uncertainty bounds on our model simulations. All these approaches have been illustrated with three case studies addressing the transport of water, vapor, and contaminants through the vadose zone.

Finally, we would like to emphasize that considerable progress in the application of inverse modeling to complex simulation models has also been reported in adjacent fields such as groundwater and surface hydrology. Methods and approaches developed there might be of great use and inspiration to solve inverse modeling problems in vadose zone hydrology. Particularly appealing are approaches that attempt to reduce the dimensionality of the optimization problem using concepts of sensitivity and principle component analysis (Tonkin and Doherty, 2005) and representerbased calibration methods (Valstar et al., 2004). Among emerging methods for uncertainty estimation not included here are Bayesian

model averaging (Hoeting et al., 1999; Raftery et al., 2005; Vrugt et al., 2006a; Vrugt and Robinson, 2007b) and its maximum likelihood variant (Neuman, 2003; Ye et al., 2004, 2005). These methods seem particularly appealing because they are relatively easy to implement, accounting jointly for conceptual model and parameter uncertainty by considering a set of alternative models that may be very different from each other while being based on a common set of data.

References

Abbaspour, K.C., M.Th. van Genuchten, R. Schulin, and E. Schläppi. 1997. A sequential uncertainty domain inverse procedure for estimating subsurface flow and transport parameters. Water Resour. Res. 33:1879–1892.

Abbaspour, K.C., R. Schulin, and M.Th. van Genuchten. 2001. Estimating unsaturated soil hydraulic parameters using ant colony optimization. Adv. Water Resour. 24:827–841.

Allen, R.G., L.S. Perreira, D. Raes, and M. Smith. 1998. Crop evapotranspiration. FAO Irrig. Drain. Pap. 56. FAO, Rome.

Anghel, I., H.J. Turin, and P.W. Reimus. 2002. Lithium sorption to Yucca Mountain tuffs. Appl. Geochem. 17:819–824.

Beven, K., D. Zhang, and A. Mermoud. 2006. On the value of local measurements for prediction of pesticide transport at the field scale. Vadose Zone J. 5:222–233.

Birdsell, K.H., A.V. Wolfsberg, D. Hollis, T.A. Cherry, and K.M. Bower. 2002. Groundwater flow and radionuclide transport calculations for a performance assessment of a low-level waste site. J. Contam. Hydrol. 46:99–129.

Bitterlich, S, W. Durner, S.C. Iden, and P. Knabner. 2004. Inverse estimation of the unsaturated soil hydraulic properties from column outflow experiments using free-form parameterizations. Vadose Zone J. 3:971–981.

Brooks, R.H., and A.T. Corey. 1964. Hydraulic properties of porous media. Hydrol. Pap. 3. Colorado State Univ., Ft. Collins.

Broxton, D.E., and D.T. Vaniman. 2005. Geologic framework of a groundwater system on the margin of a rift basin, Pajarito Plateau, north-central New Mexico. Vadose Zone J. 4:522–550.

Carrera, J., and S.P. Neuman. 1986. Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability, and solution algorithms. Water Resour. Res. 22:211–227.

Celia, M.A., E.T. Bouloutas, and R.L. Zarba. 1990. A general mass-conservative numerical solution for the unsaturated flow equation. Water Resour. Res. 26:1483–1496.

Ciollaro, G., and N. Romano. 1995. Spatial variability of the soil hydraulic properties of a volcanic soil. Geoderma 65:263–282.

Clausnitzer, V., and J.W. Hopmans. 1995. Non-linear parameter estimation: LM_OPT. General-purpose optimization code based on the Levenberg–Marquardt algorithm. Land, Air and Water Resour. Pap. 100032. Univ. of California, Davis.

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

Dagan, G., and V. Cvetkovic. 1996. Reactive transport and immiscible flow in geological media: 1. General theory. Proc. R. Soc. London Ser. A 452:285–301.

Dane, J.H., and S. Hruska. 1983. In situ determination of soil hydraulic properties during drainage. Soil Sci. Soc. Am. J. 47:619–624.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. 6:182–197.

Destouni, G., and V. Cvetkovic. 1991. Field scale mass arrival of sorptive solute into the groundwater. Water Resour. Res. 27:1315–1325.

Doherty, J. 2004. PEST model-independent parameter estimation. User manual.

5th ed. Watermark Numerical Computing, Brisbane, Australia.

Durner, W., E.B. Schultze, and T. Zurmühl. 1999. State-of-the-art in inverse modeling of inflow/outflow experiments. p. 661–681. In M.Th. van Genuchten et al. (ed.) Characterization and measurement of the hydraulic properties of unsaturated porous media. Proc. Int. Worksh., Riverside, CA. 22–24 Oct. 1997. U.S. Salinity Lab., Riverside, CA.

Eching, S.O., and J.W. Hopmans. 1993. Optimization of hydraulic functions from transient outflow and soil water pressure data. Soil Sci. Soc. Am. J. 57:1167–1175.

Eching, S.O., J.W. Hopmans, and O. Wendroth. 1994. Unsaturated hydraulic conductivity from transient multi-step outflow and soil water pressure data. Soil Sci. Soc. Am. J. 58:687–695.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99:10143–10162.

Fen, C.S., and L.M. Abriola. 2004. A comparison of mathematical formulations for organic vapor transport in porous media. Adv. Water Resour. 27:1005–1016.

Fernández, J., M. Anguita, S. Mota, A. Cañaz, E. Ortigosa, and F.J. Rojas. 2004. MPI toolbox for Octave. In VecPar'04, Valencia, Spain. 28–30 June 2004. Available at atc.ugr.es/~javier/investigacion/papers/VecPar04.pdf (verified 10 Apr. 2008) Dep. de Arquitectura y Tecnología de Computadores, Univ. de Granada, Granada, Spain.

Fetter, C.W. 1999. Contaminant hydrogeology. Prentice Hall, Upper Saddle River, NJ.

Furman, A., T.P.A. Ferre, and A.W. Warrick. 2004. Optimization of ERT surveys for monitoring transient hydrological events using perturbation sensitivity and genetic algorithms. Vadose Zone J. 3:1230–1239.

Ginn, T.R. 2001. Stochastic–convective transport with nonlinear reactions and mixing: Finite streamtube ensemble formulation for multi-component reaction systems with intra-streamtube dispersion. J. Contam. Hydrol. 47:1–28.

Griggs, R.L. 1964. Geology and groundwater resources of the Los Alamos area, New Mexico. Water-Supply Pap. 1753. USGS, Reston, VA.

Gupta, H.V., S. Sorooshian, and P.O. Yapo. 1998. Towards improved calibration of hydrologic models: Multiple and non-commensurable measures of information. Water Resour. Res. 34:751–763.

Haario, H., E. Saksman, and J. Tamminen. 2001. An adaptive Metropolis algorithm. Bernoulli 7:223–242.

Hammond, G.E., A.J. Valocchi, and P.C. Lichtner. 2005. Application of Jacobian-free Newton–Krylov with physics-based preconditioning to biogeochemical transport. Adv. Water Resour. 28:359–376.

Heimovaara, T.J., J.A. Huisman, J.A. Vrugt, and W. Bouten. 2004. Obtaining the spatial distribution of water content along a TDR probe using the SCEM-UA Bayesian inverse modeling scheme. Vadose Zone J. 3:1128–1145.

Hoeg, S., H.F. Scholer, and J. Warantz. 2004. Assessment of interfacial mass transfer in water-unsaturated soils during vapor extraction. J. Contam. Hydrol. 74:163–195.

Hoeting, J.A., D. Madigan, A.E. Raftery, and C.T. Volinsky. 1999. Bayesian model averaging: A tutorial. Stat. Sci. 14:382–417.

Hollenbeck, K.J., and K.H. Jensen. 1998. Maximum-likelihood estimation of unsaturated soil hydraulic parameters. J. Hydrol. 210:192–205.

Hopmans, J.W., and J. Šimůnek. 1999. Review of inverse estimation of soil hydraulic properties. p. 643–659. In M.Th. van Genuchten et al. (ed.) Characterization and measurement of the hydraulic properties of unsaturated porous media. Proc. Int. Worksh., Riverside, CA. 22–24 Oct. 1997. U.S. Salinity Lab., Riverside, CA.

Hopmans, J.W., J. Šimůnek, N. Romano, and W. Durner. 2002. Simultaneous determination of water transmission and retention properties: Inverse methods. p. 963–1008. In J.H. Dane and G.C. Topp (ed.) Methods of soil analysis. Part 4. Physical methods. SSSA Book Ser. 5. SSSA, Madison, WI.

Hopmans, J.W., K.L. Bristow, and J. Šimůnek. 2001. Indirect estimation of soil thermal properties and water flux from heat pulse measurements: Geometry and dispersion effects. Water Resour. Res. 38:1–14.

Huisman, J.A., W. Bouten, J.A. Vrugt, and P.A. Ferré. 2004. Accuracy of frequency domain analysis scenarios for the determination of complex dielectric permittivity. Water Resour. Res. 40:W02401, doi:10.1029/2002WR001601.

Hupet, F., S. Lambot, M. Javaux, and M. Vanclooster. 2002. On the identification of macroscopic root water uptake parameters from soil water content observations. Water Resour. Res. 38(12):1300, doi:10.1029/2002WR001556.

Iden, S.C., and W. Durner. 2007. Free-form estimation of the unsaturated soil hydraulic properties by inverse modeling using global optimization. Water Resour. Res. 43:W07451, doi:10.1029/2006WR005845.

Inoue, M., J. Šimůnek, J.W. Hopmans, and V. Clausnitzer. 1998. In situ estimation of soil hydraulic functions using a multistep soil-water extraction technique. Water Resour. Res. 34:1035–1050.

Inoue, M., J. Šimůnek, S. Shiozawa, and J.W. Hopmans. 2000. Simultaneous estimation of soil hydraulic and solute transport parameters from transient infiltration experiments. Adv. Water Resour. 23:677–688.

Izett, G.A., and J.D. Obradovich. 1994. $^{40}Ar/^{39}Ar$ age constraints for the Jaramillo Normal Subchron and the Matuyama–Brunhes geomagnetic boundary. J. Geophys. Res. 99(B2):2925–2934.

Jury, W.A., D. Russo, G. Streile, and H. Abd. 1990. Evaluation of volatilization by organic chemicals residing below the soil surface. Water Resour. Res. 26:13–20.

Katul, G.G., O. Wendroth, M.B. Parlange, C.E. Puente, M.V. Folegatti, and D.R. Nielsen. 1993. Estimation of in situ hydraulic conductivity function from nonlinear filtering theory. Water Resour. Res. 29:1063–1070.

Kennedy, J., R.C. Eberhart, and Y. Shi. 2001. Swarm intelligence. Morgan Kaufmann, San Francisco.

Kool, J.B., and J.C. Parker. 1988. Analysis of the inverse problem for transient unsaturated flow. Water Resour. Res. 24:817–830.

Kool, J.B., J.C. Parker, and M.Th. van Genuchten. 1985. Determining soil hydraulic properties for one-step outflow experiments by parameter estimation: I. Theory and numerical studies. Soil Sci. Soc. Am. J. 49:1348–1354.

Kosugi, K. 1996. Lognormal distribution model for unsaturated soil hydraulic properties. Water Resour. Res. 32:2697–2703.

Kosugi, K. 1999. General model for unsaturated hydraulic conductivity for soils with lognormal pore-size distribution. Soil Sci. Soc. Am. J. 63:270–277.

Kwicklis, E.M., A.V. Wolfsberg, P.H. Stauffer, M.A. Walvroord, and M.J. Sully. 2006. Multiphase multicomponent parameter estimation for liquid and vapor fluxes in deep arid systems using hydrologic data and natural environmental traces. Vadose Zone J. 5:934–950.

LAM Team. 2006. MPI tutorials. Available at www.lam-mpi.org/tutorials/ (verified 12 Apr. 2008). Indiana Univ., Bloomington.

Lambot, S., M. Javaux, F. Hupet, and M. Vanclooster. 2002. A global multilevel coordinate search procedure for estimating the unsaturated soil hydraulic properties. Water Resour. Res. 38(11):1224, doi:10.1029/2001WR001224.

Lehr, J.H. 2004. Wiley's remediation technologies handbook: Major contaminant chemicals and chemical groups. John Wiley & Sons, Hoboken, NJ.

Marquardt, D.W. 1963. An algorithm for least squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. 11:431–441.

McLaughlin, D., and L.R. Townley. 1996. A reassessment of the groundwater inverse problem. Water Resour. Res. 32:1131–1161.

McLin, S.G., B.D. Newman, and D.E. Broxton. 2005. Vadose zone characterization and monitoring beneath waste disposal pits using horizontal boreholes. Vadose Zone J. 4:774–788.

Mertens, J., H. Madsen, L. Feyen, D. Jacques, and J. Feyen. 2004. Including prior information in the estimation of effective soil parameters in unsaturated zone modeling. J. Hydrol. 294:251–269.

Mertens, J., H. Madsen, M. Kristensen, D. Jacques, and J. Feyen. 2005. Sensitivity of soil parameters in unsaturated zone modeling and the relation between effective, laboratory and in situ estimates. Hydrol. Processes 19:1611–1633.

Mertens, J., R. Stenger, and G.F. Barkle. 2006. Multiobjective inverse modeling for soil parameter estimation and model verification. Vadose Zone J. 5:917–933.

Mualem, Y.A. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 12:513–522.

Neeper, D.A. 2002. Investigation of the vadose zone using barometric pressure cycles. J. Contam. Hydrol. 54:59–80.

Neeper, D.A., and R.H. Gilkeson. 1996. The influence of topography, stratigraphy, and barometric venting on the hydrology of unsaturated Bandelier Tuff. p. 427–443. In F. Goff et al. (ed.) Jemez Mountains Region. Guidebook 47. New Mexico Geol. Soc., Socorro.

Neeper, D.A., and P.H. Stauffer. 2005. Unidirectional gas flow in soil porosity resulting from barometric pressure cycles. J. Contam. Hydrol. 78:281–289.

Neupaier, R.M., B. Borchers, and J.L. Wilson. 2000. Comparison of inverse methods for reconstructing the release history of a groundwater contamination source. Water Resour. Res. 36:2469–2476.

Neuman, S.P. 1973. Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty. Water Resour. Res. 9:1006–1021.

Neuman, S.P. 2003. Maximum likelihood Bayesian averaging of alternative conceptual–mathematical models. Stochastic Environ. Res. Risk Assess. 17:291–305.

Pan, L.H., and L.S. Wu. 1998. A hybrid global optimization method for inverse estimation of hydraulic parameters: Annealing-simplex method. Water Resour. Res. 34:2261–2269.

Parker, J.C., J.B. Kool, and M.Th. van Genuchten. 1985. Determining soil hydraulic properties from one-step outflow experiments by parameter estimation: II. Experimental studies. Soil Sci. Soc. Am. J. 49:1354–1359.

Parlange, M.B., G.G. Katul, M.V. Folegatti, and D.R. Nielsen. 1993. Evaporation and the field scale soil water diffusivity function. Water Resour. Res. 29:1279–1286.

Pawar, R.J., and G.A. Zyvoloski. 2006. A novel method to couple wellbore flow to reservoir flow. p. 1–8. In P.J. Binning et al. (ed.) Proc. Int. Conf. on Computational Methods in Water Resour., 16th, Copenhagen, Denmark. 19–22 June 2006. Available at proceedings.cmwr-xvi.org/getFile.py/access?contribId=305&sessionId=10&resId=0&materialId=paper&confId=a051 (verified 23 Apr. 2008). Technical Univ. of Denmark, Lyngby.

Raftery, A.E., T. Gneiting, F. Balabdaoui, and M. Polakowski. 2005. Using Bayesian model averaging to calibrate forecast ensembles. Mon. Weather Rev. 133:1155–1174.

Rainwater, K.A., W.R. Wise, and R.J. Charbeneau. 1987. Parameter estimation through groundwater tracer tests. Water Resour. Res. 23:1901–1910.

Reimus, P.W., M.J. Haga, A.I. Adams, T.J. Callahan, H.J. Turin, and D.A. Counce. 2003. Testing and parameterizing a conceptual solute transport model in saturated fractured tuff using sorbing and nonsorbing tracers in cross-hole tracer tests. J. Contam. Hydrol. 62–63:613–636.

Robinson, B.A., and H.S. Viswanathan. 2003. Application of the theory of micromixing to groundwater reactive transport models. Water Resour. Res. 39(11):1313, doi:10.1029/2003WR002368.

Romano, N., and A. Santini. 1999. Determining soil hydraulic functions from evaporation experiments by a parameter estimation approach: Experimental verifications and numerical studies. Water Resour. Res. 35:3343–3359.

Rossi, C., and J.R. Nimmo. 1994. Modeling of soil-water retention from saturation to oven dryness. Water Resour. Res. 30: 701–708.

Rucker, D.F., and T.P.A. Ferre. 2005. Automated water content reconstruction of zero-offset borehole ground penetrating radar data using simulated annealing. J. Hydrol. 309:1–16.

Russo, D. 1988. Determining soil hydraulic properties by parameter estimation: On the selection of a model for the hydraulic properties. Water Resour. Res. 24:453–459.

Santini, A., N. Romano, G. Ciollaro, and V. Comegna. 1995. Evaluation of a laboratory inverse method for determining unsaturated hydraulic properties of a soil under different tillage practices. Soil Sci. 160:340–351.

Schoups, G.H., J.W. Hopmans, C.A. Young, J.A. Vrugt, and W.W. Wallender. 2005a. Multi-criteria optimization of a regional spatially-distributed subsurface water flow model. J. Hydrol. 311:20–48.

Schoups, G.H., J.W. Hopmans, C.A. Young, J.A. Vrugt, and W.W. Wallender. 2005b. Sustainability of irrigated agriculture in the San Joaquin Valley, California. Proc. Natl. Acad. Sci. USA 102:15352–15356.

Simmons, C.S., T.R. Ginn, and B.D. Wood. 1995. Stochastic–convective transport with nonlinear reaction: Mathematical framework. Water Resour. Res. 31:2675–2688.

Šimůnek, J., and M.Th. van Genuchten. 1996. Estimating unsaturated soil hydraulic properties from tension disc infiltrometer data by numerical inversion. Water Resour. Res. 32:2683–2696.

Šimůnek, J., and M.Th. van Genuchten. 1997. Estimating unsaturated soil hydraulic properties from multiple tension disc infiltrometer data. Soil Sci. 162:383–398.

Šimůnek, J., N.J. Jarvis, M.Th. van Genuchten, and A. Gärdenäs. 2003. Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone. J. Hydrol. 272:14–35.

Šimůnek, J., M. Šejna, and M.Th. van Genuchten. 1998a. The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Version 2.0. IGWMC-TPS-70. Int. Ground Water Modeling Ctr., Colorado School of Mines, Golden.

Šimůnek, J., O. Wendroth, and M.Th. van Genuchten. 1998b. A parameter estimation analysis of the evaporation method for determining soil hydraulic properties. Soil Sci. Soc. Am. J. 62:894–905.

Si, B.C., and W. Bodhinayake. 2005. Determining soil hydraulic properties from tension infiltrometer measurements. Soil Sci. Soc. Am. J. 69:1922–1930.

Springer, E.P. 2005. Statistical exploration of matrix hydrologic properties for the Bandelier Tuff, Los Alamos, New Mexico. Vadose Zone J. 4:505–521.

Stauffer, P.H., and N.D. Rosenberg. 2000. Vapor phase transport at a hillside landfill. Environ. Eng. Geosci. 6:71–84.

Stauffer, P.H., K.H. Birdsell, M.S. Witkowski, and J.K. Hopkins. 2005. Vadose zone transport of 1,1,1-trichloroethane: Conceptual model validation through numerical simulation. Vadose Zone J. 4:760–773.

Stauffer, P.H., L.H. Auer, and N.D. Rosenberg. 1997. Compressible gas in porous media: A finite amplitude analysis of natural convection. Int. J. Heat Mass Transfer 40:1585–1589.

Stauffer, P.H., J.K. Hopkins, and T. Anderson. 2007a. A soil vapor extraction pilot study in a deep arid vadose zone: Part 2. Simulations in support of decision making processes. *In* Global Accomplishments in Environmental and Radioactive Waste Management: Waste Manage. Symp., Tucson, AZ [CD-ROM]. 25 Feb.–1 Mar. 2007.

Stauffer, P.H., J.K. Hopkins, and T. Anderson. 2007b. Soil vapor extraction pilot test at Technical Area 54, Material Disposal Area L: Numerical modeling in support of decision analysis. LA-UR-07-4890. Los Alamos Natl. Lab., Los Alamos, NM.

Storn, R., and K. Price. 1997. Differential evolution: A fast and efficient heuristic for global optimization over continuous spaces. J. Global Optim. 11:341–359.

Takeshita, Y. 1999. Parameter estimation of unsaturated soil hydraulic properties from transient outflow experiments using genetic algorithms. p. 761–768. *In* M.Th. van Genuchten et al. (ed.) Characterization and measurement of the hydraulic properties of unsaturated porous media. Proc. Int. Worksh., Riverside, CA. 22–24 Oct. 1997. U.S. Salinity Lab., Riverside, CA.

Tonkin, M.J., and J. Doherty. 2005. A hybrid regularized inversion methodology for highly parameterized environmental models. Water Resour. Res. 41:W10412, doi:10.1029/2005WR003995.

Toorman, A.F., P.J. Wierenga, and R.G. Hills. 1992. Parameter estimation of soil hydraulic properties from one-step outflow data. Water Resour. Res. 28:3021–3028.

Unsal, E., J.H. Dane, and G.V. Dozier. 2005. A genetic algorithm for predicting pore geometry based on air permeability measurements. Vadose Zone J. 4:389–397.

Valiantzas, J.D., and D.G. Kerkides. 1990. A simple iterative method for the simultaneous determination of soil hydraulic properties from one-step outflow data. Water Resour. Res. 26:143–152.

Valstar, J.R., D.B. McLaughlin, C.B.M.T. Stroet, and F.C. van Geer. 2004. A representer-based inverse method for groundwater flow and transport applications. Water Resour. Res. 40(5):W05116, doi:10.1029/2003WR002922.

van Dam, J.C., J.N.M. Stricker, and P. Droogers. 1992. Evaluation of the inverse method for determining soil hydraulic functions from one-step outflow experiments. Soil Sci. Soc. Am. J. 56:1042–1050.

van Dam, J.C., J.N.M. Stricker, and P. Droogers. 1994. Inverse method for determining soil hydraulic functions from multi-step outflow experiments. Soil Sci. Soc. Am. J. 58:647–652.

van Genuchten, M.Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44:892–898.

Vereecken, H., R. Kasteel, J. Vanderborght, and T. Harter. 2007. Upscaling hydraulic properties and soil water flow processes in heterogeneous soils: A review. Vadose Zone J. 6:1–28.

Vrugt, J.A., and W. Bouten. 2002. Validity of first-order approximations to describe parameter uncertainty in soil hydrologic models. Soil Sci. Soc. Am. J. 66:1740–1752.

Vrugt, J.A., W. Bouten, H.V. Gupta, and J.W. Hopmans. 2003a. Toward improved identifiability of soil hydraulic parameters: On the selection of a suitable parametric model. Vadose Zone J. 2:98–113.

Vrugt, J.A., W. Bouten, H.V. Gupta, and S. Sorooshian. 2002. Toward improved identifiability of hydrologic model parameters: The information content of experimental data. Water Resour. Res. 38(12):312, doi:10.1029/2001WR001118.

Vrugt, J.A., M.P. Clark, C.G.H. Diks, Q. Duan, and B.A. Robin-

son. 2006a. Multi-objective calibration of forecast ensembles using Bayesian model averaging. Geophys. Res. Lett. 33:L19817, doi:10.1029/2006GL027126.

Vrugt, J.A., and J.H. Dane. 2005. Inverse modeling of soil hydraulic properties. p. 1103–1120. *In* M.G. Anderson et al. (ed.) Encyclopedia of hydrological sciences. John Wiley & Sons, Chichester, UK.

Vrugt, J.A., S.C. Dekker, and W. Bouten. 2003b. Identification of rainfall interception model parameters from measurements of throughfall and forest canopy storage. Water Resour. Res. 39(9):1251, doi:10.1029/2003WR002013.

Vrugt, J.A., C.G.H. Diks, H.V. Gupta, W. Bouten, and J.M. Verstraten. 2005a. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. Water Resour. Res. 41:W01017, doi:10.1029/2004WR003059.

Vrugt, J.A., H.V. Gupta, W. Bouten, L.A. Bastidas, and S. Sorooshian. 2003c. Effective and efficient algorithm for multi-objective optimization of hydrologic models. Water Resour. Res. 39(8):1214, doi:10.1029/2002WR001746.

Vrugt, J.A., H.V. Gupta, W. Bouten, and S. Sorooshian. 2003d. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour. Res. 39(8):1201, doi:10.1029/2002WR001642.

Vrugt, J.A., J.W. Hopmans, and J. Šimůnek. 2001a. Calibration of a two-dimensional root water uptake model. Soil Sci. Soc. Am. J. 65:1027–1037.

Vrugt, J.A., B.Ó. Nualláin, B.A. Robinson, W. Bouten, S.C. Dekker, and P.M.A. Sloot. 2006b. Application of parallel computing to stochastic parameter estimation in environmental models. Comput. Geosci. 32:1139–1155.

Vrugt, J.A., and B.A. Robinson. 2007a. Improved evolutionary optimization from genetically adaptive multimethod search. Proc. Natl. Acad. Sci. USA 104(3):708–711.

Vrugt, J.A., and B.A. Robinson. 2007b. Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging. Water Resour. Res. 43:W01411, doi:10.1029/2005WR004838.

Vrugt, J.A., B.A. Robinson, and J.M. Hyman. 2008. Self-adaptive multimethod search for global optimization in real parameter spaces. IEEE Trans. Evol. Comput. 12:(in press).

Vrugt, J.A., B.A. Robinson, and V.V. Vesselinov. 2005b. Improved inverse modeling for flow and transport in subsurface media: Combined parameter and state estimation. Geophys. Res. Lett. 32:L18408, doi:10.1029/2005GL023940.

Vrugt, J.A., G.H. Schoups, J.W. Hopmans, C.H. Young, W. Wallender, T. Harter, and W. Bouten. 2004. Inverse modeling of large-scale spatially distributed vadose zone properties using global optimization. Water Resour. Res. 40(6):W06503, doi:10.1029/2003WR002706.

Vrugt, J.A., M.T. van Wijk, J.W. Hopmans, and J. Šimůnek. 2001b. One-, two-, and three-dimensional root water uptake functions for transient modeling. Water Resour. Res. 37:2457–2470.

Vrugt, J.A., A.H. Weerts, and W. Bouten. 2001c. Information content of data for identifying soil hydraulic parameters from outflow experiments. Soil Sci. Soc. Am. J. 65:19–27.

Whisler, F.D., and K.K. Whatson. 1968. One-dimensional gravity drainage of uniform columns of porous materials. J. Hydrol. 6:277–296.

Wildenschild, D., J.W. Hopmans, and J. Šimůnek. 2001. Flow rate dependence of soil hydraulic characteristics. Soil Sci. Soc. Am. J. 65:35–48.

Winter, C.L., E.P. Springer, K. Costigan, P. Fasel, S. Mniszewski, and G. Zyvoloski. 2004. Virtual watersheds: Simulating the water balance of the Rio Grande basin. Comput. Sci. Eng. 6(3):18–26.

Wöhling, Th., J.A. Vrugt, and G.F. Barkle. 2008. Comparison of three multiobjective optimization algorithms for inverse modeling of vadose zone hydraulic parameters. Soil Sci. Soc. Am. J. 72:305–319.

Wolfsberg, A.V., and P.H. Stauffer. 2004. Vadose-zone fluid and solute flux: Advection and diffusion at the Area 5 Radioactive Waste Management Site. LA-UR-03–4819. Los Alamos Natl. Labl, Los Alamos, NM.

Wolpert, D., and W.G. Macready. 1999. No free lunch theorems for optimization. IEEE Trans. Evol. Comput. 1:67–82.

Wu, Y.S., K. Zhang, C. Ding, K. Pruess, E. Elmroth, and G.S. Bodvarsson. 2002. An efficient parallel-computing method for modeling nonisothermal multiphase flow and multicomponent transport in porous and fractured media. Adv. Water Resour. 25:243–261.

Ye, M., S.P. Neuman, and P.D. Meyer. 2004. Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. Water Resour. Res. 40(5):W05113, doi:10.1029/2003WR002557.

Ye, M., S.P. Neuman, P.D. Meyer, and K. Pohlmann. 2005. Sensitivity analysis and assessment of prior model probabilities in MLBMA with application to unsaturated fractured tuff. Water Resour. Res. 41:W12429, doi:10.1029/2005WR004260.

Zachmann, D.W., P.C. Duchateau, and A. Klute. 1981. The calibration of the Richards flow equation for a draining column by parameter identification. Soil Sci. Soc. Am. J. 45:1012–1015.

Zachmann, D.W., P.C. Duchateau, and A. Klute. 1982. Simultaneous approximation of water capacity and soil hydraulic conductivity by parameter identification. Soil Sci. 134:157–163.

Zeleke, T.B., and B.C. Si. 2005. Parameter estimation using the falling head infiltration model: Simulation and field experiment. Water Resour. Res. 41:W02027, doi:10.1029/2004WR003407.

Zhang, D., K.J. Beven, and A. Mermoud. 2006. A comparison of non-linear least square and GLUE for model calibration and uncertainty estimation for pesticide transport in soils. Adv. Water Resour. 29:1924–1933.

Zijlstra, J., and J.H. Dane. 1996. Identification of hydraulic parameters in layered soils based on a quasi-Newton method. J. Hydrol. 181:233–250.

Zurmühl, T., and W. Durner. 1998. Determination of parameters for bimodal hydraulic functions by inverse modeling. Soil Sci. Soc. Am. J. 62:874–880.

Zwietering, T.N. 1959. The degree of mixing in continuous flow systems. Chem. Eng. Sci. 11:1–15.

Zyvoloski, G.A., B.A. Robinson, and H.S. Viswanathan. 2008. Generalized dual porosity: A numerical method for representing spatially variable sub-grid processes. Adv. Water Resour. 31:535–544.

Zyvoloski, G.A., B.A. Robinson, Z.V. Dash, and L.L. Trease. 1997. Summary of the models and methods for the FEHM application: A finite element heat- and mass-transfer code. LA-13307-MS. Los Alamos Natl. Lab., Los Alamos, NM.